Contents

Acknowledgments

It's truly difficult to express by words the motions of a heart that throbs of joy: at the end of a path as intense as what I have experienced during my PhD, what prevails is a sense of deep and serene joy. The good climber knows that the beauty of the landscape offered by the highest peaks annihilates the efforts that led him up to there: similarly, at the end of any significant stage in our own life, what must prevail is the joy that comes from knowing that time has not been spent in vain and, before this awareness, the efforts and the inevitable renunciations completely eclipse.

But the joy never walks alone: the gratitude is her faithful companion. I wish to express my sincere thanks to all those who have guided, inspired and supported me: the list is quite long, but I hope I will not forget anyone.

I desire to address my deep gratitude to my advisor, prof. Beatrice Paternoster, for believing in me, helping me, encouraging and supporting my research: if I had the chance to do all my doctoral experiences, to weave all the international contacts I have by now and to refine myself as mathematician, it is because of her deep interest and support.

I wish to express my sincere thanks to my advisor, prof. Zdzislaw Jackiewicz, for having strongly believed in our research, especially during the year I have spent in Tempe, hosted by the Department of Mathematics and Statistics of the Arizona State University, to whom I address my thanks. I am also very grateful to his family, for the warm hospitality they have given me during my stay in Tempe, for the moments of joy and conviviality spent with them.

I sincerely thank Prof. Liviu Gr. Ixaru, Dr. Dajana Conte and Dr. Giuseppe Izzo, with whom I had the opportunity to closely collaborate in the scientific research. A warm and affectionate mention to Prof. John C. Butcher and all the scientists I have met during the conferences and the

summer schools I attended in these years, for the profitable discussions.

A sincere thank to Dr. Giovanni Capobianco, who advised me for the first time and believed in me since I was a very young student in Mathematics.

I would like to thank from the depth of my heart my colleagues of the office number 6 in Fisciano, who have been my traveling companions with whom I have shared many experiences and emotions.

Last, but not least, a deep gratitude to my beloved family who has patiently supported, guided, encouraged and comforted me, showing me each day the wonder of feeling loved. Thanks to my friends, my choir and my Gentle Light for their precious presence.

Chapter 1

Introduction

If we glance through the past decades, we can outright notice a remarkable increase of interest in the area of mathematical modelling as applied to science, engineering, business and management, generally expressed through functional equations, which provide the best and most natural way to describe evolution in time and space, also in presence of memory. In fact, the spread of diseases, the growth of biological populations, the brain dynamics, elasticity and plasticity, heat conduction, fluid dynamics, scattering theory, seismology, biomechanics, game theory, control, queuing theory, design of electronic filters and many other problems from physics, chemistry, pharmacology, medicine, economics can be modelled through systems of functional equations, such as Ordinary Differential Equations (ODEs) and Volterra Integral Equations (VIEs).

For instance, ODEs based models can be found in the context of evolution of biological populations [162, 207, 211], mathematical models in physiology and medicine [22], such as oncogenesis [136, 194] and spread of infections and diseases [143], economical sciences [104], analysis of signals [166]. Concerning VIEs based models, the following books and review papers contain sections devoted to this subject in the physical and biological sciences: Brunner [24, 27], Agarwal and O'Regan [3], Corduneanu and Sandberg [78], Zhao [211]. Most of these also include extensive lists of references. Regarding some specific applications of VIEs, they are for example models of population dynamics and spread of epidemics [124, 129], wave problems [109], fluido-dynamics [130], contact problems [1], electromagnetic signals [160].

The theoretical investigation concerning systems of ODEs and VIEs has

been very widely treated in the literature, especially in terms of existence, uniqueness and asymptotic stability of the solutions but, in most of the cases, it is not possible to analytically determine the solution of a functional equation: for this reason it gets more and more important to develop numerical methods in order to solve these problems and, moreover, due to the increasing complexity of such models, it is also necessary to require some special features on these methods, especially in terms of stability properties.

1.1 Problems and motivations

The purpose of this work is the construction, the theoretical analysis and the implementation of new efficient, accurate and highly stable numerical integration methods for the approximate solution of functional equations. We focus our interest on the following classes of problems:

1. Hadamard well-posed initial value problems based on systems of first order ordinary differential equations

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

with $f : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}^d$, even though we will often refer, without loss of generality, to the autonomous problem $y'(t) = g(y(t))$, with $g : \mathbb{R}^d \to \mathbb{R}^d$, in order to provide more compact notations. This is not a restriction, since whatever nonautonomous problem (1.1.1) can be regarded in the autonomous form

$$\begin{cases} u' = h(u), \\ u(t_0) = u_0 \in \mathbb{R}^{d+1}, \end{cases}$$

where $u = [y, t]^T$, $h(u) = [f(t, y), 1]^T$, $u_0 = [y_0, t_0]^T$;

2. Hadamard well-posed initial value problems based on systems of special second order ordinary differential equations

$$\begin{cases} y''(t) = f(t, y(t)), & t \in [t_0, T], \\ y'(t_0) = y'_0, \\ y(t_0) = y_0, \end{cases} \tag{1.1.2}$$

where $f : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}^d$, with periodic or oscillating solutions;

3. Volterra Integral Equations

$$y(t) = g(t) + \int_0^x k(t, \tau, y(\tau))d\tau \qquad t \in I := [0, T], \tag{1.1.3}$$

where $k \in C(D \times \mathbb{R}^d)$, with $D := \{(t, \tau) : 0 \leq \tau \leq t \leq T\}$ and $g \in C(I)$.

There are several specific reasons justifying our interest in the above functional equations.

Problem 1. The numerical approach to the problem (1.1.1) is a quite classical subject, which has been widely discussed in the literature (compare, for instance, the classical monographes [36, 123, 147, 186, 185]). However, the interest in this area is still vivid (see, for instance, the recent monographes [42, 122, 118, 119, 138, 105]) and the need to improve the methods already considered in the literature is remarkably felt by the recent literature, aiming for the development of building blocks for modern numerical solvers able to efficiently and accurately approach systems of ODEs modeling real phenomena. In fact, such equations generally exhibit typical problems (e.g. stiffness [122], metastability [122], periodicity [135], high oscillations [118, 105], discontinuities [2]) which must efficiently be overcome by using suitable numerical integrators. Part 1 of this dissertation is devoted to the development, the analysis and the implementation of new efficient numerical methods for the integration of the problem (1.1.1): our strengths in this context are devoted to the construction of highly stable continuous multistage integration methods able to efficiently approach stiff problems. The derived methods are compared with the ones classically considered in the literature, in order to show the advantages we have obtained. The theoretical analysis of the developed formulae is also supported by the experimental one, through the implementation of these methods in fixed and variable stepsize environment, obtained by exploiting the special structure of the derived numerical methods.

Problem 2. The approximate integration of the problem (1.1.2) has more recently been investigated in the literature. Although the problem (1.1.2) 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 seems more natural and efficient. Second order ODEs, having periodic or oscillatory solutions,

often appears in many applications: celestial mechanics, seismology, molecular dynamics, and so on (see for instance [178, 203] and references therein contained). Classical numerical methods for ODEs may not be well-suited to follow a prominent periodic or oscillatory behaviour of the solutions 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 (within the round-off error) problems of type (1.1.2) whose solution can be expressed as linear combination of functions other than polynomials. In the context of linear multistep methods for second order ODEs, Gautschi [110], Stiefel-Bettis [195] considered trigonometric functions depending on one or more frequencies, while Lyche [157] derived methods exactly integrating initial value problems of order r whose solution can be expressed as linear combination of powers and exponentials; Raptis-Allison [181] and Ixaru-Rizea [133] derived special purpose linear multistep methods for the numerical treatment of the radial Schrödinger equation $y'' = (V(t) - E)y$, 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 [53], Franco [107], Simos [145, 190] and Vanden Berghe [201], while their trigonometrically-fitted version has been developed by Paternoster in [170]; mixed-collocation based Runge–Kutta–Nyström methods have been introduced by Coleman and Duxbury in [64]. Recent adaptations of the Numerov method have been provided in [106, 128, 202]. For a more extensive bibliography see [135] and references within. 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 (1.1.2) in presence of periodicity and high oscillations in the solution. We also deal with the possibility to derive wider and more general classes of numerical methods for (1.1.2), inheriting the stability properties of the most stable numerical methods already considered in the literature.

Problem 3. The numerical treatment of Volterra integral equations (1.1.3) has been introduced in the literature in very recent times: the first monography [27] dealing with this subject has appeared in the literature only in 1986. Therefore, many open problems could still be investigated in this area. Part

3 of the thesis treats the construction and the theoretical analysis of highly stable continuous multistage methods which depend on structured coefficient matrices.

1.2 Some recent models

In this section we aim to present some of the most recent models of interest in the applied sciences, which involve the functional equations (1.1.1), (1.1.2) and (1.1.3). Of course, a widespread survey of interesting models involving these problems can be found in [42, 119, 122, 138] or in the Test Set website http://pitagora.dm.uniba.it/∼testset/.

We first consider some models involving systems of first order ODEs (1.1.1), of interest in Biology, Medicine and Cluster Analysis.

1. *Cell cycles.* Cell cycles, fundamental events in the life of every organism, are established by successive coordinated and oscillating steps which allow the cell to grow and duplicate correctly, strongly studied in correlation to tumour diseases. In [11] (see also the references therein contained), the description of such biological processes has been provided by a set of kinetic equations that define the biochemical reactions together with dynamic equations, structured as systems of first order ODEs (1.1.1) of the type

$$\frac{dX_i}{dt} = F_i(X_1, X_2, \ldots, X_n; p_1, p_2, \ldots, p_m), \quad i = 1, 2, \ldots, n,$$

where X_i, $i = 1, 2, \ldots, n$, is a state variable, generally describing the concentration of a certain species in the studied organism, each function F_i describes the rate of change of the corresponding X_i and p_j, $j = 1, 2, \ldots, m$ are parameters appearing in each F_i.

2. *Cancer growth.* Oncogenesis generally intends a cancer as the result of several mutations, giving some cells a selective growth advantage. In [194] the tumour growth is reduced to a simple set of rules according to which a normal cell becomes malignant. This point of view has spread out in the field of cancer research, which is very often treated as a logical science, trying to understand and describe the behaviour of the

disease through some underlying principles. This is the reason why the investigation can be done using mathematical tools: in this context, the evolutionary model is expressed through systems of ordinary differential equations that take into account the contribute of different elements to the cancer progression. The evolutionary model considered in [194] is the following

$$\begin{aligned}\frac{d\mathbf{y}}{dt} &= \operatorname{diag}(\operatorname{diag}(\mathbf{y}^{\mathbf{T}}\mathbf{k})^T\mathbf{b})\mathbf{M} + \operatorname{diag}\big((\mathbf{b}-\mathbf{d})^T\mathbf{y}\big) \\ &\cdot \quad \mathbf{S}(1 - a(\mathbf{y})\frac{P_{NM}}{10^6})(1 - \frac{P_{NM}}{10^{13}}) + \mathbf{m}_m,\end{aligned}$$

where $\mathbf{y} \in \mathbb{R}^{17}$ is a partitioned vector describing the whole cell populations: y_1 is related to normal cells, y_2, y_3 ... y_{15} describes the populations undergoing a mutation, y_{16} is the number of primary tumour cells and y_{17} is the number of metastatic cells. The vector $\mathbf{k}$ describes the mutation rates, i.e. k_i is the mutation rate for the population y_i. The vectors $\mathbf{b}$ and $\mathbf{d}$ respectively describe the birth and death rates, while $\mathbf{m}_m$ is the metastasis rate vector. The upper triangular matrix $\mathbf{M} \in \mathbb{R}^{17\times17}$ contains the number of genes going from state i to state j, while the matrix $\mathbf{S} \in \mathbb{R}^{17\times17}$ is related to non-normal, non-metastatic cells, i.e. cells undergoing mutations, which are denoted by P_{NM}, where $P_{NM} = \sum_{i=2}^{16} y_i$. The logistic term $a(\mathbf{y})$, defined as $a(\mathbf{y}) = \begin{cases} 0, & \frac{P_A}{P_{NM}} > 10\%, \\ 1 & \text{otherwise}, \end{cases}$ expires when more than the 10% of the non-normal, non-metastatic cells are in angiogenesis mutations (the number P_A denotes the amount of cells in this situation).

3. *A model in epileptic seizures.* Complex partial epileptic seizures are thought to originate from an area of focal abnormality in the brain, usually in the hippocampal portion of the temporal lobe. In [150] paper, a system of ODEs (1.1.1) is introduced to model for a hippocampal subnetwork believed to be of importance in the generation of focal, or complex partial, seizures. The variables in this model correspond to membrane potentials for prototypical pyramidal cells and inhibitory interneurons in the CA3 region of the hippocampus, the most likely

location of the focus. The model presented in [150] is the following:

$$\begin{aligned}
\frac{dV_i}{dt} &= -g_{Ca}m_\infty(V_i-1) - g_K W_i(V_i - V_i^K) - g_L(V_i - V_i^L) \\
&+ I - \alpha_{inh}Z_i, \\
\frac{dZ_i}{dt} &= b(cI + \alpha_{exc}V_i), \\
\frac{dW_i}{dt} &= \frac{\phi(w_\infty - W_i)}{\tau_w}, \qquad i = 1, 2, \ldots,
\end{aligned}$$

where V_i and Z_i are the membrane potentials of the pyramidal and inhibitory cells, respectively, while W_i is a relaxation factor which is essentially the fraction of open potassium channels in the population of pyramidal cells. The parameters g_{Ca}, g_K and g_L are the total conductances for the populations of Ca, K and leakage channels, respectively. V_i^K is the Nernst potential for potassium in node i; this parameter is be used in coupling subnetwork populations together into a lattice. V_i^L is a leak potential, τ_w is a voltage dependent time constant for W_i, I is the applied current, while ϕ and b are temperature scaling factors. The parameter c differentially modifies the current input to the inhibitory interneuron. The functions w_∞ and m_∞ are nondimensionalized expressions that describe the voltage-regulated Ca^{2+} channels in the cell membrane, each of which is either open or closed at any given moment. The parameters α_{exc} and α_{inh} model the populations of excitatory and inhibitory synaptic connections between pyramidal cells and their interneurons in the population of cells corresponding to the subnetwork model.

4. *"Stiff equations are multiscale problems".* This sentence, contained in the first pages of the paper [58] by J. R. Cash provides an intuitive idea of the nature of stiffness. There is an extensive bibliography regarding stiff problems: the reader can refer, for instance, to the monographes [36, 42, 122, 138, 147] and the references therein contained. Stiff equations often occur in the description of coupled physical systems having components which vary on very different time-scales: several examples which elaborate on this intuition can be found in [122]. When coupling together deterministic models of processes that occur on different scales or as part of different physical systems it is tempting simply to couple existing codes for the separate models to one another. However,

this does not take into account how the inaccuracies in the values of the variables that are passed between the two models are inherited. In order to prevent these inaccuracies from occurring we should consider the whole system as a single model rather than the combination of two simpler ones. When models are coupled together in this way, obviously the number of variables becomes very large. In addition, the range of length scales and time scales will also increase. The system of governing equations is now a large system modelling multi-scale processes. Equations representing such multi-scale processes are therefore particularly stiff (compare [193]). An important multi-scale problem in the life sciences is the simulation of a beating heart. Within the heart, several coupled physical processes occur at each level and there are complex feedback mechanisms between the scales. Further, almost all of the processes occur on multiple time scales. These properties make the development of an accurate whole heart simulator very difficult, both in terms of the formulation of appropriate models and the development of appropriate numerical algorithms for their solution. State-of-the-art cardiac cell models can include a large number of nonlinear ODEs: we only cite here, for the sake of brevity, the paper [161], where the reported model contains 23 coupled ODEs.

5. *A metastable problem: the Becker-Döring equation.* The Becker-Döring model [57] (see also the monography [122]) describes the dynamics of a large system of identical particles which can coagulate to form clusters. We denote by y_k the k-particle clusters per unit volume and assume that clusters can gain or lose single particles only. The following system of ODEs (1.1.1) arises:

$$\begin{aligned} \frac{dy_1}{dt} &= -J_1 - \sum_{k=1}^{N-1} J_k, \\ \frac{dy_k}{dt} &= J_{k-1} - J_k, k = 2, 3, \ldots, N-1, \\ \frac{dy_N}{dt} &= J_{N-1}, \end{aligned}$$

where $J_k = y_1 y_k - b_{k+1} y_{k+1}$ and $b_{k+1} = \exp(k^{2/3} - (k-1)^{2/3})$. This problem is particularly interesting because, especially for large values of N, the phenomenon of *metastability* occurs, i.e. extremely slow variations in the solution appear over long time intervals (see [57, 122]).

We next present some models involving systems of second order ODEs (1.1.2), of interest in Celestial Mechanics, Climatology, Chaos Theory, Ecology, Sensor Dynamics.

1. *The Pleiades problem.* The Pleiades problem (see, for instance, [119]) 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:
$$\begin{aligned} &z'' = f(z), \quad z \in \mathbb{R}^{14}, \\ &z(0) = z_0, \\ &z'(0) = z_0', \end{aligned}$$
where $z = [x, y]^T$, $x, y \in \mathbb{R}^7$, $f(z) = [f^{(1)}(z), f^{(2)}(z)]^T$, 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. *The interaction between climate and vegetation: a minimal model.* Today climatic change is a debated topic, especially in terms of the consequences that they lead to the environment, for instance to vegetation. A study aiming to describe the interaction of climate and vegetation has been carried out in [196], where the authors introduced the following model:
$$\begin{aligned} &vT' = D_T T'' + \Psi(N) - \sigma T^4, \\ &vN' = D_N N'' + \alpha(T) N - \gamma N^2, \end{aligned}$$
where the unknown functions $T(\xi)$ and $N(\xi)$ are oscillatory functions which exhibit a wave behaviour and, therefore, the variable ξ is of the type $x + vt$, where v stands for the (constant) velocity of the wave, $x \in \Omega$ is the point of the world area Ω we are observing and t is the time of the observation. In particular, $T(\xi)$ is the temperature function and $N(\xi)$ is the density of the vegetation in the area Ω. The function
$$\Psi(N) = \frac{\hat{S}}{C}(1 - N)$$

is an insolation term depending on the solar constant $\hat{S}$, while the term $\sigma^2 T^4$ is an irradiation term depending on the Stephan-Boltzmann constant σ. The function $\alpha(T)$ is the growth vegetation function, which depends only on the temperature: this function describe the ecological niche for vegetation, i.e. its behaviour in the space of the climatic factors.

3. *Oscillations and Chaos: the Duffing problem.* The Duffing equation, introduced for the first time by Duffing in [103], 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 [199]. 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 behaviour is known as deterministic chaos, or simply *chaos* (compare, for instance, [200]).

4. *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 (taller trees may spread the most propagules and have a reproductive advantage), forest mensuration (percent of height is used as spacing indicator) and site assessment (height-age, height-diameter relations can be used as indicators of site quality). An ex-

ample of model describing the time growth of trees height is given in [151], where the authors provided the analysis of the influence of the first year's height growth in predicting heights for later ages. The model discussed in [151] assumes the following form:

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

where the function h_{ijk} regards the total tree height 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 [151] the function g_{ijk} considered by the authors is such that

$$\dot{g}_{ijk} = c_0 + c_1V(t) + (a_0 + a_1V_2(t))\dot{h}_{ijk} + (b_0 + b_1V_3(t))h_{ijk},$$

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.

5. *A taste of chocolate: modelling dynamic flavour properties with second order ODEs.* Sensory properties such as taste, aroma and flavour are among the most important quality aspects of foods. Descriptive sensory analysis is closely related to classical psychophysics which uses different mathematical models to describe perceived intensities as functions of physical stimulus concentrations (see, for instance, Chavez-Birch, 1997; Kemp-Birch, 1992; Overbosch, 1986; Portmann, Serghat and Mathlouthi, 1992). Within the area of aroma and sensory perception, mathematical models has led to the formulation of ODEs (Harrison, Hills, Bakker and Clothier, 1997), which describe the flavour release from liquid emulsions and other closely related areas. In [144], the authors describe the time variation of a certain chocolate flavour by means of second order ODEs, with piecewise constant coefficients. Suppose that $I(t)$ denotes the intensity of the flavour at time t, then model suggested in [144] takes the form

$$I''(t) + b_i(u)I'(t) + a_i(u)I(t) = 0,$$

where the functions $a_i(u)$ and $b_i(u)$ depend on the parameter u denoting the chocolate content in percentage.

We finally provide some models consisting in second-kind Volterra Integral Equations (1.1.3) regarding Biology, Immunology and Population Dynamics.

1. *A model in the spread of infectious diseases.* Endemic infectious diseases for which infectious confers permanent immunity can be described by a system of nonlinear VIEs of convolution type [124, 129]. These models include vital dynamics (births and deaths), immunizations and distributed infectious period. In the papers [124, 129] the models are shown to be well posed, the threshold criteria are determined and the asymptotic behaviour is analyzed. The population is divided into disjoint classes which change with time t: $S(t)$, $I(t)$ and $R(t)$ represent the fractions of population that are susceptible, infectious and removed, respectively. Let us denote with β the constant contact rate (i.e. the average number of contacts, sufficient for transmission, of an infective per unit time), $P(t)$ the probability of remaining infectious t units after becoming infectious, $1/\mu$ the average lifetime, φ the fraction of immunized newborns, θ the rate of immunization of susceptibles, $\sigma = \beta \int_0^\infty P(t)e^{-\mu t}dt$ the average number of contacts of an infective during the infectious period, $(S_e, I_e, R_e) = \left(\frac{1}{\sigma}, \frac{(\sigma S^* - 1)(\mu+\theta)}{\beta}, 1 - S_e - R_e\right)$ the coordinates of the equilibrium point with $S^* = \frac{(1-\varphi)\mu}{\mu+\theta}$, R_0 the initial removed fractions and $I_0(t)e^{-\mu t}$ the fraction of the population that was initially infectious and is still alive and infectious at time t. The model assumes the following form

$$X(t) = F(t) + \int_0^t A(t-s)G(X(s))ds,$$

where $X(t) = [X_1(t), X_2(t)]^T$ and

$$A(\tau) = \begin{bmatrix} \beta P(\tau)e^{-\mu\tau} & 0 \\ \beta(1-P(\tau))e^{-\mu\tau} & -\theta e^{-\mu\tau} \end{bmatrix},$$

$$G(X) = \begin{bmatrix} S_e X_1 - I_e(X_1(t) + X_2) - X_1(X_1 + X_2) \\ X_1 + X_2 \end{bmatrix},$$

$$F(t) = \begin{bmatrix} I_0(t)e^{-\mu t} - \displaystyle\int_{-\infty}^{-t} \beta S_e I_e P(-s)e^{\mu s}ds \\ e^{-\mu t}\left(R_0 + I_0(0) - I_0(t) - \dfrac{\theta}{\sigma\mu} - \varphi\right) - \displaystyle\int_{-\infty}^{-t} \beta S_e I_e \left[1 - P(-s)\right] e^{\mu s}ds \end{bmatrix}.$$

2. *Vaccine-induced immune responses on HIV infection.* The need for anti-HIV-1 vaccines is universally recognized. Although several potential vaccine formulations are being tested in clinical trials, the complexity of the viral system and the length of the experimentation required and its costs makes the goal of obtaining such a vaccine still elusive. In [115], the authors have built a mathematical model for the simulation of HIV-1 infection spreading into the body, which allows to study the effect of hypothetical anti-HIV-1 vaccines having different properties. The model allows to predict which characteristics of immunogenicity a preventive or therapeutic vaccine should possess to be efficacious, and which are the key factors that most likely will affect its ability to control the spread of the infection. The model developed in [115] assumes the form

$$P(t) = \int_0^t F(t-x)e^{-\delta(t-x)}k_i k_s V(x)S(x)dx,$$

$$V(t) = V_0 e^{-ct} + \int_0^t e^{-c(t-x)}pP(x)dx,$$

$$S(t) = S_0 e^{-\beta t} + \int_0^t e^{-\beta(t-x)}[\alpha - k_s V(x)S(x)]dx,$$

where the unknown functions $P(t)$, $V(t)$ and $S(t)$ represent three populations of cells present in a unit volume of plasma at time t: $P(t)$ is the population of infected cells producing the virus, $V(t)$ is the virus population, $S(t)$ is the population of susceptibles cells.

3. *Age-structured populations: the Lotka-McKendrick model.* The theory of population dynamics has been extensively treated by mathematical demographers and population biologist. One of the most important features to include in realistic models for the dynamics of populations is certainly given by the effects of the age. The first continuous models incorporating the age effects have been proposed by Sharpe-Lotka in 1911 and McKendrick in 1926. The Lotka-McKendrick model for an age-structured population (compare [131]) assumes the form

$$B(t) = F(t) + \int_0^t \beta(x)B(t-x)\pi(x)dx,$$

where $B(t)$ describes the density of births at time t, $u(x,t)$ the age-specific density at time t, i.e. $u(x,t)dx$ is the total numbers of indi-

viduals of age between x and $x + dx$ at time t, $\pi(x)$ is the probability to survive from birth to age x and $\beta(x)$ is the rate of fertility for an individual of age x per unit time and the forcing function $F(t)$ takes the form

$$F(t) = \int_0^\omega \beta(x+t)u(x,0)\frac{\pi(x+t)}{\pi(x)}dx,$$

where ω is the maximal possible age.

Chapter 2

An unifying framework for the numerical integration of ODEs: General Linear Methods

2.1 General Linear Methods (GLMs)

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 [113] 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 [42]) 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 multivalue-multistage integration methods have been provided by J. Butcher from 1965 on (see [42] and references therein), Gear [111], Dahlquist [85], Donelson and Hansen [102], Byrne and Lambert [52], Jackiewicz and Tracogna [140]. In very recent times the first monography totally devoted to GLMs has also been released [138].

The representation formula we present has been introduced by Burrage and Butcher in [31] and then widely used in the context of GLMs (see [42, 119, 138] and references therein): it involves four coefficient matrices $\mathbf{A} \in \mathbb{R}^{s\times s}$, $\mathbf{U} \in \mathbb{R}^{s\times r}$, $\mathbf{B} \in \mathbb{R}^{r\times s}$, $\mathbf{V} \in \mathbb{R}^{r\times r}$, which are put together in the following partitioned $(s+r)\times(s+r)$ matrix

$$\left[\begin{array}{cc} \mathbf{A} & \mathbf{U} \\ \mathbf{B} & \mathbf{V} \end{array}\right].$$

We consider the uniform grid $\{t_n = t_0+nh,\ n = 0, 1, ..., N,\ Nh = T-t_0\}$, introduce the abscissa vector $\mathbf{c} = [c_1, c_2, \ldots, c_s]$ and define the following supervectors

$$y^{[n]} = \left[\begin{array}{c} y_1^{[n]} \\ y_2^{[n]} \\ \vdots \\ y_r^{[n]} \end{array}\right] \in \mathbb{R}^{rd}, \qquad F^{[n]} = \left[\begin{array}{c} f(t_n + c_1 h, Y_1^{[n]}) \\ f(t_n + c_2 h, Y_2^{[n]}) \\ \vdots \\ f(t_n + c_s h, Y_s^{[n]}) \end{array}\right] \in \mathbb{R}^{sd}.$$

Using these notations, a GLM $(\mathbf{c}, \mathbf{A}, \mathbf{U}, \mathbf{B}, \mathbf{V})$ can then be expressed as follows:

$$\left\{\begin{array}{l} Y^{[n]} = h(\mathbf{A}\otimes\mathbf{I})F + (\mathbf{U}\otimes\mathbf{I})y^{[n]}, \\ y^{[n+1]} = h(\mathbf{B}\otimes\mathbf{I})F + (\mathbf{V}\otimes\mathbf{I})y^{[n]}, \end{array}\right. \tag{2.1.1}$$

where $A\otimes B$ denotes the usual Kronecker tensor product and $\mathbf{I}$ is the identity matrix in $\mathbb{R}^{d\times d}$. Componentwise,

$$\left\{\begin{array}{ll} Y_i^{[n]} = \sum_{j=1}^{s} a_{ij} h F_j^{[n]} + \sum_{j=1}^{r} u_{ij} y_j^{[n]}, & i = 1, 2, \ldots, s, \\ y_i^{[n+1]} = \sum_{j=1}^{s} b_{ij} h F_j^{[n]} + \sum_{j=1}^{r} v_{ij} y_j^{[n]}, & i = 1, 2, \ldots, r. \end{array}\right. \tag{2.1.2}$$

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

General Linear Methods machinery is particularly useful in order to create an unifying approach to analyse the properties of a numerical method for ODEs, e.g. convergence, consistency and stability. We report in this subsection some definitions regarding the properties of GLMs, which are nothing more than all the 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 and, for this reason, we can decide to work with the GLM formulation of the method, instead of using the original one. We skip many details which can be found, for instance, in [42, 138], since we only aim to introduce the class of GLMs to make the dissertation self-contained and use the related main results in the following chapters for our specific purposes.

Definition 2.1.1 *A GLM* $(\mathbf{c}, \mathbf{A}, \mathbf{U}, \mathbf{B}, \mathbf{V})$ *is preconsistent if there exists a vector* $\mathbf{q}_0 \in \mathbb{R}^r$ *(named preconsistency vector) such that*

$$\mathbf{U}\mathbf{q}_0 = \mathbf{e}, \qquad \mathbf{V}\mathbf{q}_0 = \mathbf{q}_0, \tag{2.1.3}$$

where $\mathbf{e} = [1, \dots, 1] \in \mathbb{R}^r$.

Definition 2.1.2 *A preconsistent GLM* $(\mathbf{c}, \mathbf{A}, \mathbf{U}, \mathbf{B}, \mathbf{V})$ *is consistent if there exists a vector* $\mathbf{q}_1 \in \mathbb{R}^r$ *(named consistency vector) such that*

$$\mathbf{B}\mathbf{e} + \mathbf{V}\mathbf{q}_1 = \mathbf{q}_0 + \mathbf{q}_1, \tag{2.1.4}$$

with $\mathbf{e} = [1, \dots, 1] \in \mathbb{R}^r$*, while it is stage-consistent if*

$$\mathbf{A}\mathbf{e} + \mathbf{U}\mathbf{q}_1 = \mathbf{c}, \tag{2.1.5}$$

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

Butcher [42] 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. This idea can be

nicely explained by means of diagrams commutation. For this purpose, we consider two initial value problems

$$y'(t) = f(y(t)), \qquad y(t_0) = y_0, \tag{2.1.6}$$
$$z'(t) = f(z(t) - \eta), \qquad z(t_0) = y_0 + \eta, \tag{2.1.7}$$

for $x \in [x_0, \bar{x}]$, which differ each other for the presence of the perturbation $\eta \in \mathbb{R}^n$, representing a shift in the origin x_0. Assuming that (2.1.6) is a well-posed problem with solution $y(t)$, it is possible to prove that also (2.1.7) is well-posed and its solution is obtained translating each point by η on the trajectory described by $y(t)$. Of course, it is natural to require that a numerical method could preserve this property of the exact solution. We denote by ν the operation representing the computation of a numerical approximation to $y(t)$ in $[x_0, \bar{x}]$ and σ a shift of coordinates by η. Such a method is covariant if the diagram in Fig. 2.1 commutes, i.e. if $\sigma \circ \nu = \nu \circ \sigma$.

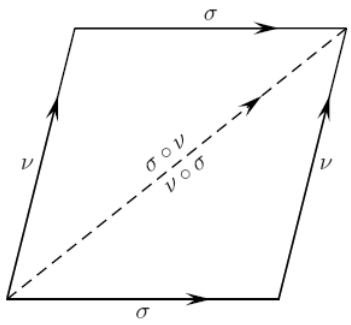

Figure 2.1: A commutative diagram for covariance (see [42])

Basic requirements in the context of the numerical integration of ODEs are, together with consistency, also zero-stability and convergence, which are defined for GLMs in the following way (see [42, 138]).

Definition 2.1.3 *A GLM* $(\mathbf{c}, \mathbf{A}, \mathbf{U}, \mathbf{B}, \mathbf{V})$ *is zero-stable if there exists a constant* C *such that*

$$\|\mathbf{V}^n\| \leq C, \tag{2.1.8}$$

for all $n \geq 0$.

A more practical way to analyze the zero-stability of a GLM arises from the following result (compare [36, 42, 138]).

Theorem 2.1.1 *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.*

Definition 2.1.4 *A GLM* $(\mathbf{c}, \mathbf{A}, \mathbf{U}, \mathbf{B}, \mathbf{V})$ *is convergent if for any initial value problem subject to the Lipschitz condition* $\|f(y) - f(z)\| \leq L\|y - z\|$, *there exist a nonzero vector* $u \in \mathbb{R}^r$ *and a starting procedure* $\phi : (0, \infty) \to \mathbb{R}^r$ *such that, for all* $i = 1, 2, \ldots, r$, $\lim_{h \to 0} \phi_i(h) = u_i y(t_0)$, *and such that for any* $\bar{x} > x_0$, *the sequence of vectors* $y^{[n]}$, *computed using* n *steps with stepsize* $h = (\bar{x} - x0)/n$ *and using* $y^{[0]} = \phi(h)$ *in each case, converges to* $uy(\bar{x})$.

Proving the convergence of a numerical method is generally a quite tedious and nontrivial task. However, a very helpful result in the convergence analysis of a finite difference approximation scheme is the Lax's equivalence theorem, also known as Lax-Richtmyer theorem [182]: this 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 linear multistep methods and Runge–Kutta methods (see, for instance, [36, 42, 147]) and, more in general, for a generic difference equation [182]. An analogous result has also been proved in the context of GLMs [42, 138] and can be stated as follows.

Theorem 2.1.2 *A GLM* $(\mathbf{c}, \mathbf{A}, \mathbf{U}, \mathbf{B}, \mathbf{V})$ *is convergent if and only if it is consistent and zero-stable.*

2.1.2 Order and stage-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 [5, 6, 7, 8], Butcher [32, 33, 36, 42, and references therein contained], Hairer and Wanner [120, 119, and references therein], using different approaches. We report in this section the main ideas concerning the derivation of order conditions for high stage order GLMs, which will next be useful in the remainder of this dissertation. The order

theory for GLMs has been developed by J. Butcher (see [36, 42] and references therein), by considering rooted trees associated to the numerical methods. However, in the case of high stage order methods, a different (but anyway related at the same time) approach to derive order conditions can be used. This approach has been discussed by Butcher himself in [39], in the context of diagonally implicit multistage integration methods (see [138]).

To formulate the stage order and order conditions for GLMs (2.1.2) we assume that the components of the input vector $y_i^{[n-1]}$ for the next step satisfy

$$y_i^{[n]} = \sum_{k=0}^{p} q_{ik} h^k y^{(k)}(t_n) + O(h^{p+1}), \quad i = 1, 2, \ldots, r, \tag{2.1.9}$$

for some real parameters q_{ik}, $i = 1, 2, \ldots, r$, $k = 0, 1, \ldots, p$. 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 (1.1.1) at the points $t_{n-1} + c_i h$, i.e.,

$$Y_i^{[n]} = y(t_n + c_i h) + O(h^{q+1}), \quad i = 1, 2, \ldots, r, \tag{2.1.10}$$

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

$$y_i^{[n+1]} = \sum_{k=0}^{p} q_{ik} h^k y^{(k)}(t_{n+1}) + O(h^{p+1}), \quad i = 1, 2, \ldots, r. \tag{2.1.11}$$

The integers q and p are called the stage order and order, respectively, of GLM (2.1.2). We collect the parameters q_{ik} appearing in (2.1.9) and (2.1.11) in the vectors $\mathbf{q}_k$ defined by

$$\mathbf{q}_k = \left[\begin{array}{cccc} q_{1k} & q_{2k} & \cdots & q_{rk} \end{array} \right]^T \in \mathbb{R}^r, \quad k = 0, 1, \ldots, p.$$

We also introduce the notation $e^{cz} = \left[\begin{array}{cccc} e^{c_1 z} & e^{c_2 z} & \cdots & c^{c_s z} \end{array} \right]^T$, 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 2.1.3 *(Butcher [39]). Assume that $y^{[n-1]}$ satisfies (2.1.9). Then the GLM (2.1.2) of order p and stage order $q = p$ satisfies (2.1.10) and (2.1.11) if and only if*

$$e^{cz} = z\mathbf{A}e^{cz} + \mathbf{U}\mathbf{w}(z) + O(z^{p+1}), \tag{2.1.12}$$

and

$$e^z \mathbf{w} = z\mathbf{B}e^{cz} + \mathbf{V}\mathbf{w}(z) + O(z^{p+1}). \tag{2.1.13}$$

Expanding e^{cz} and e^z in (2.1.12) and (2.1.13) 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 $\mathbf{e} = [1, \ldots, 1] \in \mathbb{R}^{s+2}$. 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

$$\frac{c^k}{k!} - \frac{\mathbf{A}c^{k-1}}{(k-1)!} - \mathbf{U}\mathbf{q}_k = 0, \quad k = 1, 2, \ldots, p, \tag{2.1.14}$$

and

$$\sum_{l=0}^{k} \frac{\mathbf{q}_{k-l}}{l!} - \frac{\mathbf{B}c^{k-1}}{(k-1)!} - \mathbf{V}\mathbf{q}_k = 0, \quad k = 1, 2, \ldots, p, \tag{2.1.15}$$

compare also [138]. In correspondence of $k = 1$, the stage-consistency and consistency conditions (2.1.5) and (2.1.4) result from (2.1.14) and (3.2.15) respectively.

2.1.3 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 are classically provided with respect to the scalar linear test equation

$$y' = \xi y, \quad t \geq 0, \tag{2.1.16}$$

where $\xi \in \mathbb{C}$ and $\mathrm{Re}(\xi) \leq 0$, considered for the first time by Dahlquist [82]. 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 (see [147]): 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 (2.1.1) to the linear test equation (2.1.16), 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

$$\mathbf{M}(z) = \mathbf{V} + z\mathbf{B}(\mathbf{I} - z\mathbf{A})^{-1}\mathbf{U}. \tag{2.1.17}$$

The characteristic polynomial of the stability matrix

$$p(\eta, z) = \det\big(\eta\mathbf{I} - \mathbf{M}(z)\big) \tag{2.1.18}$$

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 2.1.5 *A GLM (2.1.2) 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 (2.1.18) lie in the unit circle.*

Definition 2.1.6 *The region $\mathcal{A}$ of absolute stability of (2.1.2) is the set*

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

Definition 2.1.7 *The GLM (2.1.2) is said to be A-stable if its region of absolute stability includes the negative complex plane, i.e.*

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

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 (2.1.18) has to lie in the unit circle for all $z \in \mathbb{C}$ with $\mathrm{Re}(z) \leq 0$. 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 $p(\eta, iy)$ are in the unit circle for all $y \in \mathbb{R}$. This last condition can be investigated using the

Schur criterion [184] (see also [147]). This criterion, for a general k^{th} degree polynomial, can be formulated as follows. Consider the polynomial

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

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

$$\hat{\phi}(w) = \bar{d}_0 w^k + \bar{d}_1 w^{k-1} + \cdots + \bar{d}_{k-1} w + \bar{d}_k,$$

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

$$\phi_1(w) = \frac{1}{w}\Big(\hat{\phi}(0)\phi(w) - \phi(0)\hat{\phi}(w)\Big)$$

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

Theorem 2.1.4 *(Schur [184]). $\phi(w)$ is a Schur polynomial if and only if*

$$|\hat{\phi}(0)| > |\phi(0)|$$

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

Roughly speaking, the Schur criterion allows us to investigate the stability properties of a k^{th} degree polynomial, looking at the roots of a polynomial of lower degree (i.e. $k-1$). Iterating this process, the last step consists in the investigation of the root of a linear polynomial, plus some additional conditions.

Dahlquist introduced the concept of A-stability in his famous paper [82] of 1963, which has also been recently celebrated by J. Butcher [41]. When a numerical method is A-stable, there are no stability restriction on the stepsize in the implementation, which is a desirable property especially for the integration of stiff systems [36, 42, 81, 122, 147]. There is a stronger definition, which is required in order to damp the very stiff components of the numerical solution [36, 42, 122, 138, 147].

Definition 2.1.8 *A GLM (2.1.2) is said to be L-stable if it is A-stable and, in addition,*

$$\lim_{z\to\infty} \rho(\mathbf{M}(z)) = 0, \tag{2.1.19}$$

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 (2.1.18), tend to zero as $z \to -\infty$. Therefore, such methods satisfy the nonlinear system of equations

$$\begin{cases} \lim\limits_{z\to-\infty} \dfrac{p_0(z)}{p_r(z)} = 0, \\ \vdots \\ \lim\limits_{z\to-\infty} \dfrac{p_{r-1}(z)}{p_r(z)} = 0. \end{cases} \tag{2.1.20}$$

2.1.4 Nonlinear stability concepts

We have considered the concept of A-stability, which is based on the linear test equation (2.1.16). A-stable methods can reliably approach many stiff problems: however, additional difficulties can arise in the integration of nonlinear stiff problems. It is therefore reasonable to analyze the stability properties of numerical methods with respect to more difficult test equations: this is the foundation of the theory of nonlinear stability, whose father is G. Dahlquist, who introduced in 1976 the concept of G-stability [83], providing the starting point for many successive papers (e.g. [30, 31, 37, 38, 80, 84, 86]; for a more extensive bibliography see [42, 138] and references therein contained) concerning related nonlinear stability concepts. Also this work has been celebrated by J. Butcher in the review paper [40] of 2006.

We now report the main concepts concerning nonlinear stability of GLMs. Consider the initial-value problem

$$\begin{cases} y'(t) = g\big(t, y(t)\big), \quad t \geq 0, \\ y(0) = y_0, \end{cases} \tag{2.1.21}$$

$g : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^m$, where the function g satisfies the one-sided Lipschitz condition of the form

$$\big(g(t, y_1) - g(t, y_2)\big)^T (y_1 - y_2) \leq 0 \tag{2.1.22}$$

for all $t \geq 0$ and $y_1, y_2 \in \mathbb{R}^m$. Denote by $y(t)$ and $\tilde{y}(t)$ two solutions to (2.1.21) with initial conditions y_0 and $\tilde{y}_0$, respectively. Then it is known that

the condition (2.1.22) implies that

$$\big\|y(t_2) - \tilde{y}(t_2)\big\| \leq \big\|y(t_1) - \tilde{y}(t_1)\big\| \tag{2.1.23}$$

for $0 \leq t_1 \leq t_2$ (compare [40], [99]). Here, $\|\cdot\|$ is any norm in $\mathbb{R}^m$. The differential systems (2.1.21) with this property are called *dissipative*.

Let $\mathbf{G} = [g_{ij}]_{i,j=1}^r$ be a real, symmetric and positive definite matrix, and for a vector $y \in \mathbb{R}^{mr}$

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_r \end{bmatrix}, \quad y_i \in \mathbb{R}^m, \quad i = 1, 2, \ldots, r,$$

define the inner product norm $\|\cdot\|_G$ by

$$\|y\|_G^2 = \sum_{i=1}^r \sum_{j=1}^r g_{ij} y_i^T y_j. \tag{2.1.24}$$

Denote by $\{z^{[n]}\}_{n=0}^N$ the solution to (2.1.2) with initial value $z^{[0]}$, and by $\{\tilde{z}^{[n]}\}_{n=0}^N$ the solution obtained by perturbing (2.1.2) or by using a different initial value $\tilde{z}^{[0]}$. The numerical method which inherits the dissipativity property (2.1.23) of the solution $y(t)$ to (2.1.21) in the norm (2.1.24) is said to be *G-stable*. To be more precise, the GLM (2.1.2) is G-stable if there exists a real, symmetric and positive definite matrix $\mathbf{G} \in \mathbb{R}^{r\times r}$ such that for two numerical solutions $\{z^{[n]}\}_{n=0}^N$ and $\{\tilde{z}^{[n]}\}_{n=0}^N$ we have

$$\big\|z^{[n+1]} - \tilde{z}^{[n+1]}\big\|_G \leq \big\|z^{[n]} - \tilde{z}^{[n]}\big\|_G, \tag{2.1.25}$$

for all stepsizes $h > 0$ and for all differential systems (2.1.21) with the function g satisfying (2.1.22). The definition of G-stability is not of practical utility at all and, therefore, we look for more practical but equivalent concepts consisting, for instance, in a set of algebraic conditions involving the coefficient matrices of the methods we aim to analyze. Such conditions have been derived in the literature and lead to the concept of *algebraic stability*.

Definition 2.1.9 *The GLM (2.1.2) is said to be algebraically stable, if there exist a real, symmetric and positive definite matrix* $\mathbf{G} \in \mathbb{R}^{r\times r}$ *and a real, diagonal and positive definite matrix* $\mathbf{D} \in \mathbb{R}^{s\times s}$ *such that the matrix* $\mathbf{M} \in \mathbb{R}^{(s+r)\times(s+r)}$ *defined by*

$$\mathbf{M} = \left[\begin{array}{c|c} \mathbf{DA} + \mathbf{A}^T\mathbf{D} - \mathbf{B}^T\mathbf{GB} & \mathbf{DU} - \mathbf{B}^T\mathbf{GV} \\ \hline \mathbf{U}^T\mathbf{D} - \mathbf{V}^T\mathbf{GB} & \mathbf{G} - \mathbf{V}^T\mathbf{GV} \end{array}\right] \tag{2.1.26}$$

is nonnegative definite.

The significance of this definition follows from the result proved by Butcher [36], [38] (see also [122]), that for a preconsistent and non-confluent GLMs (2.1.2), i.e., methods with distinct abscissas c_i, $i = 1, 2, \ldots, s$, algebraic stability is equivalent to G-stability.

In general, it is quite difficult to verify if a given GLM is algebraically stable, and even more difficult to construct new classes of GLMs which are algebraically stable. In searching for such methods, it is worthwhile using the fact, proved in [31], that for a preconsistent and algebraically stable GLM (2.1.2) the matrices $\mathbf{G}$ and $\mathbf{D}$ are not independent but related by the equation

$$\mathbf{De} = \mathbf{B}^T\mathbf{Gq}_0, \tag{2.1.27}$$

where $\mathbf{q}_0$ is the preconsistency vector and $\mathbf{e} = [1, \ldots, 1]^T \in \mathbb{R}^s$. Moreover, $\mathbf{Gq}_0$ is a left eigenvector of the coefficient matrix $\mathbf{V}$ corresponding to the eigenvalue equal to one, i.e.,

$$(\mathbf{I} - \mathbf{V}^T)\mathbf{Gq}_0 = 0, \tag{2.1.28}$$

compare part ii) of Lemma 9.5 in [122].

We will write $\mathbf{M} \geq 0$ if the matrix $\mathbf{M}$ is nonnegative definite. It was observed by Hewitt and Hill [125], [126] that the verification if the matrix $\mathbf{M}$ is nonnegative definite can be simplified by the use of the result proved by Albert [4]. This result states that the matrix $\mathbf{M}$ given by

$$\mathbf{M} = \left[\begin{array}{c|c} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \hline \mathbf{M}_{12}^T & \mathbf{M}_{22} \end{array}\right]$$

satisfies $\mathbf{M} \geq 0$ if and only if

$$\mathbf{M}_{11} \geq 0, \quad \mathbf{M}_{22} - \mathbf{M}_{12}^T\mathbf{M}_{11}^+\mathbf{M}_{12} \geq 0, \quad \mathbf{M}_{11}\mathbf{M}_{11}^+\mathbf{M}_{12} = \mathbf{M}_{12}. \tag{2.1.29}$$

Here, $\mathbf{A}^+$ stands for the Moore-Penrose pseudoinverse of the matrix $\mathbf{A}$: we refer to [101] or [112] for the definition of this notion. In practice it may be more convenient to apply this result to the matrix $\tilde{\mathbf{M}}$ defined by

$$\tilde{\mathbf{M}} = \left[\begin{array}{c|c} \mathbf{M}_{22} & \mathbf{M}_{12}^T \\ \hline \mathbf{M}_{12} & \mathbf{M}_{11} \end{array}\right],$$

which is nonnegative definite if and only if $\mathbf{M}$ is nonnegative definite. Then $\mathbf{M} \geq 0$ if and only if

$$\mathbf{M}_{22} \geq 0, \quad \mathbf{M}_{11} - \mathbf{M}_{12}\mathbf{M}_{22}^{+}\mathbf{M}_{12}^{T} \geq 0, \quad \mathbf{M}_{22}\mathbf{M}_{22}^{+}\mathbf{M}_{12}^{T} = \mathbf{M}_{12}^{T}. \tag{2.1.30}$$

Although the criteria based on Albert theorem can be used to verify if specific examples of GLMs are algebraically stable, these criteria are not very practical to search for algebraically stable GLMs which depend on a number of unknown parameters. In such searches it is necessary to examine many inequalities which depend on the unknown coefficients of the matrix $\mathbf{G}$ and the remaining free parameters of GLMs and this task often exceeds the capabilities of symbolic manipulation packages such as Mathematica or Maple. However, there is a more practical approach, where this search can be done numerically, using the criterion for algebraic stability which is based on the Nyquist stability function defined by

$$\mathbf{N}(\xi) = \mathbf{A} + \mathbf{U}(\xi\mathbf{I} - \mathbf{V})^{-1}\mathbf{B}, \quad \xi \in \mathbb{C} - \sigma(\mathbf{V}). \tag{2.1.31}$$

Here, $\sigma(\mathbf{V})$ stands for the spectrum of the matrix $\mathbf{V}$. This terminology of the Nyquist stability function was suggested by Hill [127], although this function in the context of GLMs was first introduced by Butcher [38], who did not assign to it any specific name.

Denote by $\tilde{\mathbf{w}}$ a principal left eigenvector of $\mathbf{V}$, i.e., the vector such that

$$\tilde{\mathbf{w}}^{T}\mathbf{V} = \tilde{\mathbf{w}}^{T}, \quad \tilde{\mathbf{w}}^{T}\mathbf{q}_0 = 1, \tag{2.1.32}$$

where $\mathbf{q}_0$ is the preconsistency vector of GLM (2.1.2). Following [127] define the diagonal matrix $\tilde{\mathbf{D}}$ by

$$\tilde{\mathbf{D}} = \operatorname{diag}(\mathbf{B}^{T}\tilde{\mathbf{w}}), \tag{2.1.33}$$

and following [38], define by $\operatorname{He}(\mathbf{Q})$ the Hermitian part of a complex square matrix $\mathbf{Q}$, i.e.,

$$\operatorname{He}(\mathbf{Q}) = \frac{1}{2}\Big(\mathbf{Q} + \mathbf{Q}^{*}\Big),$$

where $\mathbf{Q}^{*}$ stands for the conjugate transpose of $\mathbf{Q}$. Then it was demonstrated in [38] and [127] that a consistent GLM (2.1.2) is algebraically stable if the following conditions are satisfied:

1. The coefficient matrix $\mathbf{V}$ is power-bounded.

2. $\mathbf{Ux} \neq \mathbf{0}$ for all right eigenvectors of $\mathbf{V}$ and $\mathbf{B}^T\mathbf{x} \neq \mathbf{0}$ for all left eigenvectors of $\mathbf{V}$.

3. $\tilde{\mathbf{D}} > 0$ and $\tilde{\mathbf{D}}\mathbf{A} \geq 0$.

4. $\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big) \geq 0$ for all ξ such that $|\xi| = 1$ and $\xi \in \mathbb{C} - \sigma(\mathbf{V})$.

2.2 A special class of GLMs: Two-Step Runge-Kutta Methods

In the context of GLMs we place special emphasis on the family of two-step Runge-Kutta (TSRK) methods

$$\begin{cases} y_{n+1} = \theta y_{n-1} + \tilde{\theta} y_n + h \sum_{j=1}^{s} \Big(v_j f(Y_j^{[n-1]}) + w_j f(Y_j^{[n]}) \Big), \\ Y_i^{[n]} = u_i y_{n-1} + \tilde{u}_i y_n + h \sum_{j=1}^{s} \Big(a_{ij} f(Y_j^{[n-1]}) + b_{ij} f(Y_j^{[n]}) \Big), \end{cases} \tag{2.2.1}$$

with $i = 1, 2, \ldots, s$, introduced by Jackiewicz and Tracogna [140] and further investigated by several authors (see, for instance, [14], [16], [17], [19], [48], [60], [121], [138] and references therein contained, [141], [142], [197], [198]). In (2.2.1), y_n is an approximation of order p to $y(t_n)$, $t_n = t_0 + nh$, and $Y_i^{[n]}$ are approximations of stage order q to $y(t_{n-1} + c_i h)$, $i = 1, 2, \ldots, s$, where $y(t)$ is the solution to (1.1.1) and $c = [c_1, \ldots, c_s]^T$ is the abscissa vector.

The peculiarity of two-step Runge-Kutta methods (2.2.1) 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 [140]), 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.

TSRK methods (2.2.1) can be represented by the abscissa vector c and

the table of their coefficients

$$\begin{array}{c|c|c} u & A & B \\ \hline \theta & v^T & w^T \end{array} = \begin{array}{c|ccc|ccc} u_1 & a_{11} & \cdots & a_{1s} & b_{11} & \cdots & b_{1s} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ u_s & a_{s1} & \cdots & a_{ss} & b_{s1} & \cdots & b_{ss} \\ \hline \theta & v_1 & \cdots & v_s & w_1 & \cdots & w_s \end{array},$$

and, moreover, they can also be regarded as GLMs (2.1.1) with $r = s + 2$ in correspondence of the vector of external approximations

$$y^{[n]} = \begin{bmatrix} y_n \\ y_{n-1} \\ hf(Y^{[n]}) \end{bmatrix} = \begin{bmatrix} y(t_n) \\ y(t_n - h) \\ y'(t_n + (c-e)h) \end{bmatrix} + O(h^{p+1}), \tag{2.2.2}$$

where $e = [1, \ldots, 1] \in \mathbb{R}^s$. Therefore, the Butcher tableau of TSRK methods re-casted as GLMs takes the form

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{U} \\ \hline \mathbf{B} & \mathbf{V} \end{array}\right] = \left[\begin{array}{c|ccc} B & e-u & u & A \\ \hline w^T & 1-\theta & \theta & v^T \\ 0 & 1 & 0 & 0 \\ I_s & 0 & 0 & 0 \end{array}\right] \in \mathbb{R}^{(2s+2)\times(2s+2)}. \tag{2.2.3}$$

We next report the main properties of TSRK methods in terms of order of convergence and linear stability: this analysis is carried out taking into account the formulation of TSRK methods as GLMs and, therefore, exploiting the results given in the general framework of GLMs.

We first aim to derive the set of order conditions for TSRK methods, under the hypothesis that we assume these methods to have *high stage order*, i.e. we consider methods of order p and stage order $q = p$. This choice is coherent with the spirit of the following sections, where we will consider continuous methods within the class of TSRK methods having *uniform* order of convergence $p = q$ over all the integration interval and, moreover, allows us to obtain the set of order conditions in a simple way: the only needed tool is the Taylor series expansion. The existing literature on TSRK methods does not always point particular importance to the request of *high stage order* and, as a consequence, order conditions are derived for any value of p and q, by using the Albrecht approach [140], or trees and B-series [48, 121].

In order to obtain the requested order conditions, we first need to determine the vectors $\mathbf{q}_k$, $k = 0, 1, \ldots, p$, appearing in (2.1.14) and (3.2.15). Taking into account the expression of the vector of external approximations $y^{[n]}$ given by (2.2.2), we expand $y(t_n - h)$ and $y'(t_n + (c - e)h)$ into Taylor series around the point t_n, obtaining

$$y^{[n]} = \begin{bmatrix} y(t_n) \\ y(t_n) - hy'(t_n) + \dfrac{h^2}{2!}y''(t_n) + \cdots + (-1)^p \dfrac{h^p}{p!}y^{(p)}(t_n) \\ hy'(t_n)e + h^2 y''(t_n)\dfrac{(c-e)}{1!} + \cdots + h^p y^{(p)}(t_n)\dfrac{(c-e)^{p-1}}{(p-1)!} \end{bmatrix} + O(h^{p+1}),$$

where the power $(c - e)^\nu$, $\nu = 0, 1, ..., p - 1$ is intended componentwise. Comparing this expression with (2.1.11) leads to the following formulas for the vectors $\mathbf{q}_k$

$$\mathbf{q}_0 = [1\ 1\ \mathbf{0}^T]^T, \quad \mathbf{q}_k = \begin{bmatrix} 0 & \dfrac{(-1)^k}{k!} & \left(\dfrac{(c-e)^{k-1}}{(k-1)!}\right)^T \end{bmatrix}^T,$$

$k = 1, 2, \ldots, p$, where $\mathbf{0}$ in $\mathbf{q}_0$ stands for the zero vector of dimension s. As a consequence, the preconsistency conditions (2.1.3) are automatically satisfied for TSRK methods. At this point, is much more convenient to express the stage order and order conditions (2.1.14) and (3.2.15) directly in terms of the coefficients c, θ, u, v, w, A, and B of the original TSRK methods. Theorem 2.1.3 implies the following result.

Theorem 2.2.1 *(compare [16], [140]) Assume that the TSRK method (2.2.1) has order p and stage order $q = p$. Then the order and stage order conditions take the form*

$$C_k := \frac{c^k}{k!} - \frac{(-1)^k}{k!}u - \frac{A(c-e)^{k-1}}{(k-1)!} - \frac{Bc^{k-1}}{(k-1)!} = 0, \tag{2.2.4}$$

$k = 1, , 2, \ldots, p$, *and*

$$\hat{C}_k := \frac{1}{k!} - \frac{(-1)^k}{k!}\theta - \frac{v^T(c-e)^{k-1}}{(k-1)!} - \frac{w^T c^{k-1}}{(k-1)!} = 0, \tag{2.2.5}$$

$k = 1, 2, \ldots, p$, *where* $c^\nu := [c_1^\nu, \ldots, c_s^\nu]^T$.

Proof: To reformulate (2.1.14) and (3.2.15) in terms of the coefficients of TSRK method (2.2.1) we use the representation of the matrices $\mathbf{A}$, $\mathbf{U}$, $\mathbf{B}$ and $\mathbf{V}$ given in (2.2.3). The stage order conditions (2.2.4) follow directly from (2.1.14). To reformulate (3.2.15) observe that

$$\sum_{l=0}^{k} \frac{\mathbf{q}_{k-l}}{l!} = \left[\begin{array}{cc} \frac{1}{k!} & \sum_{l=0}^{k} \frac{(-1)^l}{(k-l)!l!} \end{array} \quad \left(\sum_{l=0}^{k-1} \frac{(c-e)^l}{(k-1-l)!l!} \right)^T \right]^T.$$

Since

$$\sum_{l=0}^{k} \frac{(-1)^l}{(k-l)!l!} = \frac{1}{k!} \sum_{l=0}^{k} \left(\begin{array}{c} k \\ l \end{array} \right) (-1)^l = 0,$$

and

$$\sum_{l=0}^{k-1} \frac{(c-e)^l}{(k-1-l)!l!} = \frac{1}{(k-1)!} \sum_{l=0}^{k-1} \left(\begin{array}{c} k-1 \\ l \end{array} \right) (c-e)^l = \frac{c^{k-1}}{(k-1)!}$$

it follows that the last $s+1$ components of the left hand side of (3.2.15) are automatically equal to zero, and comparing the first components of (3.2.15) we obtain order conditions (2.2.5). This completes the proof. □

Setting $k = 1$ in (2.1.14) and (3.2.15), the stage consistency and consistency conditions (2.1.5) and (2.1.4) take the form $(A + B)e - u = c$, $(v^T + w^T)e = 1 + \theta$, respectively.

We next derive the conditions for zero-stability of TSRK methods, using the result in Theorem 2.1.1. For this purpose, we derive the minimal polynomial of the coefficient matrix $\mathbf{V}$ in (2.2.3), i.e.

$$p(\omega) = \omega(\omega^2 - (1-\theta)\omega - \theta),$$

whose roots are $\omega = 0$, $\omega = 1$ and $\omega = -\theta$. Therefore, a TSRK is zero-stable if and only if $-1 < \theta \leq 1$.

We finally focus our attention on the linear stability properties of TSRK methods. We have already provided the expression of the stability matrix (2.1.17) of a GLM: in the case of TSRK methods, the matrix (2.1.17) takes the form

$$\mathbf{M}(z) = \left[\begin{array}{ccc} 1 - \theta + w^T Q(z)(e-u) & \theta + w^T Q(z)u & v^T + w^T Q(z)A \\ 1 & 0 & 0 \\ Q(z)(e-u) & Q(z)u & Q(z)A \end{array} \right] \in \mathbb{R}^{(s+2)\times(s+2)}, \tag{2.2.6}$$

where $z \in \mathbb{C}$ and $Q(z) = (I - zB)^{-1}$. We also define the stability function $\tilde{p}(\omega, z)$ as the characteristic polynomial of $\mathbf{M}(z)$, i.e.,

$$\tilde{p}(\omega, z) = \det\Big(\omega I - \mathbf{M}(z)\Big). \tag{2.2.7}$$

This is a polynomial of degree $s + 2$ with respect to ω whose coefficients are rational functions with respect to z.

Many examples of highly-stable TSRK methods (2.2.1) can be found in [140, 138] and references therein contained.

Part I

Numerical solution of differential systems of the first order

Chapter 3

Two-step collocation and almost collocation methods

Collocation is a widely applied and powerful technique in the construction of numerical methods for functional equations. As it is well known, a collocation method is based on the idea of approximating the exact solution of a given functional equation with a suitable approximant, the *collocation function*, belonging to a chosen finite dimensional space, usually a piecewise algebraic polynomial, which exactly satisfies the equation on a certain subset of the integration interval (i.e. the set of the so-called *collocation points*).

This technique, when applied to problems based on functional equations, allows the derivation of methods having many desirable properties. In fact, collocation methods provide an approximation over the entire integration interval to the solution of the equation. Moreover, the collocation function can be expressed as linear combination of functions *ad hoc* for the problem we are integrating, in order to better reproduce the qualitative behaviour of the solution. The usage of the collocation polynomial is also particularly useful for the implementation of the related methods in a variable stepsize environment, as it will be discussed in Chapter 6. It is also worthwhile to mention that collocation has also an important theoretical relevance: in fact, many numerical methods are difficult to be analyzed as discrete schemes while, re-casted as collocation-based methods, their analysis is reasonably simplified and can be carried out in a very elegant way.

The systematic study of collocation methods for ODEs (1.1.1)-(1.1.2), VIEs (1.1.3), Volterra integro-differential equations (VIDEs) and related

functional equations has its origin, respectively, in the late '60s, the early '70s and the early '80s (an interesting historical review can be found, for instance, in the Notes 1.11 of the monography [24] on collocation methods by H. Brunner). The idea of multistep collocation was first introduced by Lie and Norsett in [154], and further extended and investigated by several authors [48, 70, 75, 76, 77, 88, 89, 90, 92, 116, 153, 158].

Multistep collocation methods depend on more parameters than classical ones, without any significant increase in the computational cost: therefore, there are much more degrees of freedom to be spent in order to obtain strong stability properties and higher order and stage order of convergence. As a direct consequence, the effective order of multistep collocation methods is generally higher with respect to one stage collocation methods with the same number of stages. Moreover, in force of their high stage order, they do not suffer from the order reduction phenomenon (see [42, 122]), which occurs in the integration of stiff systems.

This chapter consists of two sections: the first one reviews the classical idea of collocation for first order ODEs (1.1.1) and its successive modifications introduced in the literature, providing the background in which our new results are merged into and the framework which constituted the starting point of our research; the second part presents the new ideas of two-step collocation and almost collocation we have introduced and analyzed in [70, 88, 89, 92, 94, 95, 98], with special emphasis on the results concerning representation formulae for the coefficients of the corresponding methods, the analysis of the local discretization error, the related set of continuous order conditions, the attainable order of convergence, the linear stability properties.

3.1 Collocation based methods for first order ODEs: state of the art

In this section we focus our attention on the historical background and concerning the collocation technique: we will analyze the classical idea of collocation for the numerical integration of ODEs, which is placed in the context of Runge–Kutta methods, together with some famous extensions and modifications which have been developed in the literature during the

'80s and '90s.

3.1.1 Classical one-step collocation methods

Let us suppose that the integration interval $[t_0, T]$ is discretized in an uniform grid $\{t_h : t_0 < t_1 < ... < t_N = T\}$. Classical collocation methods (see [24, 36, 42, 118, 119, 147, 208]) are determined by means of a continuous approximant, generally an algebraic polynomial $P(t)$, satisfying some opportune conditions: in order to advance from t_n to t_{n+1}, the polynomial $P(t)$ interpolates the numerical solution in t_n, and exactly satisfies the ODE (1.1.1) - i.e. *co-locates* - in the set of points $\{t_n + c_i h,\ i = 1, 2, \dots m\}$, where $c_1, c_2, \dots, c_m$ are m real numbers (named *collocation nodes*), that is

$$\begin{cases} P(t_n) = y_n, \\ P'(t_n + c_i h) = f(t_n + c_i h, P(t_n + c_i h)), \quad i = 1, 2, ..., m. \end{cases} \tag{3.1.1}$$

The solution in t_{n+1} can then be computed from the function evaluation

$$y_{n+1} = P(t_{n+1}). \tag{3.1.2}$$

The classical framework in which collocation methods must be placed is certainly constituted by implicit Runge-Kutta methods (IRK). In fact, Guillou and Soule in [116] and Wright in [208] independently proved that one step collocation methods form a subset of implicit Runge-Kutta methods

$$y_{n+1} = y_n + h \sum_{i=1}^{m} b_i f(t_n + c_i h, Y_i) \tag{3.1.3}$$

$$Y_i = y_n + h \sum_{j=1}^{m} a_{ij} f(t_n + c_j h, Y_j), \qquad i = 1, 2, ..., m, \tag{3.1.4}$$

where

$$a_{ij} = \int_0^{c_i} L_j(s) ds, \qquad b_j = \int_0^1 L_j(s) ds, \qquad i, j = 1, 2, ..., m \tag{3.1.5}$$

and $L_j(s)$, $j = 1, ..., m$, are fundamental Lagrange polynomials. The maximum attainable order of such methods is at most $2m$, and it is obtained by using Gaussian collocation points [119, 147]. Anyway, unfortunately, the order

$2m$ is gained only at the mesh points: the uniform order of convergence over the entire integration interval is only m. As a consequence, they suffer from order reduction showing effective order equal to m (see [36, 42, 119, 122]).

Butcher (see [36] and references therein) gave an interesting characterization of collocation methods in terms of easy algebraic conditions, and analogous results are also reported in [119, 147]. This characterization, together with many other results regarding the main properties of collocation methods, comes out as natural consequence of an interesting interpretation of collocation methods in terms of quadrature formulae. In fact, if $f(t, y(t)) = f(t)$, equations (3.1.3)-(3.1.4) can be respectively interpreted as quadrature formulae for $\int_{t_n}^{t_n+h} f(t)dt$ and $\int_{t_n}^{t_n+c_i h} f(t)dt$, for i=1,2,...,m. We next consider the following linear systems

$$A(q) \quad : \quad \sum_{j=1}^{m} a_{ij} c_j^{k-1} = \frac{c_i^k}{k}, \qquad k = 1, 2, ..., q, \ i = 1, 2, ..., m, \tag{3.1.6}$$

$$B(p) \quad : \quad \sum_{i=1}^{m} b_i c_i^{k-1} = \frac{1}{k}, \qquad k = 1, 2, ..., p. \tag{3.1.7}$$

Next, the following result holds (see [118, 147]):

Theorem 3.1.1 *If the condition $B(p)$ holds for some $p \geq m$, then the collocation method (3.1.1) has order p.*

As a consequence, a collocation method has the same order of the underlying quadrature formula (see [118], p. 28). Finally, the following result characterizing classical collocation methods arises (see [36, 118, 119, 147]).

Theorem 3.1.2 *An implicit m-stage Runge-Kutta method, satisfying $B(m)$ and having distinct collocation abscissae, is a collocation method if and only if conditions $A(m)$ holds.*

The most used collocation methods are those based on the zeros of some orthogonal polynomials, that is Gauss, Radau, Lobatto [36, 42, 119, 122, 147], having respectively order of convergence $2m$, $2m-1$, $2m-2$, where m is the number of collocation points (or the number of stages, regarding the collocation method as an implicit Runge-Kutta). Concerning their stability properties, it is known that Runge-Kutta methods based on Gaussian collocation points are A-stable, while the ones based on Radau IIA points are L-stable

and, moreover, they are also both algebraically stable (see [42, 122, 138] and references therein contained); Runge-Kutta methods based on Lobatto IIIA collocation points, instead, are A-stable but they are not algebraically stable (see [36, 118, 119, 147]).

3.1.2 Perturbed collocation

As remarked by Hairer and Wanner in [122], only some IRK methods are of collocation type, i.e. Gauss, Radau IIA, and Lobatto IIIA methods. An extension of the collocation idea, the so-called *perturbed collocation* is due to Norsett and Wanner (see [164, 165]), which applies to all IRK methods.

We denote by Π_m the linear space of polynomials of degree at most m and consider the polynomial $N_j \in \Pi_m$ defined by

$$N_j(t) = \frac{1}{j!}\sum_{i=0}^{m}(p_{ij} - \delta_{ij})t^i, \quad j = 1, 2, ..., m,$$

where d_{ij} is the usual Kronecker delta. We next define the *perturbation operator* $P_{t_0,h} : \Pi_m \to \Pi_m$ by

$$(P_{t_0,h}u)(t) = u(t) + \sum_{j=1}^{n} N_j\left(\frac{t - t_0}{h}\right)u^{(j)}(t_0)h^j.$$

Next, the following definition is given (see [164, 165]).

Definition 3.1.1 *Let $c_1, \ldots, c_m$ be given distinct collocation points. Then the corresponding perturbed collocation method is defined by*

$$\begin{aligned} u(t_0) &= y_0, \quad u \in \Pi_m, \\ u'(t_0 + c_i h) &= f(t_0 + c_i h, (Pu)(t_0 + c_i h)), \quad i = 1, 2, ..., m, \\ y_1 &= u(t_0 + h). \end{aligned}$$

As the authors remark in [165], if all N_j's are identically zero, then P is the identical map and the definition coincides with classical collocation. In the same paper the authors provide the equivalence result between the family of perturbed collocation methods and Runge-Kutta methods (see [165]). The interest of this results, as again is stated in [165], is that the properties of collocation methods, especially in terms of order, linear and nonlinear stability, can be derived in a reasonable short, natural and very elegant way, while it is known that, in general, these properties are very difficult to handle and investigate outside collocation.

3.1.3 Discontinuous collocation

In the literature, perturbed collocation has been considered as a modification of the classical collocation technique, in such a way that much more Runge-Kutta methods could be regarded as perturbed collocation based methods, rather than classically collocation based. There are other possible extensions of the collocation idea, which apply to wider classes of Runge-Kutta methods, such as the so-called *discontinuous collocation* (see [118]).

Definition 3.1.2 *Let $c_2, \ldots, c_{m-1}$ be distinct real numbers (usually between 0 and 1), and let b_1, b_m be two arbitrary real numbers. The corresponding discontinuous method is then defined via a polynomial of degree $m-2$ satisfying*

$$\begin{aligned} u(t_0) &= y_0 - hb_1(\dot{u}(t_0) - f(t_0, u(t_0)), \\ \dot{u}(t_0 + c_i h) &= f(t_0 + c_i h, u(t_0 + c_i h)), \quad i = 2, \ldots, m-1, \\ y_1 &= u(t_1) - hb_s(\dot{u}(t_1) - f(t_1, u(t_1)). \end{aligned}$$

Discontinuous collocation methods fall inside a large class of implicit Runge-Kutta methods, as stated by the following result (see [118]).

Theorem 3.1.3 *The discontinuous collocation method given in Definition 3.1.2 is equivalent to an m-stage Runge-Kutta method with coefficients determined by $c_1 = 0$, $c_m = 1$ and*

$$a_{i1} = b_1, \qquad a_{im} = 0, \qquad i = 1, \ldots, m,$$

while the other coefficients result as solutions of the linear systems $A(m-2)$ and $B(m-2)$ defined in (3.1.6) and (3.1.7).

As a consequence of this result, if $b_1 = 0$ and $b_m = 0$, then the discontinuous collocation method in Definition 3.1.2 is equivalent to the $(m-2)$-collocation method based on $c_2, \ldots, c_{m-1}$. An interesting example of implicit Runge-Kutta method which is not collocation based but is of discontinuous collocation type is the Lobatto IIIB method (see [36, 118, 119, 147]), which plays an important rule in the context of geometric numerical integration, together with Lobatto IIIA method (see [118], p. 33). They are both nonsymplectic methods (see Theorem 4.3 in [118]) but, considered as a pair, the resulting method is symplectic. This is a nice example of methods which possess very strong properties, but are difficult to investigate as discrete scheme (they

cannot be studied as collocation methods, because they are not both collocation based); however, re-casted as discontinuous collocation based methods, their analysis is reasonably simplified and very elegant [118].

3.1.4 Multistep collocation

The successive results which appeared in the literature (see [68, 116, 122, 153, 154]) have been devoted to the construction of multistep collocation methods. Guillou and Soulé introduced multistep collocation methods [116], by adding interpolation conditions in the previous k step points, so that the collocation polynomial is defined by

$$\begin{cases} P(t_{n-i}) = y_{n-i} \quad i = 0, 1, ..., k-1, \\ P'(t_n + c_j h) = f(t_n + c_j h, P(t_n + c_j h)), \quad j = 1, ..., m. \end{cases} \tag{3.1.8}$$

The numerical solution is given, as usual, by the evaluation

$$y_{n+1} = P(t_{n+1}). \tag{3.1.9}$$

Hairer-Wanner [122] and Lie-Norsett [154] derived different strategies to obtain multistep collocation methods. In [122] the Hermite problem with incomplete data (3.1.8) is solved by means of the introduction of a generalized Lagrange basis

$$\{\varphi_i(s),\ \psi_j(s),\ i = 1, 2, \ldots, k,\ j = 1, 2, \ldots, m\}$$

and, correspondingly, the collocation polynomial is expressed as linear combination of this set of functions, i.e.

$$P(t_n + sh) = \sum_{j=1}^{k} \varphi_j(s) y_{n-k+j} + h \sum_{i=1}^{s} \psi_j(s) P'(t_n + c_j h),$$

where $s = \frac{t - t_n}{h}$ is the scaled time variable. Therefore, the problem (3.1.8) is transformed in the problem of deriving $\{\varphi_i(s),\ \psi_j(s),\ i = 1, 2, \ldots, k,\ j = 1, 2, \ldots, m\}$ in such a way that the corresponding polynomial $P(t_n + sh)$ satisfies the conditions (3.1.8).

Lie-Norsett in [154] completely characterized multistep collocation methods, giving the expressions of the coefficients of collocation based multistep Runge-Kutta methods in closed form, as stated by the following

Theorem 3.1.4 *The multistep collocation method (3.1.8)-(3.1.9) is equivalent to the multistep Runge-Kutta method*

$$\begin{aligned} Y_j &= \sum_{i=0}^{k-1} \varphi(t_{n+k-1} + c_j h) y_{n+k-1-i} \\ &+ h \sum_{i=1}^{m} \psi(t_{n+k-1} + c_j h) f(t_{n+k-1} + c_j h, Y_i), \qquad j = 1, ..., m, \end{aligned}$$

$$y_{n+k} = \sum_{i=0}^{k-1} \varphi(t_{n+k}) y_{n+k-1-i} + h \sum_{i=1}^{m} \psi(t_{n+k-1}) f(t_{n+k-1} + c_j h, Y_i),$$

where the expression of the polynomials $\varphi_i(s)$, $\psi_i(s)$ are provided in Lemma 1 of [154]. □

Lie and Norsett in [154] also provided a complete study of the order of the resulting methods, stating order conditions by means of the study of variational matrices, and showing that the maximum attainable order of a k-step m-stage collocation method is $2m + k - 1$. They also proved that there exist $\binom{m+k-1}{k-1}$ nodes that allow superconvergence and, in analogy with Runge-Kutta methods, they are named *multistep Gaussian* collocation points. As Hairer-Wanner stated in [122], these methods are not stiffly stable and, therefore, they are not suited for stiff problems: in order to obtain better stability properties, they derived methods of highest order $2m+k-2$, imposing $c_m = 1$ and deriving the other collocation abscissa in a suited way to achieve this highest order and named the corresponding methods of "Radau"-type, studied their stability properties, deriving also many A-stable methods.

3.2 Two-step collocation and almost collocation methods

In the context of collocation-based methods for the numerical solution of the problem (1.1.1), our strengths have been devoted to the extension of the multistep collocation technique to the class of TSRK methods (2.2.1) we have discussed in Section 2.2, pursuing the aim of deriving highly stable

collocation-based methods belonging to the family of TSRK formulae. Different approaches to the construction of continuous TSRK methods outside collocation have been presented in literature in the papers [17], [19] and [141]. We introduce the continuous approximant

$$\begin{cases} P(t_n + sh) = \varphi_0(s)y_{n-1} + \varphi_1(s)y_n \\ \quad + h\sum_{j=1}^{m}\Big(\chi_j(s)f(P(t_{n-1} + c_j h)) + \psi_j(s)f(P(t_n + c_j h))\Big), \ s \in [0,1], \\ y_{n+1} = P(t_{n+1}), \end{cases} \tag{3.2.1}$$

which is modeled on the shape of the numerical method under consideration, i.e. the TSRK method (2.2.1): in fact, in the advancing from the grid point t_n to the point t_{n+1}, it depends on the approximations y_{n-1} and y_n to the solution $y(t)$ of (1.1.1) in the two consecutive step points t_{n-1} and t_n, but also on the stage derivatives related to the last two subintervals of integration. We assume that $P(t_n + sh)$ is an algebraic polynomial and we denote it as *collocation polynomial*, while the corresponding method is called *two-step collocation method*.

Similarly as in the case of multistep collocation methods, the collocation polynomial (3.2.1) is expressed as linear combination of the basis functions

$$\{\varphi_0(s),\ \varphi_1(s),\ \chi_j(s),\ \psi_j(s),\ j = 1, 2, \ldots, m\}, \tag{3.2.2}$$

which are unknown algebraic polynomials to be suitably determined. It is required that the polynomial $P(t_n+sh)$ interpolates the solution in the points t_{n-1} and t_n, i.e.

$$P(t_{n-1}) = y_{n-1}, \quad P(t_n) = y_n, \tag{3.2.3}$$

and collocates it in the points $t_{n-1} + c_i h$, $t_n + c_i h$, $i = 1, 2, \ldots, m$, i.e.

$$\begin{aligned} P'(t_{n-1} + c_i h) &= f(t_{n-1} + c_i h, P(t_{n-1} + c_i h)),\ i = 1, 2, ..., m, \\ P'(t_n + c_i h) &= f(t_n + c_i h, P(t_n + c_i h)), \quad i = 1, 2, ..., m. \end{aligned} \tag{3.2.4}$$

We observe that the first of the (3.2.4) is peculiar of TSRK-type methods, because of the special dependency on the internal points $t_{n-1} + c_i h$: we will next analyze how this dependency affects the stability properties of the resulting methods in the context of collocation-based methods.

Since the collocation polynomial (3.2.1) is a linear combination of the functions (3.2.2), the conditions (3.2.3) and (3.2.4) are then reflected on the

basis polynomials in the following way: the counterpart of the interpolation conditions (3.2.3) on (3.2.2) is

$$\begin{array}{llll} \varphi_0(-1)=1, & \varphi_1(-1)=0, & \chi_j(-1)=0, & \psi_j(-1)=0, \\ \varphi_0(0)=0, & \varphi_1(0)=1, & \chi_j(0)=0, & \psi_j(0)=0, \end{array} \tag{3.2.5}$$

while the analog of the collocation conditions (3.2.4) on (3.2.2) is

$$\begin{array}{llll} \varphi_0'(c_{i-1})=0, & \varphi_1'(c_{i-1})=0, & \chi_j'(c_{i-1})=\delta_{ij}, & \psi_j'(c_{i-1})=0, \\ \varphi_0'(c_i)=0, & \varphi_1'(c_i)=0, & \chi_j'(c_i)=0, & \psi_j'(c_i)=\delta_{ij}, \end{array} \tag{3.2.6}$$

where δ_{ij} is the usual Kronecker delta, $i,j=1,2,\ldots,m$.

The following subsections aim to analyze the properties of two-step collocation methods (3.2.1), providing the main results concerning their possible representations, the error and order analysis, the linear stability properties. These original results have been included in the papers [69, 70, 88, 89, 92, 95].

3.2.1 Representation of two–step collocation methods

The first step of our analysis aims to create a link between two-step collocation methods (3.2.1) and the family of TSRK formulae (2.2.1). The technique we apply in order to succeed in this direction is a suitable modification of the procedure introduced by Lie and Norsett [154] in the context of collocation-based multistep Runge–Kutta methods. The result we are going to prove characterizes the family of two-step collocation methods as a special class of TSRK methods (2.2.1) and, moreover, provides a constructive tool for the derivation of the basis functions (3.2.2): this theorem can be therefore considered as the analog of the result introduced by Guillou-Soulé and Wright in the context of collocation-based Runge–Kutta methods, and reported in Section 3.1.1.

Theorem 3.2.1 *The method defined by (3.2.1) is equivalent to the TSRK method (2.2.1) with*

$$\begin{array}{lll} \theta=\varphi_0(1), & v_j=\chi_j(1), & w_j=\psi_j(1), \\ u_i=\varphi_0(c_i), & a_{ij}=\chi_j(c_i), & b_{ij}=\psi_j(c_i), \end{array}$$

$i,j=1,\ldots,m$, *where*

$$\psi_j(s)=\int_0^s l_j(\tau)d\tau-\frac{\int_{-1}^0 l_j(\tau)d\tau}{\int_{-1}^0 M(\tau)d\tau}\int_0^s M(\tau)d\tau, \quad j=1,\ldots,m, \tag{3.2.7}$$

$$\chi_j(s) = \int_0^s \tilde{l}_j(\tau)d\tau - \frac{\int_{-1}^0 \tilde{l}_j(\tau)d\tau}{\int_{-1}^0 M(\tau)d\tau}\int_0^s M(\tau)d\tau, \quad j = 1, \ldots, m, \tag{3.2.8}$$

$$\begin{aligned} \varphi_0(s) &= -\frac{\int_0^s M(\tau)d\tau}{\int_{-1}^0 M(\tau)d\tau}, \\ \varphi_1(s) &= 1 + \frac{\int_0^s M(\tau)d\tau}{\int_{-1}^0 M(\tau)d\tau}, \end{aligned} \tag{3.2.9}$$

with

$$l_i(s) = \prod_{\substack{j=1 \\ j\neq i}}^{2m} \frac{s - d_j}{d_i - d_j}, \quad M(s) = \prod_{j=1}^{2m}(s - d_j), \quad \tilde{l}_i(s) = \prod_{\substack{j=1 \\ j\neq i}}^{2m} \frac{s - \tilde{d}_j}{\tilde{d}_i - \tilde{d}_j}, \tag{3.2.10}$$

in correspondence of the abscissas

$$\begin{cases} d_i = c_i, \\ d_{m+i} = c_i - 1, \end{cases} \qquad \begin{cases} \tilde{d}_i = c_i - 1, \\ \tilde{d}_{m+i} = c_i, \end{cases} \tag{3.2.11}$$

$i = 1, 2, \ldots, m$.

Proof: We first exhibit the form of the polynomials $\psi_j(s)$, $j = 1, 2, ...m$, assuming that the interpolation conditions (3.2.5) and the collocation ones (3.2.6) hold. Following [154], the collocation conditions (3.2.6) are surely satisfied by a polynomial of the form

$$\psi_j'(s) = l_j(s) + \frac{\alpha_0}{a_j}M(s) \tag{3.2.12}$$

with

$$a_j \;=\; \prod_{\substack{i=1 \\ i\neq j}}^{2m}(d_j - d_i),$$

and $\alpha_0 \in \mathbb{R}$. We set $\bar{\alpha}_0 = \frac{\alpha_0}{a_j}$ in (3.2.12) and integrate both of its sides, obtaining

$$\psi_j(s) = \int_0^s l_j(\tau)d\tau + \bar{\alpha}_0 \int_0^s M(\tau)d\tau. \tag{3.2.13}$$

Imposing the interpolation conditions (3.2.5) on this function, we derive $\bar{\alpha}_0$ as solution of the linear equation

$$\bar{\alpha}_0 \int_{-1}^{0} M(\tau)d\tau = -\int_{-1}^{0} l_j(\tau)d\tau.$$

The computed value of $\bar{\alpha}_0$ replaced in Equation (3.2.13) leads to (3.2.7).

The form (3.2.8) of the polynomials $\chi_j(s)$, $j = 1, 2, ...m$, is obtained in a similar way, by referring to the collocation abscissae $\tilde{d}_i$ in (3.2.11) instead of d_i.

We finally investigate on the form of the basis polynomials $\varphi_i(s)$, $i = 0, 1$. In order to fulfill the collocation conditions (3.2.6), the functions $\varphi_i(s)$, $i = 0, 1$, have to assume the form

$$\varphi_i(s) = \gamma_0^{(i)} + \gamma_1^{(i)} \int_0^s M(\tau)d\tau,$$

with $\gamma_j^{(i)} \in \mathbb{R}$. Let us treat $\varphi_0(s)$ and $\varphi_1(s)$ separately. Concerning $\varphi_0(s)$, imposing the interpolation conditions (3.2.5), we have

$$0 = \varphi_0(0) = \gamma_0^{(0)} + \gamma_1^{(0)} \int_0^0 M(\tau)d\tau$$

and, as a consequence, we obtain $\gamma_0^{(0)} = 0$. Moreover, since

$$1 = \varphi_0(-1) = \gamma_1^{(0)} \int_0^{-1} M(\tau)d\tau = -\gamma_1^{(0)} \int_{-1}^{0} M(\tau)d\tau$$

we obtain

$$\gamma_1^{(0)} = -\frac{1}{\int_{-1}^{0} M(\tau)d\tau}.$$

In analogous way, concerning the polynomial $\varphi_1(s)$, the interpolation conditions (3.2.5) lead to

$$\gamma_0^{(1)} = 1, \qquad \gamma_1^{(1)} = \frac{1}{\int_{-1}^{0} M(\tau)d\tau}.$$

□

3.2.2 Continuous order conditions

This section is devoted to the construction of continuous order conditions for two-step collocation methods (3.2.1), under the assumption that the collocation polynomial $P(t_n+sh)$ provides an uniform approximation to $y(t_n+sh)$, $s \in [0,1]$, of order p. As a consequence, the stage values $P(t_n + c_j h)$ have (stage) order $q = p$.

To this end, we investigate on the local discretization error $\xi(t_n + sh)$ associated to (3.2.1), i.e. the residuum obtained by replacing $P(t_n + sh)$ by $y(t_n + sh)$, $P(t_n + c_j h)$ by $y(t_n + c_j h)$, $j = 1, 2, \ldots, m$, y_{n-1} by $y(t_{n-1})$ and y_n by $y(t_n)$ in (3.2.1), where $y(t)$ is the true solution to (1.1.1). This leads to

$$\begin{aligned} \xi(t_n + sh) &= y(t_n + sh) - \varphi_0(s)y(t_n - h) - \varphi_1(s)y(t_n) \\ &- h\sum_{j=1}^{m}\Big(\chi_j(s)y'(t_n + (c_j - 1)h) + \psi_j(s)y'(t_n + c_j h)\Big), \end{aligned} \tag{3.2.14}$$

$s \in (0,1]$, $n = 1, 2, \ldots, N-1$. We have the following result.

Theorem 3.2.2 *Assume that the function $f(y)$ is sufficiently smooth. Then the method (3.2.1) has uniform order p if the following conditions are satisfied*

$$\begin{cases} \varphi_0(s) + \varphi_1(s) = 1, \\ \dfrac{(-1)^k}{k!}\varphi_0(s) + \displaystyle\sum_{j=1}^{m}\left(\chi_j(s)\frac{(c_j-1)^{k-1}}{(k-1)!} + \psi_j(s)\frac{c_j^{k-1}}{(k-1)!}\right) = \frac{s^k}{k!}, \end{cases} \tag{3.2.15}$$

$s \in [0,1]$, $k = 1, 2, \ldots, p$. Moreover, the local discretization error (3.2.14) takes the form

$$\xi(t_n + sh) = h^{p+1}C_p(s)y^{(p+1)}(t_n) + O(h^{p+2}), \tag{3.2.16}$$

as $h \to 0$, where the error function $C_p(s)$ is defined by

$$C_p(s) = \frac{s^{p+1}}{(p+1)!} - \frac{(-1)^{p+1}}{(p+1)!}\varphi_0(s) - \sum_{j=1}^{m}\left(\chi_j(s)\frac{(c_j-1)^p}{p!} + \psi_j(s)\frac{c_j^p}{p!}\right). \tag{3.2.17}$$

Proof:. Expanding $y(t_n + sh)$, $y(t_n - h)$, $y'(t + (c_j - 1)h)$ and $y(t_n + c_j h)$ into Taylor series around the point t_n and collecting terms with the same powers

of h we obtain

$$\begin{aligned}\xi(t_n+sh) &= \big(1-\varphi_0(s)-\varphi_1(s)\big)y(t_n)\\ &+ \sum_{k=1}^{p+1}\left(\frac{s^k}{k!}-\frac{(-1)^k}{k!}\varphi_0(s)\right)h^k y^{(k)}(t_n)\\ &- \sum_{k=1}^{p+1}\sum_{j=1}^{m}\left(\chi_j(s)\frac{(c_j-1)^{k-1}}{(k-1)!}+\psi_j(s)\frac{c_j^{k-1}}{(k-1)!}\right)h^k y^{(k)}(t_n)\\ &+ O(h^{p+2}).\end{aligned}$$

Equating to zero the terms of order k, $k=0,1,\ldots,p$, we obtain order conditions (3.2.15). Comparing the terms of order $p+1$ we obtain (3.2.16) with error function $C_p(s)$ defined by (3.2.17). □

The condition

$$\varphi_0(s)+\varphi_1(s)=1, \quad s\in[0,1],$$

is the generalization of preconsistency conditions for TSRK methods (2.2.1), compare [138]: this condition, in the context of TSRK methods, implies that θ, $\tilde{\theta}$, u_j and $\tilde{u}_j$ appearing in (2.2.1) satisfy the conditions

$$\theta+\tilde{\theta}=1,\ u_j+\tilde{u}_j=1,\ j=1,2,\ldots,m.$$

The set of order conditions (3.2.15) constitute a linear system of $p+1$ equations in $2m+2$ unknowns, i.e. the $2m+2$ basis functions (3.2.2). As a consequence, in order to ensure the compatibility of the system (3.2.15), p can be at most equal to $2m+1$. This remark leads to the following result.

Corollary 3.2.1 *The maximum attainable uniform order of convergence for two-step collocation methods (3.2.1) is* $2m+1$.

Let us now discuss on the unicity to the solution of the system (3.2.15) when $p=2m+1$, i.e. when we aim to achieve the maximum attainable uniform order. For this purpose, we introduce the recurrence relation

$$\begin{aligned}r^{(0)}(j) &= \frac{(-1)^j}{j},\\ r^{(k)}(j) &= r^{(k-1)}(j)-r^{(k-1)}(j-1)d_k,\end{aligned} \tag{3.2.18}$$

for $k,j=1,2,\ldots,2m$, where

$$d_k=\begin{cases} c_k-1, & k=1,2,\ldots,m,\\ c_{k-m}, & k=m+1,\ldots,2m.\end{cases} \tag{3.2.19}$$

Theorem 3.2.3 *Assume that $c_i \neq c_j$, $c_i \neq c_j - 1$ for $i \neq j$ and, moreover, that $r^{(k)}(j) \neq 0$, $k, j = 1, 2, \ldots, 2m$. Then, the system (3.2.15) corresponding to $p = 2m + 1$ has a unique solution $\varphi_0(s)$, $\varphi_1(s)$, $\chi_j(s)$, and $\psi_j(s)$, $j = 1, 2, \ldots, m$, which are polynomials of degree $\leq 2m + 1$.*

Proof: The coefficient matrix of the system (3.2.15) has the form

$$H = \begin{bmatrix} 1 & 1 & 0 & \ldots & 0 \\ -1 & 0 & 1 & \ldots & 1 \\ \frac{1}{2!} & 0 & d_1 & \ldots & d_{2m} \\ -\frac{1}{3!} & 0 & \frac{d_1^2}{2} & \ldots & \frac{d_{2m}^2}{2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{(-1)^{j-1}}{(j-1)!} & 0 & \frac{d_1^{j-2}}{(j-2)!} & \ldots & \frac{d_{2m}^{j-2}}{(j-2)!} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{(-1)^{2m+1}}{(2m+1)!} & 0 & \frac{d_1^{2m}}{(2m)!} & \ldots & \frac{d_{2m}^{2m}}{(2m)!} \end{bmatrix},$$

where the abscissas d_i are given in (3.2.19). In order to achieve the thesis, we need to prove that the matrix H is nonsingular by computing its determinant. First of all, we consider the Laplace expansion of $detH$ along the second column, obtaining $\det H = -\det H'$, where H' is the matrix we derive by dropping the first row and the second column of H, i.e.

$$H' = \begin{bmatrix} -1 & 1 & \ldots & 1 \\ \frac{1}{2!} & d_1 & \ldots & d_{2m} \\ -\frac{1}{3!} & \frac{d_1^2}{2} & \ldots & \frac{d_{2m}^2}{2} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{(-1)^j}{j!} & \frac{d_1^{j-1}}{(j-1)!} & \ldots & \frac{d_{2m}^{j-1}}{(j-1)!} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{(-1)^{2m}}{(2m)!} & \frac{d_1^{2m}}{(2m)!} & \ldots & \frac{d_{2m}^{2m}}{(2m)!} \end{bmatrix}.$$

Using the elementary properties of determinants (compare [149]), obtain we obtain

$$detH' = \frac{1}{2!} \cdot \frac{1}{3!} \cdot \ldots \cdot \frac{1}{(2m)!} \cdot detH''$$

where

$$H'' = \begin{bmatrix} -1 & 1 & \dots & 1 \\ \frac{1}{2} & d_1 & \dots & d_{2m} \\ -\frac{1}{3} & d_1^2 & \dots & d_{2m}^2 \\ \vdots & \vdots & \vdots & \vdots \\ \frac{(-1)^j}{j} & d_1^{j-1} & \dots & d_{2m}^{j-1} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{(-1)^{2m}}{2m} & d_1^{2m} & \dots & d_{2m}^{2m} \end{bmatrix}.$$

By setting $r^{(0)}(j) = \frac{(-1)^j}{j}$, the matrix H'' takes the form

$$H'' = \begin{bmatrix} r^{(0)}(1) & 1 & \dots & 1 \\ r^{(0)}(2) & d_1 & \dots & d_{2m} \\ r^{(0)}(3) & d_1^2 & \dots & d_{2m}^2 \\ \vdots & \vdots & \vdots & \vdots \\ r^{(0)}(j) & d_1^{j-1} & \dots & d_{2m}^{j-1} \\ \vdots & \vdots & \vdots & \vdots \\ r^{(0)}(2m) & d_1^{2m} & \dots & d_{2m}^{2m} \end{bmatrix}.$$

With this positions, the form of H'' is now clearly of Vandermonde type, except the first column: in order to compute its determinant, we can apply with suitable modifications the procedure introduced to compute the determinant of the Vandermonde matrix [158]. By means of $2m$ switches, we can put the first column of H'' in the last position, obtaining the matrix

$$H''' = \begin{bmatrix} 1 & \dots & 1 & r^{(0)}(1) \\ d_1 & \dots & d_{2m} & r^{(0)}(2) \\ d_1^2 & \dots & d_{2m}^2 & r^{(0)}(3) \\ \vdots & \vdots & \vdots & \vdots \\ d_1^{j-1} & \dots & d_{2m}^{j-1} & r^{(0)}(j) \\ \vdots & \vdots & \vdots & \vdots \\ d_1^{2m} & \dots & d_{2m}^{2m} & r^{(0)}(2m) \end{bmatrix},$$

and, as a consequence, $\det H''' = \det H''$ and

$$\det H = -\frac{1}{2!} \cdot \frac{1}{3!} \cdot \ldots \cdot \frac{1}{(2m)!} \cdot \det H'''.$$

We now aim to compute the determinant of H'''. We subtract d_1 times the first row from the other $2m$ rows, obtaining the matrix

$$\begin{bmatrix} 1 & 1 & \dots & 1 & r^{(0)}(1) \\ 0 & d_2 - d_1 & \dots & d_{2m} - d_1 & r^{(0)}(2) - r^{(0)}(1)d_1 \\ 0 & d_2^2 - d_1 d_2 & \dots & d_{2m}^2 - d_1 d_{2m} & r^{(0)}(3) - r^{(0)}(2)d_1 \\ \vdots & \vdots & \vdots & \vdots & \\ 0 & d_2^{j-1} - d_1 d_2^{j-2} & \dots & d_{2m}^{j-1} - d_1 d_{2m}^{j-2} & r^{(0)}(j-1) - r^{(0)}(j-2)d_1 \\ \vdots & \vdots & \vdots & \vdots & \\ 0 & d_2^{2m} - d_1 d_2^{2m-1} & \dots & d_{2m}^{2m} - d_1 d_{2m}^{2m-1} & r^{(0)}(2m) - r^{(0)}(2m-1)d_1 \end{bmatrix},$$

whose determinant is the equal to the determinant of the matrix

$$\begin{bmatrix} d_2 - d_1 & \dots & d_{2m} - d_1 & r^{(1)}(2) \\ d_2(d_2 - d_1) & \dots & d_{2m}(d_{2m} - d_1) & r^{(1)}(3) \\ \vdots & \vdots & \vdots & \\ \vdots & \vdots & \vdots & \\ d_2^{2m-1}(d_2 - d_1) & \dots & d_{2m}^{2m-1}(d_2 - d_1) & r^{(1)}(2m) \end{bmatrix},$$

where

$$r^{(1)}(j) = r^{(0)}(j) - r^{(0)}(j-1)d_1, \qquad j = 2, ...2m.$$

As a consequence, we obtain

$$\det H''' = (d_2 - d_1)(d_3 - d_1)\dots(d_{2m} - d_1) \cdot \det \begin{bmatrix} 1 & \dots & 1 & r^{(1)}(2) \\ d_2 & \dots & d_{2m} & r^{(1)}(3) \\ \vdots & \vdots & \vdots & \\ \vdots & \vdots & \vdots & \\ d_2^{2m-1} & \dots & d_{2m}^{2m-1} & r^{(1)}(2m) \end{bmatrix},$$

and, iterating this process, we get

$$\det H''' = \prod_{1 \le i \le k \le 2m} (d_k - d_i) r^{(2m-1)}(2m)$$

where

$$r^{(2m-1)}(2m) = r^{(2m-2)}(2m) - r^{(2m-2)}(2m-1)d_{2m-1}.$$

According to the hypothesis, it is $r^{(2m-1)}(2m) \neq 0$ and, as a consequence, it follows that H''' is nonsingular and, finally, that H is nonsingular. $\square$

The next result shows that the polynomials $\varphi_0(s)$, $\varphi_1(s)$, $\chi_j(s)$, and $\psi_j(s)$, $j = 1, 2, \ldots, m$, corresponding to the methods of order $p = 2m + 1$ satisfy the interpolation and collocation conditions (3.2.5)-(3.2.6).

Theorem 3.2.4 *Assume that $\varphi_0(s)$, $\varphi_1(s)$, $\chi_j(s)$, and $\psi_j(s)$, $j = 1, 2, \ldots, m$, satisfy the system of order conditions (3.2.15) for $p = 2m + 1$. Then these polynomials fulfill the interpolation conditions (3.2.5) and the collocation conditions (3.2.6).*

Proof:. The conditions (3.2.5) follow immediately by substituting $s = 0$ and $s = -1$ into the system (3.2.15) corresponding to $p = 2m + 1$. To show (3.2.6) we differentiate (3.2.15), obtaining

$$\begin{cases} \varphi_0'(s) + \varphi_1'(s) = 0, \\ \dfrac{(-1)^k}{k!}\varphi_0'(s) + \displaystyle\sum_{j=1}^{m} \left(\chi_j'(s) \frac{(c_j - 1)^{k-1}}{(k-1)!} + \psi_j'(s) \frac{c_j^{k-1}}{(k-1)!} \right) = \frac{s^{k-1}}{(k-1)!}, \end{cases} \tag{3.2.20}$$

$k = 1, 2, \ldots, 2m + 1$. Substituting $s = c_i$ and $s = c_i - 1$, $i = 1, 2, \ldots, m$, into (3.2.20) we obtain (3.2.6). □

As a consequence of this theorem, we can observe that the basis functions computed by solving the linear system of the order conditions (3.2.15) for $p = 2m + 1$ automatically satisfy all the interpolation conditions (3.2.5) and all the collocation ones (3.2.6) and, as a consequence, the collocation polynomial (3.2.1) satisfies all those conditions. This aspect can be used as a constructive tool of two-step collocation methods, since they are univocally determined by the system of the order conditions (3.2.15).

3.2.3 Superconvergence

The order conditions for two-step collocation methods (3.2.1) can be inferred also in a different way, using some tools borrowed from the Calculus of Variations, such as the Alekseev–Gröbner theorem [9, 114].

Theorem 3.2.5 *(Alekseev-Gröbner, see [119]). Denote by y and z the solutions of*

$$y' = f(x, y), \quad y(x_0) = y_0, \tag{3.2.21}$$

$$z' = f(x, z) + g(x, z), \quad z(x_0) = y_0, \tag{3.2.22}$$

respectively and suppose that $\partial f/\partial y$ exists and is continuous. Then the solutions of (3.2.21) and of its perturbed" version (3.2.22) are connected by the relation

$$z(x) = y(x) + \int_{x_0}^{x} \frac{\partial y}{\partial y_0}(x, s, z(s)) \cdot g(s, z(s))ds. \tag{3.2.23}$$

Then, following result holds.

Theorem 3.2.6 *Let*

$$G_i = \det \begin{bmatrix} \int_{-1}^{0} M(\tau)\tau^i d\tau & \int_{-1}^{0} M(\tau)\tau^{i+1} d\tau \\ \int_{0}^{1} M(\tau)\tau^i d\tau & \int_{0}^{1} M(\tau)\tau^{i+1} d\tau \end{bmatrix}, \tag{3.2.24}$$

where the function $M(\tau)$ is given by (3.2.10). Then the corresponding two-step collocation method has order $2m + \sigma$ if and only if $G_i = 0$ for $i = 0, 1, \ldots, \sigma - 1$.

Proof: We proceed along the lines drawn by Lie and Norsëtt in [154]. By using the Alekseev-Gröbner formula, we define the error associated to the method (3.2.1) by

$$P(t) - y(t) = \int_{t_n}^{t} \Phi(t, \tau, P(\tau)) \cdot (P'(\tau) - f(\tau, P(\tau)))d\tau$$

where Φ is a suitable variational matrix. The interpolation conditions

$$P(t_{n-i}) = y(t_{n.i}) \qquad i = 0, 1,$$

lead to

$$\int_{t_n}^{t_{n-1}} \Phi(t, \tau, P(\tau)) \cdot (P'(\tau) - f(\tau, P(\tau)))d\tau = 0,$$

and moreover, assuming $h = 1$ and $t_n = 0$ and using the scaled variable s, we obtain

$$\int_{-1}^{0} g(\tau) = 0,$$

under the assumption that

$$g(\tau) = \Phi(t, \tau, P(\tau)) \cdot (P'(\tau) - f(\tau, P(\tau))).$$

Let us consider the linear subspace

$$V = \{g : \int_{-1}^{0} g(\tau)d\tau = 0\},$$

and let us derive a quadrature formula integrating functions belonging to V. We approximate any function $p \in V$ by an interpolation polynomial $\rho \in V$, such that $p(d_i) = \rho(d_i)$, where the nodes d_i are defined in (3.2.11). As a consequence,

$$p(s) - \rho(s) = (s - d_1) \cdot (s - d_2) \dots (s - d_{2m}) \cdot r(s),$$

i.e. $p(s) - \rho(s) = M(s) \cdot r(s)$, where $M(s)$ is defined in (3.2.10). Since V is a linear subspace and $p(s), \rho(s) \in V$, also $M(s) \cdot r(s) \in V$. Suppose that $r(s)$ is a polynomial of the type

$$r(s) = \sum_{l=0}^{1} \alpha_l s^{i+l}, \qquad i = 0, 1, ...$$

Then, the condition $M(s) \cdot r(s) \in V$ becomes

$$\begin{bmatrix} \int_{-1}^{0} M(s)s^i ds & \int_{-1}^{0} M(s)s^{i+1} \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} = 0,$$

while the error $\int_0^1 M(s)r(s)$ associated to the quadrature formula is zero if

$$\begin{bmatrix} \int_{0}^{1} M(s)s^i ds & \int_{0}^{1} M(s)s^{i+1} \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} = 0.$$

Then, the resulting system

$$\begin{bmatrix} \int_{-1}^{0} M(s)s^i ds & \int_{-1}^{0} M(s)s^{i+1} \\ \int_{0}^{1} M(s)s^i ds & \int_{0}^{1} M(s)s^{i+1} \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

has a nontrivial solution if and only if the coefficient matrix

$$\begin{bmatrix} \int_{-1}^{0} M(s)s^i ds & \int_{-1}^{0} M(s)s^{i+1} \\ \int_{0}^{1} M(s)s^i ds & \int_{0}^{1} M(s)s^{i+1} \end{bmatrix}$$

is singular. We denote by G_i its determinant. Let us suppose that $G_i = 0$ for $i = 0, 1, ..., \sigma - 1$. This means that the quadrature formula is able to exactly integrate polynomials of the type $M(s) \sum_{l=0}^{1} \alpha_l s^{i+l}$. The maximum degree of these polynomials is $2m + \sigma$ and, therefore, the quadrature formula and, as a consequence, the corresponding method, have both the same order $2m + \sigma$. □

The following theorem explores the possibility to achieve superconvergence.

Theorem 3.2.7 *The maximum attainable order of two-step collocation methods (3.2.1) is* $3m$.

Proof: The set of conditions $G_i = 0$ of Theorem 3.2.6 leads to nonlinear system in the unknowns $c_1, c_2, ..., c_m$. Let us consider a subset $\Gamma(\zeta)$ of ℓ equations of the system, where $\Gamma(\zeta) : \mathbb{R}^m \to \mathbb{R}^\ell$, which has unique solution if and only if $\ell = m$. As a consequence, σ can be at most equal to m and, for this reason, the maximum attainable order is $3m$. □

3.2.4 Two-step almost collocation methods

We have asserted in Section 3.2.2 that the set of order conditions (3.2.15) for $p = 2m + 1$ provides as solution the $2m + 2$ basis functions (3.2.2) corresponding to a two-step collocation method (3.2.1) totally fulfilling the sets of interpolation/collocation conditions (3.2.5)-(3.2.6). $p = 2m + 1$ is the maximum attainable uniform order but, however, we will also be interested in methods having lower uniform order than $2m + 1$, especially for stability reasons, since we aim for strong stability properties (e.g. A-stability and L-stability). In fact, we remind that two-step collocation methods are implicit methods and, "nothing less than A-stability can be admitted for implicit numerical methods[1]".

In the remainder, we will mainly be interested in methods of order $p = m + r$, where $r = 1, 2, \ldots, m$, and the next theorem examines the solvability of the linear systems of equations (3.2.15) corresponding to these orders.

[1]I was very impressed by this sentence, here reported as well as I can remember, which was pronounced by John C. Butcher during the conference SciCADE 2007 in Saint Malo, which was the first conference on Numerical Analysis I attended in my life

Theorem 3.2.8 *Assume that $c_i \neq c_j$, and $c_i \neq c_j - 1$ for $i \neq j$. Then the system of continuous order conditions (3.2.15) corresponding to $p = m + r$, where $r = 1, 2, \ldots, m$, has a unique solution $\varphi_1(s)$, $\chi_j(s)$, $j = m-r+1, m-r+2, \ldots, m$, and $\psi_j(s)$, $j = 1, 2, \ldots, m$, for any given polynomials $\varphi_0(s)$ and $\chi_j(s)$, $j = 1, 2, \ldots, m-r$.*

Proof:. Observe that the polynomial $\varphi_1(s)$ is uniquely determined from the first equation of (3.2.15). The proof follows from the fact that the matrices of these systems (3.2.15) corresponding to $\chi_j(s)$, $j = m-r+1, m-r+2, \ldots, m$, are Vandermonde matrices and, therefore, the solution exists and is unique. □

For the methods of order $p = m + r$, $r = 1, 2, \ldots, m$, we will choose $\varphi_0(s)$ and $\chi_j(s)$, $j = 1, 2, \ldots, m-r$, as polynomials of degree $\leq m+r$ which satisfy the interpolation conditions

$$\varphi_0(0) = 0, \quad \chi_j(0) = 0, \quad j = 1, 2, \ldots, m - r, \tag{3.2.25}$$

and the collocation conditions

$$\varphi_0'(c_i) = 0, \quad \chi_j'(c_i) = 0, \quad j = 1, 2, \ldots, m - r. \tag{3.2.26}$$

This leads to the polynomials $\varphi_0(s)$ and $\chi_j(s)$, $j = 1, 2, \ldots, m - r$, of the form

$$\varphi_0(s) = s\big(q_0 + q_1 s + \cdots + q_{m+r-1} s^{m+r-1}\big),$$

$$\chi_j(s) = s\big(r_{j,0} + r_{j,1} s + \cdots + r_{j,m+r-1} s^{m+r-1}\big),$$

$j = 1, 2, \ldots, m - r$, where

$$q_0 + 2q_1 c_i + \cdots + (m + r) q_{m+r-1} c_i^{m+r-1} = 0,$$

$$r_{j,0} + 2r_{j,1} c_i + \cdots + (m + r) r_{j,m+r-1} c_i^{m+r-1} = 0,$$

$j = 1, 2, \ldots, m - r$, $i = 1, 2, \ldots, m$. The methods we will obtain in this way satisfy only some of the interpolation/collocation conditions (3.2.5) and (3.2.6). In particular we aim to prove that the conditions (3.2.25) and (3.2.26) imposed on the fixed basis functions among $\varphi_0(s)$ and $\chi_j(s)$ are inherited via order conditions by all the other basis functions, as asserted by the following theorem.

Theorem 3.2.9 *Assume that $\varphi_0(s)$ and $\chi_j(s)$, $j = 1, 2, \ldots, m-r$, satisfy (3.2.25) and (3.2.26). Then the solution $\varphi_1(s)$, $\chi_j(s)$, $j = m-r+1, m-r+2, \ldots, m$, and $\psi_j(s)$, $j = 1, 2, \ldots, m$ of (3.2.15) satisfy the interpolation conditions*

$$\begin{array}{l} \varphi_1(0) = 1, \quad \chi_j(0) = 0, \quad j = m-r+1, m-r+2, \ldots, m, \\ \psi_j(0) = 0, \quad j = 1, 2, \ldots, m, \end{array} \tag{3.2.27}$$

and the collocation conditions

$$\begin{array}{l} \varphi_1'(c_i) = 0, \quad \chi_j'(c_i) = 0, \quad j = m-r+1, m-r+2, \ldots, m, \\ \psi_j'(c_i) = \delta_{ij}, \quad j = 1, 2, \ldots, m, \end{array} \tag{3.2.28}$$

$i = 1, 2, \ldots, m$.

Proof:. Substituting $s = 0$ into (3.2.15) corresponding to $p = m + r$, $r = 1, 2, \ldots, m$, and taking into account that the solution to (3.2.15) is unique, the condition (3.2.27) follows. Differentiating (3.2.15) with respect to s and substituting $s = c_i$, $i = 1, 2, \ldots, m$, into the resulting relations for $k = 1, 2, \ldots, m + r$, we obtain (3.2.28). This completes the proof. □

The formulas obtained by imposing the conditions (3.2.25) and (3.2.26) will be then called *two-step almost collocation methods*. Two-step collocation methods and two-step almost collocation methods do not differ from the general expression of the collocation polynomial (3.2.1), but only on the fact that the former have the maximum attainable uniform order $p = 2m+1$ and, in force of Theorem 3.2.4 satisfy all the interpolation/collocation conditions (3.2.3)-(3.2.4), while the latter do not possess the highest attainable uniform order and satisfy only some of the mentioned conditions. We will next observe that two-step almost collocation methods will allow a proper balance between high order of convergence and strong stability properties. It follows from Theorem 3.2.9 that the polynomial $P(t)$ defined by the method (3.2.1) of order $p = m + r$, $r = 1, 2, \ldots, m$, satisfies the interpolation condition

$$P(t_n) = y_n$$

and the collocation conditions at the points c_i, i.e.,

$$P'(t_n + c_i h) = f(P(t_n + c_i h)), \quad i = 1, 2, \ldots, m.$$

However, in general, these methods do not satisfy the interpolation condition

$$P(t_{n-1}) = y_{n-1}$$

and the collocation conditions

$$P'(t_{n-1} + c_i h) = f(P(t_{n-1} + c_i h)), \quad i = 1, 2, \ldots, m.$$

In our search for highly stable methods (A-stability, L-stability) we will mainly be concerned with methods of order $p = 2m$ and $p = 2m - 1$. The advantage of these methods as compared, for example, with methods of low stage order, consists of the fact that they provide a uniform approximation $P(t)$ of order $p = 2m$ to the solution $y(t)$ of (1.1.1) over the entire interval of integration $[t_0, T]$. As a result these methods do not suffer from the order reduction phenomenon in the integration of stiff systems (see [36, 42, 122]). This is in contrast to implicit Runge-Kutta methods with m stages of order $p = 2m$, $p = 2m - 1$, or $p = 2m - 2$ for which the continuous approximation to $y(t)$ is only of (stage) order m. This leads to the reduction of order for stiff systems of ODEs for which the effective order is equal only to the stage order m.

3.2.5 Error propagation

It was demonstrated in Section 3.2.2 that the local discretization error at the point t_{n+1} of the m-stage method (3.2.1) of order p is given by

$$\xi(t_{n+1}) = C_p(1) h^{p+1} y^{(p+1)}(t_n) + O(h^{p+2}), \tag{3.2.29}$$

where the error constant $C_p(1)$ is defined by (3.2.17) for $s = 1$. We will also consider local error $\mathrm{le}(t_{n+1})$ defined by

$$\mathrm{le}(t_{n+1}) = C_p(1) h^{p+1} \tilde{y}^{(p+1)}(t_n) + O(h^{p+2}), \tag{3.2.30}$$

where $\tilde{y}(t)$ is the so-called local solution, i.e., the solution to the initial-value problem

$$\begin{cases} \tilde{y}'(t) = f(\tilde{y}(t)), & t \in [t_n, t_{n+1}], \\ \tilde{y}(t_n) = y_n. \end{cases} \tag{3.2.31}$$

Assuming that the function $f(y)$ appearing in (1.1.1) and (3.2.31) satisfies the Lipschitz condition of the form

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

with Lipschitz constant $L \geq 0$, by subtracting the integral forms of (1.1.1) and (3.2.31), we obtain

$$\|y(t) - \tilde{y}(t)\| \leq \|y(t_n) - y_n\| + L \int_{t_n}^{t} \|y(s) - \tilde{y}(s)\| ds,$$

$t \in [t_n, t_{n+1}]$. Using the Gronwall's lemma (compare, for example, [185]) yields

$$\|y(t) - \tilde{y}(t)\| \leq \|y(t_n) - y_n\| e^{L(t-t_n)}.$$

Hence,

$$\|y(t) - \tilde{y}(t)\| = O(h^p), \quad t \in [t_n, t_{n+1}].$$

Assuming that the function $f(y)$ is sufficiently smooth we have a similar conclusion for the derivatives of $y(t)$ and $\tilde{y}(t)$

$$\|y^{(i)}(t) - \tilde{y}^{(i)}(t)\| = O(h^p), \quad t \in [t_n, t_{n+1}], \quad i = 1, 2, \ldots,$$

compare [152], [186]. Therefore, we can conclude that the principal parts, i.e. the terms of order $p+1$, of the local discretization error (3.2.29) and the local error (3.2.30) are the same. This remark can be suitably exploited in the construction of an estimation to the local discretization error (3.2.29), by estimating the term $h^{p+1}\tilde{y}^{(p+1)}(t_n)$. The derivation of such estimate, which is necessary in order to implement the corresponding methods in a variable stepsize environment, will be discussed in Section 6.2.

We have previously observed that, if $P(t_n + sh)$ is an approximation of uniform order p to $y(t_n + sh)$, $s \in [0, 1)$, the stage order is also equal to p. Hence, the stage values $P(t_{n-1} + c_j h)$ and $P(t_n + c_j h)$ in (3.2.1) satisfy the relations

$$P(t_{n-1} + c_j h) = y(t_{n-1} + c_j h) - \eta_j h^{p+1} y^{(p+1)}(t_n) + O(h^{p+2}), \quad (3.2.32)$$
$$P(t_n + c_j h) = y(t_n + c_j h) - \eta_j h^{p+1} y^{(p+1)}(t_n) + O(h^{p+2}), \quad (3.2.33)$$

where

$$\eta_j = C_p(c_j), \quad j = 1, 2, \ldots, m,$$

are the stage error constants, which we put together in the vector

$$\eta = \left[\begin{array}{cccc} \eta_1 & \eta_2 & \ldots & \eta_m \end{array} \right]^T.$$

Let us focus our attention on the analysis of the local discretization error (3.2.29), considering the terms up to order $p+2$. The following result arises.

Theorem 3.2.10 *Assume that the function f appearing in (1.1.1) is sufficiently smooth and suppose that $P(t_n + sh)$ is an approximation of uniform order p to the solution $y(t_n + sh)$, $s \in [0, 1]$, of the problem (1.1.1). Then the local truncation error (3.2.29) of the method (3.2.1) takes the form*

$$\begin{aligned}\xi(t_n + sh) &= h^{p+1}C_p(s)y^{(p+1)}(t_n) + h^{p+2}C_{p+1}(s)y^{(p+2)}(t_n) \\ &+ h^{p+2}G_{p+1}(s)\frac{\partial f}{\partial y}\Big(y(t_n)\Big)y^{(p+1)}(t_n) + O(h^{p+3}), \quad (3.2.34)\end{aligned}$$

where

$$C_\nu(s) = \frac{s^{\nu+1}}{(\nu+1)!} - \frac{(-1)^{\nu+1}}{(\nu+1)!}\varphi_0(s) - \sum_{j=1}^{m}\left(\chi_j(s)\frac{(c_j-1)^\nu}{\nu!} + \psi_j(s)\frac{c_j^\nu}{\nu!}\right), \quad (3.2.35)$$

with $\nu = p,\ p+1$, and

$$G_{p+1}(s) = \sum_{j=1}^{m}\eta_j\Big(\chi_j(s) + \psi_j(s)\Big). \quad (3.2.36)$$

Proof: Substituting the relations (3.2.32) and (3.2.33) into (3.2.29), we obtain

$$\begin{aligned}\xi(t_n + sh) &= y(t_n + sh) - \varphi_0(s)y(t_n - h) - \varphi_1(s)y(t_n) \\ &- h\sum_{j=1}^{m}\left(\chi_j(s)f\left(y(t_{n-1} + c_jh) - \eta_jh^{p+1}y^{(p+1)}(t_n)\right)\right. \\ &+ \left.\psi_j(s)f\left(y(t_n + c_jh) - \eta_jh^{p+1}y^{(p+1)}(t_n)\right)\right) + O(h^{p+3}),\end{aligned}$$

and since f is sufficiently smooth, this formula can be rewritten as

$$\begin{aligned}\xi(t_n + sh) &= y(t_n + sh) - \varphi_0(s)y(t_n - h) - \varphi_1(s)y(t_n) \\ &- h\sum_{j=1}^{m}\left(\chi_j(s)y'(t_{n-1} + c_jh) + \psi_j(s)y'(t_n + c_jh)\right) \\ &+ h^{p+2}\sum_{j=1}^{m}\left(\eta_j\Big(\chi_j(s) + \psi_j(s)\Big)\frac{\partial f}{\partial y}\Big(y(t_n)\Big)y^{(p+1)}(t_n)\right) + O(h^{p+3}).\end{aligned}$$

Expanding $y(t_n + sh)$, $y(t_n - h)$, $y'(t_{n-1} + c_jh)$ and $y'(t_n + c_jh)$ into Taylor series around t_n and collecting the terms with the same powers of h, we

obtain

$$\begin{aligned}\xi(t_n+sh) &= (1-\varphi_0(s)-\varphi_1(s))y(t_n)\\ &+ \sum_{k=1}^{p+2}\left(\frac{s^k}{k!}-\frac{(-1)^k}{k!}\varphi_0(s)\right)h^k y^{(k)}(t_n)\\ &- \sum_{k=1}^{p+2}\sum_{j=1}^{m}\left(\chi_j(s)\frac{(c_j-1)^{k-1}}{(k-1)!}+\psi_j(s)\frac{c_j^{k-1}}{(k-1)!}\right)h^k y^{(k)}(t_n)\\ &+ \sum_{j=1}^{m}\left(\eta_j\Big(\chi_j(s)+\psi_j(s)\Big)h^{p+2}\frac{\partial f}{\partial y}\Big(y(t_n)\Big)y^{(p+1)}(t_n)\right)+O(h^{p+3}).\end{aligned}$$

Equating to zero terms of order $O(h^k)$, $k=0,1,\ldots,p$, we obtain the continuous order conditions (3.2.15), while the surviving terms provide the expression (3.2.34), with $C_\nu(s)$, $\nu=p,\ p+1$, and $G_{p+1}(s)$ are given by (3.2.35) and (3.2.36) respectively. □

The analysis of the high order terms appearing in the local discretization error (3.2.29) is necessary in view of a variable order implementation of two-step collocation/almost collocation methods. In particular, it would be useful to narrow the contribution of high order terms in (3.2.34), for instance by deriving methods such that the stage error constant $G_{p+1}(1)$ appearing in (3.2.34) is equal to zero. If this condition holds, the terms of order $p+2$ will only depend on the derivatives of the solution and not on the form of the equation. Moreover, this feature is of practical utility in the implementation of such methods in a variable order environment, since it simplifies the order changing strategy. The construction of two-step almost collocation methods (3.2.1) with $G_{p+1}(1)=0$ is discussed in Section 4.3.

3.2.6 Linear Stability Analysis

We conclude this chapter by analyzing the linear stability properties of two-step collocation/almost collocation methods (3.2.1). In order to achieve this purpose, we use the standard test equation

$$y'=\lambda y,\quad t\geq 0, \tag{3.2.37}$$

where λ is a complex parameter with negative real part. Applying the method (3.2.1) to the problem (3.2.37) and computing the resulting expression at the

points $s = c_i$, $i = 1, 2, \ldots, m$, and $s = 1$ we obtain

$$\begin{cases} P(t_n + c_i h) = \varphi_0(c_i) y_{n-1} + \varphi_1(c_i) y_n \\ \quad + h\lambda \sum_{j=1}^{m} \Big(\chi_j(c_i) P(t_{n-1} + c_j h) + \psi_j(c_i) P(t_n + c_j h) \Big), \\ y_{n+1} = \varphi_0(1) y_{n-1} + \varphi_1(1) y_n \\ \quad + h\lambda \sum_{j=1}^{m} \Big(\chi_j(1) P(t_{n-1} + c_j h) + \psi_j(1) P(t_n + c_j h) \Big), \end{cases} \tag{3.2.38}$$

$i = 1, 2, \ldots, m$, $n = 1, 2, \ldots, N-1$. Introducing the notation $z = h\lambda$,

$$P(t_n + ch) = \begin{bmatrix} P(t_n + c_1 h) \\ \vdots \\ P(t_n + c_m h) \end{bmatrix}, \quad \varphi_0(c) = \begin{bmatrix} \varphi_0(c_1) \\ \vdots \\ \varphi_0(c_m) \end{bmatrix}, \quad \varphi_1(c) = \begin{bmatrix} \varphi_1(c_1) \\ \vdots \\ \varphi_1(c_m) \end{bmatrix},$$

$$v^T = \begin{bmatrix} \chi_1(1) & \cdots & \chi_m(1) \end{bmatrix}^T, \quad w^T = \begin{bmatrix} \psi_1(1) & \cdots & \psi_m(1) \end{bmatrix}^T,$$

and

$$A = [\chi_j(c_i)]_{i,j=1}^{m}, \quad B = [\psi_j(c_i)]_{i,j=1}^{m},$$

(compare also Section 3.2.1 for the definition of v, w, A, and B) the relation (3.2.38) can be written in the vector form

$$\begin{cases} P(t_n + ch) = \varphi_0(c) y_{n-1} + \varphi_1(c) y_n + z\Big(AP(t_{n-1} + ch) + BP(t_n + ch) \Big), \\ y_{n+1} = \varphi_0(1) + \varphi_1(1) y_n + z\Big(v^T P(t_{n-1} + ch) + w^T P(t_n + ch) \Big), \end{cases} \tag{3.2.39}$$

$n = 1, 2, \ldots, N-1$. Hence, the stage values satisfy the relation

$$P(t_n + ch) = \big(I - zB \big)^{-1} \Big(\varphi_0(c) y_{n-1} + \varphi_1(c) y_n + zAP(t_{n-1} + ch) \Big), \tag{3.2.40}$$

which, replaced in the expression for y_{n+1} appearing in (3.2.39), leads to

$$\begin{aligned} y_{n+1} &= \Big(\varphi_0(1) + z w^T (I - zB)^{-1} \varphi_0(c) \Big) y_{n-1} \\ &+ \Big(\varphi_1(1) + z w^T (I - zB)^{-1} \varphi_1(c) \Big) y_n \\ &+ z\Big(v^T + z w^T (I - zB)^{-1} A \Big) P(t_{n-1} + ch). \end{aligned} \tag{3.2.41}$$

The relations (3.2.40) and (3.2.41) are equivalent to the recurrence relation

$$\begin{bmatrix} y_{n+1} \\ y_n \\ P(t_n + ch) \end{bmatrix} = M(z) \begin{bmatrix} y_n \\ y_{n-1} \\ P(t_{n-1} + ch) \end{bmatrix}, \tag{3.2.42}$$

where

$$M(z) = \begin{bmatrix} M_{11}(z) & M_{12}(z) & M_{13}(z) \\ 1 & 0 & 0 \\ Q\varphi_1(c) & Q\varphi_0(c) & zQA \end{bmatrix} \in \mathbb{C}^{(m+2)\times(m+2)}, \tag{3.2.43}$$

with

$$\begin{aligned} M_{11}(z) &= \varphi_1(1) + zw^T Q\varphi_1(c), \\ M_{12}(z) &= \varphi_0(1) + zw^T Q\varphi_0(c), \\ M_{13}(z) &= z(v^T + zw^T QA), \end{aligned}$$

and

$$Q = (I - zB)^{-1} \in \mathbb{C}^{m\times m}.$$

The matrix $M(z)$ is the *stability matrix* of the method (3.2.1) and its characteristic polynomial

$$p(\omega, z) = \det\Big(\omega I - M(z)\Big). \tag{3.2.44}$$

is the *stability function* of the method (3.2.1).

We observe that the expression (3.2.43) for the stability matrix of a two-step collocation/almost collocation method is coherent with the general form of the stability matrix (2.2.6) of a TSRK method (2.2.1): this is not surprising, in force of the equivalence result reported in Theorem 3.2.1, where the link between two-step collocation methods and TSRK methods has been clarified, and according to Theorem 3.2.8 for the case of two-step almost collocation methods.

The construction of A-stable and L-stable two-step almost collocation methods (3.2.1) is object of the following chapter, where many examples of such methods will also be provided.

Chapter 4

Families of two-step modified collocation methods

We focus our attention on the construction of families of highly stable two-step collocation/almost collocation methods (3.2.1) having some desired properties. These classes of methods have been introduced and analyzed in the papers [92, 94, 95, 98]. The chapter is organized as follows: Section 4.1 is devoted to the construction of two-step collocation/almost collocation methods (3.2.1) of order from $m+1$ up to $2m+1$, aiming for the derivation of A-stable and L-stable formulae within the class (3.2.1); Section 4.2 contains an extensive analysis of highly stable two-step continuous formulae of uniform order of convergence $p = m$, having up $m = 4$ stages; in Section 4.3 we discuss the construction of two-step almost collocation methods with narrowed contribution of the high order terms in the local discretization error, providing the analysis and practical construction of highly stable formulae with up to $m = 4$ stages; Section 4.4 concerns with the construction and analysis of two-step almost collocation methods equivalent to TSRK methods with structured coefficient matrices.

4.1 Practical construction of two-step collocation and almost collocation methods

This section is devoted to the construction of two-step collocation and almost collocation methods (3.2.1), with special attention to the derivation of highly

stable formulae. In particular, we present an extensive analysis of methods with one and two stages of order $m+1$, ..., $2m+1$. The tools we are going to use in this section in order to carry out our investigation have been provided in Chapter 3.

4.1.1 Analysis of methods with $m = 1$

We first concentrate on methods (3.2.1) having the maximum attainable uniform order, i.e. $p = 2m + 1 = 3$: as a consequence, according to Theorem 3.2.4, the resulting two-step collocation methods satisfy all the interpolation conditions (3.2.3) and the collocation ones (3.2.4). Solving the order conditions (3.2.15) corresponding to $m = 1$ and $p = 3$, we obtain a one-parameter family of two-step collocation methods (3.2.1) methods depending on the abscissa c. The coefficients of these methods are

$$\begin{aligned}
\varphi_0(s) &= \frac{s(6c(c-1)+3(1-2c)s+2s^2)}{1-6c^2},\\
\varphi_1(s) &= -\frac{(1+s)(6c^2-1+(1-6c)s+2s^2)}{1-6c^2},\\
\chi(s) &= -\frac{s(1+s)(2c+3c^2-(1+2c)s)}{1-6c^2},\\
\psi(s) &= \frac{s(1+s)(1-4c+3c^2+(1-2c)s)}{1-6c^2},
\end{aligned}$$

and the error constant $C_3(1)$ is given by

$$C_3(1) = \frac{1-3c-3c^2+12c^3-6c^4}{6(1-6c^2)},$$

with $c \neq \pm\sqrt{6}/6$. To investigate stability properties of the derived methods, it is more convenient to work with the polynomial obtained by multiplying the stability function (3.2.44) by its denominator. The resulting polynomial, which will be denoted by the same symbol $p(\omega, z)$, for this family of methods takes the form

$$p(\omega, z) = p_3(z)\omega^3 + p_2(z)\omega^2 + p_1(z)\omega + p_0(z), \tag{4.1.1}$$

where the polynomials $p_i(z)$, $i = 0, 1, 2, 3$, assume the form

$$\begin{aligned}
p_0(z) &= -(c-1)^2c^2z,\\
p_1(z) &= 5 - 12c + 6c^2 + (2 - 5c + 6c^2 - 6c^3 + 3c^4)z,\\
p_2(z) &= -4 + 12c - 12c^2 + (4 - 8c - 3c^2 + 6c^3 - 3c^4)z,\\
p_3(z) &= -1 + 6c^2 + (1 - 2c - 2c^2 + c^3)cz.
\end{aligned}$$

We will next investigate if there exist A-stable methods in this class of two-step collocation formulae of order $p = 3$. Let us restrict the stability function (4.1.1) to the axes of the pure imaginary z, by considering the function

$$\tilde{p}(\omega, y) := p(\omega, iy),$$

and compute the constant polynomial $\tilde{p}_0(y)$ with respect to ω, using the recursive procedure described in Section 2.1.3. This polynomial takes the form

$$\tilde{p}_0(y) = \alpha(c)y^4 + \beta(c)y^6 + \gamma(c)y^8,$$

where $\alpha(c)$, $\beta(c)$ and $\gamma(c)$ are polynomials with respect to the abscissa c. It follows from the Schur criterion in Theorem 2.1.4 that the condition

$$\tilde{p}_0(y) \geq 0, \quad \text{for all} \quad y \geq 0,$$

is the necessary condition for A-stability. However, it can be verified that the polynomials $\alpha(c)$, $\beta(c)$ and $\gamma(c)$ are not simultaneously greater or equal to zero for any c. This proves that A-stable methods do not exist in this class of methods of order $p = 3$. In fact the regions of stability of such methods are all bounded. This is illustrated in Fig.4.1 for $m = 1$ and $p = 3$, where we have plotted, in the (c, z)-plane, the stability interval of the methods corresponding to each value of c, considering $c \geq \frac{1}{2}$ in order to be $-1 \leq \theta < 1$ for zero-stability requirements.

Consider next the methods (3.2.1) of order $p = 2m = 2$, i.e. we relax one order conditions, falling into the class of two-step almost collocation methods (3.2.1) of order $p = m + r$ with $r = m$. We choose the polynomial $\varphi_0(s)$ of degree less than or equal to two which satisfies the interpolation condition $\varphi_0(0) = 0$ and the collocation one $\varphi_0'(c) = 0$. This leads to the polynomial $\varphi_0(s)$ of the form

$$\varphi_0(s) = q_0 s\left(1 - \frac{1}{2c}s\right), \tag{4.1.2}$$

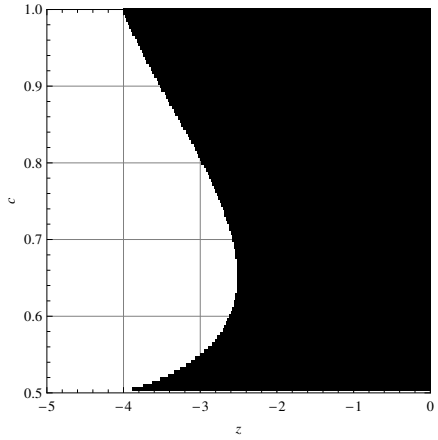

Figure 4.1: Region of stability in the (c, z)-plane for the two-step methods (3.2.1) with $m = 1$ and $p = 3$.

where q_0 is a real parameter. The free parameter q_0 can next be exploited in order to enforce high stability. Solving the order conditions (3.2.15) corresponding to $m = 1$ and $p = 2$, where $\varphi_0(s)$ is given by (4.1.2), we obtain a two-parameter family of methods depending on the parameter q_0 and the abscissa c. The coefficients of these formulae are given by

$$\begin{aligned}
\varphi_1(s) &= 1 - q_0 s + \frac{q_0}{2c} s^2, \\
\chi(s) &= \left(c + \frac{q_0}{2} + c q_0 \right) s - \left(\frac{1}{2} + \frac{q_0}{2} + \frac{q_0}{4c} \right) s^2, \\
\psi(s) &= \left(1 - c + \frac{q_0}{2} - q_0 c \right) s + \left(\frac{1}{2} + \frac{q_0}{2} - \frac{q_0}{4c} \right) s^2,
\end{aligned}$$

and the error constant $C_2(1)$ takes the form

$$C_2(1) = \frac{10c - 24c^2 + 12c^3 + q_0 - 2q_0 c - 6q_0 c^2 + 12q_0 c^3}{24c}.$$

The stability polynomial of this family of methods is

$$p(\omega, z) = \omega\big(p_2(z)\omega^2 + p_1(z)\omega + p_0(z)\big), \tag{4.1.3}$$

where the polynomials $p_0(z)$, $p_1(z)$ and $p_2(z)$ are now given by

$$p_0(z) = 2q_0 - 4q_0c + (2c - 4c^2 + 2c^3 + q_0 - 2q_0c - q_0c^2 + 2q_0c^3)z,$$
$$p_1(z) = -4c - 2q_0 + 4q_0c - (6c - 8c^2 + 4c^3 - q_0 + 2q_0c - 2q_0c^2 + 4q_0c^3)z,$$
$$p_2(z) = 4c - c^2(4 - 2c + q_0 - 2q_0c)z.$$

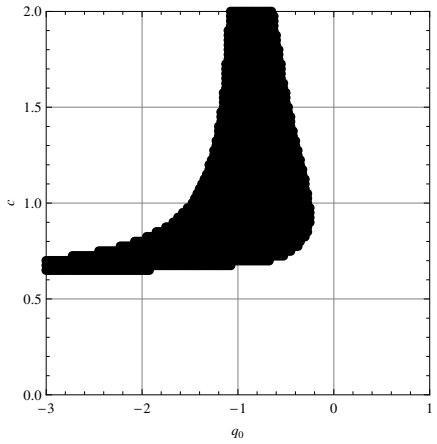

Figure 4.2: Region of A-stability in the (q_0, c)-plane for the two-step methods (3.2.1) with $m = 1$ and $p = 2$.

We have performed a computer search based on the Schur criterion using the polynomial $p(\omega, z)$ given by (4.1.3) with $p_0(z)$, $p_1(z)$ and $p_2(z)$ defined above. This search was performed in the parameter space (q_0, c) and the results are presented in Fig. 4.2 for $-3 \le q_0 \le 1$ and $0 \le c \le 2$, where the shaded region corresponds to the A-stable formulae. Choosing, for example, $q_0 = -1$ and $c = \frac{3}{4}$ we obtain the order 2 A-stable two-step almost collocation method (3.2.1) with coefficients given by

$$\varphi_0(s) = \frac{(2s-3)s}{3}, \quad \varphi_1(s) = \frac{3 + 3s - 2s^2}{3},$$
$$\chi(s) = \frac{(2s-3)s}{6}, \quad \psi(s) = \frac{(2s+3)s}{6}.$$

For this method the stability polynomial $p(w, z)$ is given by

$$p(w,z) = w\left(\left(3 - \frac{27}{16}z\right)w^2 - \left(4 + \frac{5}{8}z\right)w + \left(1 + \frac{5}{16}z\right)\right),$$

and the error constant is $C_2(1) = -\frac{17}{144}$.

We will look next for L-stable methods, i.e. methods for which all the roots of the polynomial $p(\omega, z)/p_2(z)$, where $p(\omega, z)$ is given by (4.1.3), are equal to zero as $z \to -\infty$. Such methods correspond to the solutions of the nonlinear system of equations

$$\lim_{z->-\infty} \frac{p_0(z)}{p_2(z)} = 0, \quad \lim_{z->-\infty} \frac{p_1(z)}{p_2(z)} = 0.$$

It can be verified that this system takes the form

$$\begin{cases} (c-1)(2c - 2c^2 + q_0 - q_0 c - 2q_0 c^2) = 0, \\ 6c - 8c^2 + 4c^3 - q_0 + 2q_0 c - 2q_0 c^2 + 4q_0 c^3 = 0, \end{cases}$$

and has solutions

$$q_0 = -\frac{2}{3}, \quad c = 1 \qquad \text{and} \qquad q_0 = -\frac{4}{9}, \quad c = 2.$$

The coefficients of the method corresponding to the first set of the above parameters are

$$\varphi_0(s) = \frac{(s-2)s}{3}, \quad \varphi_1(s) = \frac{3 + 2s - s^2}{3}, \quad \chi(s) = 0, \quad \psi(s) = \frac{(s+1)s}{3},$$

while the coefficients of the method related to the second set of the parameters q_0 and c are

$$\varphi_0(s) = \frac{s(s-4)}{9}, \ \varphi_1(s) = \frac{9 + 4s - s^2}{9}, \ \chi(s) = \frac{2(s-4)s}{9}, \ \psi(s) = \frac{(s-1)s}{9}.$$

It can be verified that for $s = 1$ both of the above methods reduce to backward differentiation method of order $p = 2$, compare [42],[147].

4.1.2 Analysis of methods with $m = 2$

We now consider two-step collocation methods (3.2.1) of maximum uniform order $p = 2m + 1 = 5$. Solving the system of order conditions (3.2.15) corresponding to $m = 2$ and $p = 5$ we obtain a family of methods depending on the abscissae c_1 and c_2. We have plotted in Fig. 4.3 the contour plots of error constant $C_5(1)$ of these formulas for $0 \le c_1 \le 1$ and $0 \le c_2 \le 1$. Choosing, for example, $c_1 = \frac{1}{2}$ and $c_2 = 1$ we obtain the two-step formula of

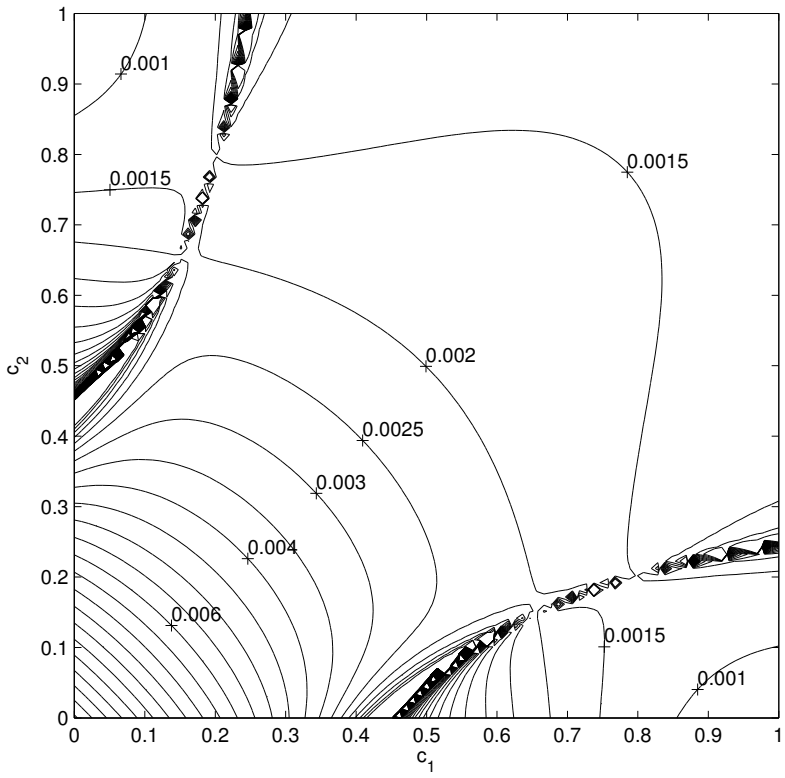

Figure 4.3: Contour plots of error constant $C_5(1)$ for $0 \leq c_1 \leq 1$ and $0 \leq c_2 \leq 1$.

uniform order $p = 5$ with coefficients given by

$$\varphi_0(s) = -\frac{(15 - 10s - 30s^2 + 24s^3)s^2}{29},$$
$$\varphi_1(s) = \frac{(1+s)(29 - 29s + 44s^2 - 54s^3 + 24s^4)}{29},$$
$$\chi_1(s) = -\frac{s^2(1+s)(89 - 187s + 96s^2)}{87},$$
$$\chi_2(s) = \frac{s(1+s)(29 - 31s - 16s^2 + 20s^3)}{29},$$
$$\psi_1(s) = \frac{s^2(1+s)(19 + 7s - 16s^2)}{29},$$
$$\psi_2(s) = -\frac{s^2(1+s)(7 - 2s - 12s^2)}{87}.$$

The error constant of this method is $C_5(1) = \frac{113}{83520}$. The stability polynomial of two parameter family of methods takes the form

$$p(\omega, z) = p_4(z)\omega^4 + p_3(z)\omega^3 + p_2(z)\omega^2 + p_1(z)\omega + p_0(z),$$

where $p_i(z)$, $i = 0, 1, 2, 3, 4$ are quadratic polynomials with respect to z. These polynomials depend also on c_1 and c_2. We have performed an extensive computer search based on Schur criterion in the two dimensional space (c_1, c_2) looking for methods with good stability properties but so far we were not able to find methods which are A-stable, because the regions of stability of such methods are all bounded, as it is illustrated in Fig. 4.4 for $m = 2$ and $p = 5$, where we have plotted, in the (c_1, z)-plane, the stability interval of the methods corresponding to each value of c_1, considering $c_2 = 1$ and taking $c_1 > \frac{-5+\sqrt{65}}{10}$, in order to satisfy the zero-stability requirement. We suspect that A-stable methods do not exist in the class of formulas with $m = 2$ and $p = 5$, also with respect to other values of c_2.

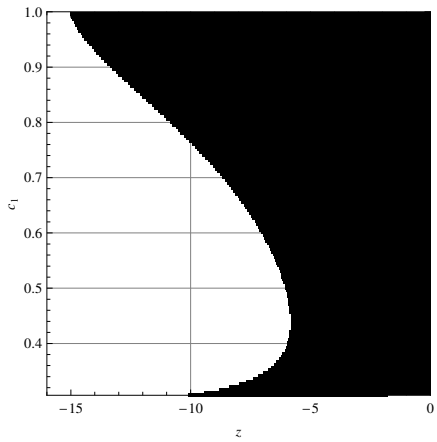

Figure 4.4: Region of stability in the (c_1, z)-plane for the two-step methods (3.2.1) with $m = 2$, $p = 5$ and $c_2 = 1$.

We consider next the methods of order $p = 2m = 4$. We choose the polynomial $\varphi_0(s)$ which satisfies the interpolation condition $\varphi_0(0) = 0$ and collocation conditions $\varphi_0'(c_i) = 0$, $i = 1, 2$. This leads to the polynomial of the form

$$\varphi_0(s) = s(q_0 + q_1 s + q_2 s^2 + q_3 s^3),$$

where q_2 and q_3 are given by

$$q_2 = -\frac{7q_0 + 6q_1}{3}, \quad q_3 = \frac{3q_0 + 2q_1}{2}.$$

The stability polynomial of this parameter family of methods of order $p = 4$ takes the form

$$p(\omega, z) = \omega(p_3(z)\omega^3 + p_2(z)\omega^2 + p_1(z)\omega + p_0(z)),$$

where $p_i(z)$, $i = 0, 1, 2, 3$ are quadratic polynomials with respect to z. These polynomials depend also on the parameters q_0, q_1, c_1, and c_2. We have performed an extensive computer search based on the Schur criterion in the four dimensional space (q_0, q_1, c_1, c_2) but so far we were not able to find methods which are A-stable. We suspect again that such methods do not exist in this class of formulas with $m = 2$ and $p = 4$.

Finally, we consider the two-step almost collocation methods of order $p = m + 1 = 3$. We choose the polynomials $\varphi_0(s)$ and $\chi_1(s)$ of degree less than or equal to three which satisfy conditions (3.2.25) and (3.2.26), i.e.,

$$\varphi_0(0) = 0, \quad \chi_1(0) = 0, \quad \varphi_0'(c_i) = 0, \quad \chi_1'(c_i) = 0, \quad i = 1, 2.$$

These polynomials take the form

$$\varphi_0(s) = s(q_0 + q_1 s + q_2 s^2), \quad \chi_1(s) = s(r_0 + r_1 s + r_2 s^2),$$

where

$$q_1 = -\frac{(c_1 + c_2)q_0}{2c_1c_2}, \quad q_2 = \frac{q_0}{3c_1c_2},$$

and

$$r_1 = -\frac{(c_1 + c_2)r_0}{2c_1c_2}, \quad r_2 = \frac{r_0}{3c_1c_2}.$$

Solving the order conditions (3.2.15) corresponding to $m = 2$ and $p = 3$ we obtain a four parameter family of methods (3.2.1) depending on q_0, r_0, c_1 and c_2. The stability polynomial of this family of methods is given by

$$p(\omega, z) = \omega^2(p_2(z)\omega^2 + p_1(z)\omega + p_0(z)),$$

where $p_i(z)$, $i = 0, 1, 2$, are polynomials of degree less than or equal to two with respect to z. These polynomials depend also on q_0, r_0, c_1 and c_2.

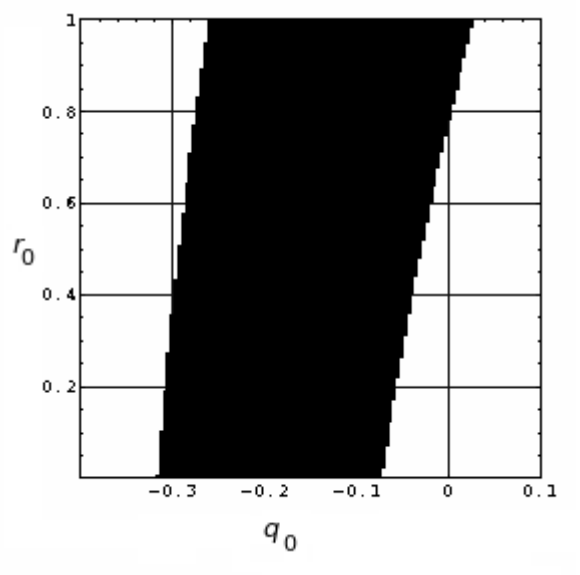

Figure 4.5: Region of A-stability in the (q_0, r_0)-plane for the two-step methods (3.2.1) with $m = 2$ and $p = 3$.

We have performed again an extensive computer search looking for methods which are A-stable. We have found such methods only if both components of the abscissa vector are outside of the interval $[0, 1]$. The results of this search for $c_1 = \frac{5}{2}$ and $c_2 = \frac{9}{2}$ are presented in Fig. 4.5 for $-0.4 \leq q_0 \leq 0.1$ and $0 \leq r_0 \leq 1$, where the shaded region corresponds to A-stable methods.

The coefficients of the resulting methods are given by

$$\varphi_0(s) = \frac{q_0 s(135 - 42s + 4s^2)}{135}, \quad \varphi_1(s) = \frac{135 - 135q_0 s + 42q_0 s^2 - 4q_0 s^3}{135},$$

$$\chi_1(s) = \frac{r_0 s(135 - 42s + 4s^2)}{135}, \quad \chi_2(s) = -\frac{(135 + 181q_0 - 36r_0)(135 - 42s + 4s^2)s}{1620},$$

$$\psi_1(s) = \left(\frac{63}{8} + \frac{241}{24}q_0 - 3r_0\right)s - \left(2 + \frac{1687}{540}q_0 - \frac{14}{15}r_0\right)s^2 + \left(\frac{1}{6} + \frac{241}{810}q_0 - \frac{4}{45}r_0\right)s^3.$$

$$\psi_2(s) = \left(\frac{35}{8} + \frac{145}{24}q_0 - r_0\right)s - \left(\frac{3}{2} + \frac{203}{108}q_0 - \frac{14}{45}r_0\right)s^2 + \left(\frac{1}{6} + \frac{29}{162}q_0 - \frac{4}{135}r_0\right)s^3.$$

The error constant $C_3(1)$ is

$$C_3(1) = \frac{4494825 + 6019723q_0 - 1229184r_0}{77760}.$$

We have also found methods in this class which are L-stable. Such methods correspond to solutions of the nonlinear system

$$\lim_{z\to-\infty} \frac{p_0(z)}{p_2(z)} = 0, \quad \lim_{z\to-\infty} \frac{p_1(z)}{p_2(z)} = 0.$$

One such solution is

$$q_0 \approx -\frac{21225899}{77647080} \approx -0.273364, \quad r_0 \approx \frac{113887980}{163068619} \approx 0.698405,$$

and the resulting method is A-stable and L-stable.

4.2 Continuous two-step Runge–Kutta methods

It is the purpose of this section to introduce new classes of continuous TSRK methods based on two-step almost collocation technique, of order $p = m$. The extensive analysis carried out in the following pages regards methods of the type (3.2.1) with $m = 1, 2, 3, 4$.

4.2.1 Construction of methods

The methods we aim to derive can be recognized as a class of two-step almost collocation methods, in the sense that, in order to advance from the point t_n to the point t_{n+1}, the continuous approximant (3.2.1) is derived by imposing some appropriate interpolation and collocation conditions, only with respect to the points t_n and $t_n + c_j h$, i.e.

$$P(t_n) = y_n, \tag{4.2.1}$$
$$P'(t_n + c_j h) = f(P(t_n + c_j h)), \quad j = 1, 2, \ldots, m. \tag{4.2.2}$$

These conditions imply that

$$\varphi_0(0) = 0, \quad \varphi_1(0) = 1, \quad \chi_j(0) = 0, \quad \psi_j(0) = 0, \tag{4.2.3}$$
$$\varphi\prime_0(c_i) = 0, \quad \varphi\prime_1(c_i) = 0, \quad \chi\prime_j(c_i) = 0, \quad \psi\prime_j(c_i) = \delta_{ij}, \tag{4.2.4}$$

for $i,\ j= 1,\ 2,\ \ldots,\ m$, where δ_{ij} is the usual Kronecker delta. We propose two different classes of continuous methods of this form. First, we impose the interpolation condition (4.2.1) only, obtaining a family of *interpolation based TSRK methods*. Then we follow the more general approach and impose the whole set of conditions (4.2.1), (4.2.2), achieving a class of *interpolation-collocation based TSRK methods*. The strategy we follow in the construction of these methods can be summarized as follows.

First of all, we fix the polynomials $\varphi_0(s)$, $\chi_j(s)$, j=1, 2, ..., m. Interpolation based TSRK methods with $m = 1$ (which can be exploited in a complete systematic way) are derived imposing

$$\varphi_0(s) = p_0 s, \quad \chi(s) = q_0 s, \tag{4.2.5}$$

with $p_0,\ q_0 \in \mathbb{R}$, while for interpolation-collocation TSRK methods with $m = 1$, we infer from (4.2.4) that

$$\varphi_0'(s) = \chi_j'(s) = (s - c),$$

therefore, if $\pi_1(s)$ is the primitive function of $\varphi_0'(s)$ and $\chi'(s)$ such that $\varphi_0(0) = \chi(0) = 0$, we impose

$$\varphi_0(s) = \alpha_0 \pi_1(s), \qquad \chi(s) = \beta_0 \pi_1(s), \tag{4.2.6}$$

with $\alpha_0,\ \beta_0 \in \mathbb{R}$. In the case $m \geq 2$, in order to carry out a more general analysis, we also ask for methods such that $\theta = \varphi_0(1) = 0$ and $u_j = \varphi_0(c_j) = 0,\ j = 1, 2, \ldots, m$. This choice is desirable in order to simplify the systems of order conditions, without loss in terms of stability and order, as Jackiewicz and Tracogna themselves stated in [140], where they first introduced TSRK methods. As a consequence of these choices, new conditions on $\varphi_0(s)$ arise, i.e.

$$\varphi_0(1) = 0, \quad \varphi_0(c_i) = 0, \quad i = 1, 2, \ldots, m, \tag{4.2.7}$$

which can be fulfilled in the following ways:

- for interpolation based methods, we choose $c_1 = 0,\ c_m = 1$, in order to obtain a family of FSAL methods with

  $$\varphi_0(s) = s(s - c_2) \cdots (s - c_{m-1})(s - 1), \quad \chi_j(s) = q_j s, \quad j = 1, 2, \ldots, m, \tag{4.2.8}$$

  with $q_j \in \mathbb{R}$, $j = 1, 2, \ldots, m$. The acronym FSAL stands for *first same as last* and identifies multistage methods such that the first stage of

the next step is the same as the last stage of the current step. It is well known that FSAL methods are suitable for efficient implementation (see [42, 147]);

- for interpolation-collocation based methods with $m \geq 2$, we impose

$$\varphi_0(s) = 0, \quad \chi_j(s) = \beta_j \pi_m(s), \quad j = 1, 2, \ldots, m, \tag{4.2.9}$$

with $\beta_j \in \mathbb{R}$, $j = 1, 2, \ldots, m$, where $\pi_m(s)$ is the primitive function of $\chi_j'(s)$ such that $\chi_j(0) = 0$.

As we aim for methods of order $p = m$, we impose the corresponding set of order conditions (3.2.15), and the unknown basis functions $\varphi_1(s)$, $\psi_j(s)$, j=1, 2, ..., m, can then be computed as solutions of the system (3.2.15).

We next compute the stability polynomial (3.2.44) of the obtained methods, in order to investigate the stability properties of the developed methods. It is possible to prove that, in correspondence of the above stated choices of the basis functions, the stability polynomial is of the type

$$p(\omega, z) = \omega^m (p_2(z)\omega^2 + p_1(z)\omega + p_0(z)), \tag{4.2.10}$$

and, therefore, the stability properties of the corresponding methods depend on the quadratic function (compare [72])

$$\tilde{p}(\omega, z) = p_2(z)\omega^2 + p_1(z)\omega + p_0(z). \tag{4.2.11}$$

Such an expression of the stability polynomial is suitable and desirable for many reasons, as it will be extensively described in Chapter 5. Intuitively, it is reasonable to think that we can succeed in finding A-stable methods implementing the Schür criterion in a symbolic environment to exactly determine its coefficients, only if the stability polynomial has low degree (e.g. at most 4). For TSRK methods, as it will be discussed in Chapter 5, it is natural to investigate the conditions to impose in order to force the stability properties to depend on a polynomial of degree 2, as discussed in [72], or on a linear polynomial. For this reason, we also discuss the conditions to accomplish in order to achieve the so-called Runge–Kutta stability (i.e. the stability matrix has one nonzero eigenvalues and, therefore, the stability properties of the corresponding methods depend on a linear polynomial).

Once we have obtained A-stability, we look for L-stable methods, by requiring that the parameters satisfy the nonlinear system of equations

$$\lim_{z \to -\infty} \frac{p_0(z)}{p_2(z)} = 0, \quad \lim_{z \to -\infty} \frac{p_1(z)}{p_2(z)} = 0. \tag{4.2.12}$$

In the following sections, we discuss the details of the construction of highly stable m-stage methods, with $m = 1, 2, 3, 4$.

4.2.2 Analysis of methods with $m = 1$

In this section we focus our attention on one-stage continuous methods of the form (3.2.1), having order $p = m = 1$. We first assume that the polynomial $P(t_n + sh)$ in (3.2.1) satisfies the interpolation condition (4.2.1) only. Correspondingly, according to the assumption (4.2.5), we fix

$$\varphi_0(s) = p_0 s, \quad \chi(s) = q_0 s, \tag{4.2.13}$$

and derive $\varphi_1(s)$ and $\psi(s)$ imposing the order conditions (3.2.15), obtaining

$$\varphi_1(s) = 1 - p_0 s, \quad \psi(s) = (1 + p_0 - q_0)s. \tag{4.2.14}$$

Therefore, the basis functions depend on the values of the parameters p_0 and q_0, which must be determined in order to achieve high stability properties. We know from the general theory of TSRK (see [140]), that a TSRK method is zero–stable if and only if $-1 < \theta \leq 1$: in our case, as $\theta = \varphi_0(1) = p_0$, it must be

$$-1 < p_0 \leq 1. \tag{4.2.15}$$

We next compute the stability function (3.2.44) of the method: in this case we have

$$p(\omega, z) = p_0(z) + p_1(z)\omega + p_2(z)\omega^2 \tag{4.2.16}$$

with

$$\begin{aligned} p_0(z) &= -p_0 - (1 - c)q_0 z, & (4.2.17)\\ p_1(z) &= -1 + p_0 + (-1 + c - p_0 + cp_0 + q_0 - 2cq_0)z, & (4.2.18)\\ p_2(z) &= 1 - c(1 + p_0 + q_0)z. & (4.2.19) \end{aligned}$$

In order to determine the values of the parameters p_0, q_0 and c achieving A-stability, we apply the Schur criterion to the stability function (4.2.16), obtaining the following result.

Theorem 4.2.1 *One-stage interpolation based continuous methods (3.2.1), with $p = m = 1$ and coefficients (4.2.13), (4.2.14) are A-stable if and only if*

the parameters p_0, q_0 and c satisfy the following system of inequalities

$$\begin{cases} -1 < p_0 < 1, \\ p_0 + 2c(1+p_0) > 1 + 2q_0, \\ (-1+2c)(1+p_0-2q_0) > 0, \\ (c+cp_0-q_0)\big(c(1+p_0-2q_0)+q_0\big) > 0. \end{cases} \tag{4.2.20}$$

Proof: In order to achieve A-stability, the roots of the polynomial (4.2.16) must lie inside the unit circle for any $z \in \mathbb{C}$ such that $\mathrm{Re}(z) < 0$. By the maximum principle, this is the case if

1. the polynomial (4.2.16) has no poles in the negative half plane;

2. the roots of $p(\omega, iy)$ are inside the unit circle $\forall y \in \mathbb{R}$.

Condition 1 is trivially satisfied. We analyse condition 2 by applying the Schur criterion to the polynomial $p(\omega, iy)$ that we will next denote as $p(\omega, y)$, therefore we first compute the polynomial

$$\hat{p}(\omega, y) = \bar{p_2}(z)\omega^2 + \bar{p_1}(z)\omega + \bar{p_0}(z)$$

where $\bar{p_2}(z)$, $\bar{p_1}(z)$ and $\bar{p_0}(z)$ are the complex conjugate polynomials associated to $p_2(z)$, $p_1(z)$ and $p_0(z)$ respectively. We next compute the polynomial

$$\alpha(\omega, y) = \frac{1}{\omega}(\hat{p}(0,y)\tilde{p}(\omega,y) - \tilde{p}(0,y)\hat{p}(0,y))$$

of degree 1. According to the Schur criterion, $p(\omega, y)$ is a Schur polynomial if and only if

$$|\hat{p}(0,y)| > |p(0,y)| \tag{4.2.21}$$

and $\alpha(\omega, y)$ is a Schur polynomial. Condition (4.2.21) is satisfied if and only if

$$1 - p_0^2 > 0, \qquad (c+cp_0-q_0)\big(c(1+p_0-2q_0)+q_0\big) > 0. \tag{4.2.22}$$

In order to investigate on the polynomial $\alpha(\omega, y)$, we apply the same procedure, i.e. we derive the corresponding polynomials $\hat{\alpha}(\omega, y)$ and

$$\beta(\omega, y) \;=\; \frac{1}{\omega}(\hat{\alpha}(0,y)\tilde{\alpha}(\omega,y) - \tilde{\alpha}(0,y)\hat{\alpha}(0,y)),$$

and the conditions imposed by the Schur criterion are satisfied for

$$-1 + p_0 + 2c(1 + p_0) - 2q_0 > 0, \qquad (-1 + 2c)(1 + p_0 - 2q_0) > 0. \quad (4.2.23)$$

Conditions (4.2.22) and (4.2.23) together give the system of inequalities (4.2.20). □

Fig. 4.1 shows some regions of A-stability in the parameter space (p_0, q_0), in correspondence of some values of the collocation abscissa c.

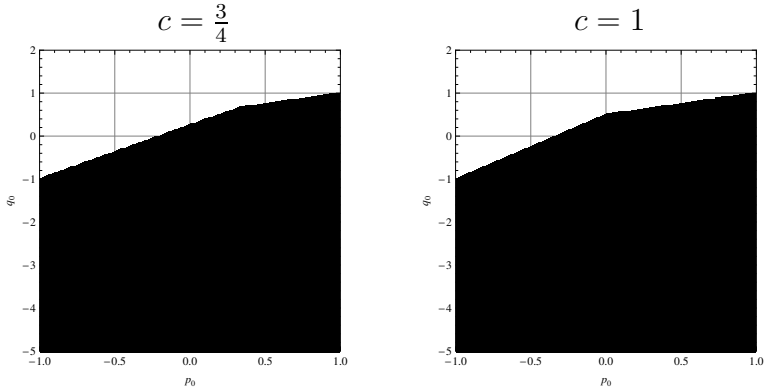

Figure 4.1: Regions of A-stability in the (p_0, q_0)-plane for two-step methods (3.2.1) with $p = m = 1$, for some values of the abscissa c.

We next look for L-stable methods: in this case, conditions (4.2.12) take the form

$$q_0(1 - c) = 0, \qquad 1 - c + p_0(1 - c) - q_0(1 - 2c) = 0,$$

whose solution is $(p_0, q_0) = (-1, 1)$, which is not acceptable because it violates the zero-stability requirement (4.2.15). However, if we set $c = 1$ and $q_0 = 0$, the above system is automatically satisfied, for any $p_0 \in (-1, 1]$, i.e. the corresponding methods are L-stable. In other words, if $c = 1$ and the basis polynomial $\chi(s)$ is identically zero, the resulting methods are all L-stable.

We now assume that the polynomial $P(t_n + sh)$ in (3.2.1) satisfies the whole set of conditions (4.2.3), (4.2.4), i.e.

$$\begin{array}{llll} \varphi_0(0) = 0, & \varphi_1(0) = 1, & \chi(0) = 0, & \psi(0) = 0, \\ \varphi_0'(c) = 0, & \varphi_1'(c) = 0, & \chi(c) = 0, & \psi(c) = 1. \end{array}$$

Correspondingly, we assume

$$\varphi_0(s) = \alpha\pi_1(s), \quad \chi(s) = \beta\pi_1(s), \tag{4.2.24}$$

and compute $\varphi_1(s)$ and $\psi(s)$ from the order conditions (3.2.15), obtaining

$$\varphi_1(s) = 1 + \alpha c s^2 - \frac{1}{2}\alpha s^3, \quad \psi(s) = s - (\alpha - \beta)cs^2 + \frac{1}{2}(\alpha - \beta)s^3. \tag{4.2.25}$$

As $\theta = \varphi_0(1) = \alpha(\frac{1}{2} - c)$, zero-stability is accomplished if and only if

$$-1 < \alpha(\frac{1}{2} - c) \leq 1. \tag{4.2.26}$$

We now investigate on the stability properties of the methods in analysis: following the lines drawn in Theorem 4.2.1, we obtain the following result.

Theorem 4.2.2 *One-stage interpolation-collocation based continuous methods (3.2.1), with $p = m = 1$ and coefficients (4.2.24)-(4.2.25) are A-stable if and only if the parameters α, β and c satisfy the following system of inequalities*

$$\begin{cases} \alpha^2(1-2c)^2 < 4, \\ (-\beta + 2c + \alpha c + 2\beta c - 2\alpha c^2)(\beta + 2c - \alpha c - 2\beta c + 2\alpha c^2 - 2\alpha c^3 + 2\beta c^3) > 0, \\ (-2 - \alpha + 2\alpha c)(2 + \alpha - 2\beta - 4c + 4\beta c - 4\alpha c^2 + 4\alpha c^3 - 4\beta c^3) > 0, \\ (-1 + 2c)(\alpha - 2(1+\beta) + 2\alpha c)(2 + \alpha(-1 + 2c)) < 0. \end{cases}$$

We have drawn some regions of A-stability in the parameter space (α, β), which result from a computer search based on the Schur criterion: Fig. 4.2 shows the results we have obtained for particular values of c.

We finally compute the values of α and β achieving L-stability, solving the system (4.2.12): it can be easily proved that those values are

$$\alpha = 2\frac{-1 + c + c^2}{1 - 2c - c^2 + 2c^3}, \quad \beta = \frac{2c}{-1 + c + 2c^2}.$$

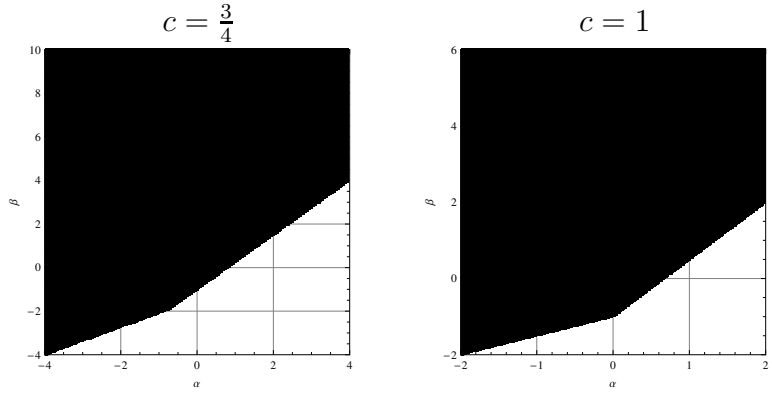

Figure 4.2: Regions of A-stability in the (α, β)-plane for interpolation-collocation based methods (3.2.1) with $p = m = 1$, for different values of the abscissa c.

4.2.3 Analysis of methods with $m = 2$

We now consider two-step continuous methods (3.2.1) with $p = m = 2$ and general abscissa vector, satisfying the interpolation condition (4.2.1). We assume

$$\varphi_0(s) = 0, \quad \chi_1(s) = q_1 s, \quad \chi_2(s) = q_2 s, \tag{4.2.27}$$

solve the system of order conditions (3.2.15) with respect to $\varphi_1(s)$, $\psi_1(s)$, $\psi_2(s)$, and compute the corresponding stability function (4.2.11), where $p_0(z)$, $p_1(z)$, $p_2(z)$ are polynomials of degree 2 with respect to z. In this case, the system (4.2.12) can be solved by setting $c_2 = 1$ and $q_2 = -q_1$. We finally apply the Schur criterion in order to localize the whole set of possible values of c_1 and q_1 such that the corresponding methods are L-stable. The results are shown in Fig. 4.3.

We next consider interpolation-collocation methods of type (3.2.1) with $p = m = 2$. According to assumption (4.2.9), we set

$$\varphi_0(s) = 0, \quad \chi_1(s) = \beta_1 \pi_2(s), \quad \chi_2(s) = \beta_2 \pi_2(s),$$

and derive $\varphi_1(s)$, $\psi_1(s)$ and $\psi_2(s)$ from the set of order conditions (3.2.15). We compute the stability function (4.2.11) where $p_0(z)$, $p_1(z)$ and $p_2(z)$ are polynomials of degree 2 with respect to z, and look for values of the parame-

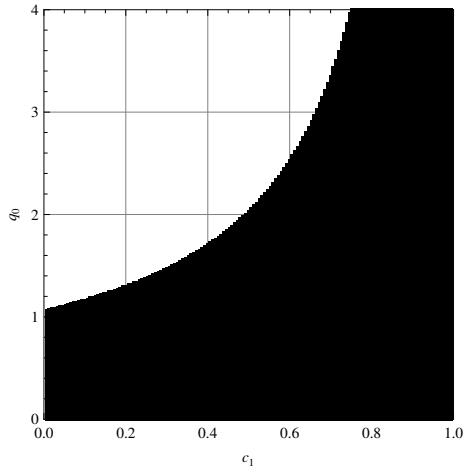

Figure 4.3: Region of L-stability in the (c_1, q_1)-plane for two-step methods (3.2.1) with $p = m = 2$ and $c_2 = 1$.

ters β_1 and β_2 achieving L-stability, by solving the system (4.2.12), obtaining

$$\beta_1 = \frac{3}{\gamma}(-1 + c_1)(-1 + c_2)^2, \quad \beta_2 = \frac{3}{\gamma}(-1 + 2c_1 - c_1^2 + c_2 - 2c_1c_2 + c_1^2c_2),$$

where

$$\gamma = 4 - 8c_1 + 5c_1^2 - c_1^3 - 8c_2 + 14c_1c_2 - 7c_1^2c_2 + 2c_1^3c_2 + 5c_2^2 - 7c_1c_2^2 - c_2^3 + 2c_1c_2^3.$$

Finally, we draw the L–stability region in the parameter space (c_1, c_2), performing a computer search based on the Schur criterion. Fig. 4.4 shows the result we obtained.

4.2.4 Construction of methods with $m = 3$

We now consider three stage continuous methods (3.2.1) with $p = m = 3$. Let us first derive FSAL interpolation based methods of order 3, corresponding to the abscissa vector $(c_1,\ c_2,\ c_3) = (0,\ c_2,\ 1)$. Following the assumptions in (4.2.8), we set

$$\varphi_0(s) = s(s - c_2)(s - 1), \quad \chi_1(s) = q_1 s, \quad \chi_2(s) = q_2 s, \quad \chi_3(s) = q_3 s,$$

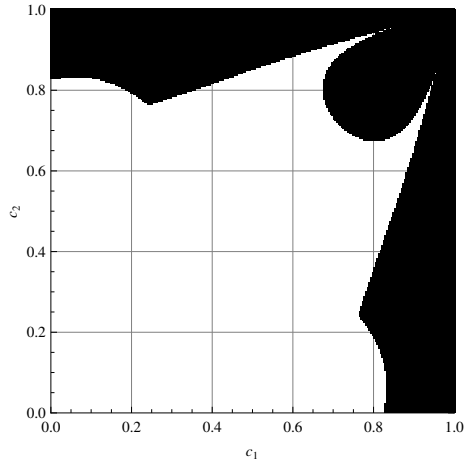

Figure 4.4: Region of L-stability in the (c_1, c_2)-plane for two-step methods (3.2.1) with $p = m = 2$.

and derive $\varphi_1(s),\ \psi_1(s),\ \psi_2(s),\ \psi_3(s)$, imposing the set of conditions (3.2.15). We omit for brevity the expression of the resulting basis polynomials, which can be easily recognized. We then derive the stability function (4.2.11), where $p_0(z),\ p_1(z),\ p_2(z)$ are polynomials of degree 3 with respect to z, and next look for the values of the parameters achieving L–stability. In this case, the solutions of the system (4.2.12) are

$$q_1 = -\frac{1 - c_2^2}{18c_2}, \quad q_2 = \frac{1 - 2c_2}{18c_2}.$$

By using the Schur criterion, it can be recognized that (4.2.11) is a Schur polynomial if and only if $c_2 > \frac{1}{2}$. We notice that q_3 has no influence in the construction, so we set it equal to 0.

We next develop interpolation-collocation based methods, assuming

$$\varphi_0(s) = 0, \quad \chi_1(s) = \beta_1\pi_3(s), \quad \chi_2(s) = \beta_2\pi_3(s), \quad \chi_3(s) = \beta_3\pi_3(s).$$

We impose the set of order conditions (3.2.15), derive $\psi_1(s)$, $\psi_2(s)$, $\psi_3(s)$, and compute β_1 and β_2 as solutions of the system (4.2.12), in order to gain L-stability. Finally, using the Schur criterion, we draw some regions of L–stability in the parameters space (c_1, c_2), for different values of β_3 and $c_3 = 1$, as shown in Fig. 4.5.

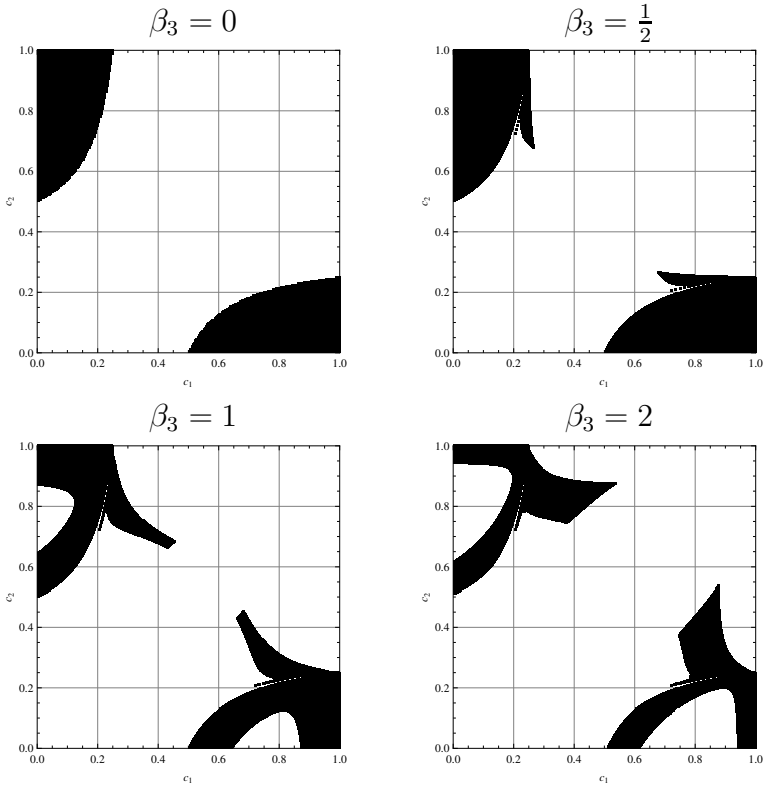

Figure 4.5: Regions of L-stability in the (c_1, c_2)-plane for two-step methods (3.2.1) with $p = m = 3$, $c_3 = 1$ and different values of the parameter β_3.

4.2.5 Construction of methods with $m = 4$

We now focus our attention on the development of two-step continuous methods (3.2.1) with $p = m = 4$, first considering interpolation based methods. According to assumptions (4.2.8), we fix

$$\varphi_0(s) = s(s - c_2)(s - c_3)(s - 1),$$
$$\chi_1(s) = sq_1, \quad \chi_2(s) = sq_2, \quad \chi_3(s) = sq_3, \quad \chi_4(s) = sq_4,$$

and derive $\varphi_1(s),\ \psi_1(s),\ \psi_2(s),\ \psi_3(s),\ \psi_4(s)$, imposing the set of order conditions (3.2.15) for $p = 4$. We omit for brevity the expression of the resulting basis polynomials, which can be easily recognized also in this case. We next derive the stability polynomial (4.2.11), where $p_0(z),\ p_1(z),\ p_2(z)$ are polynomials of degree 4 with respect to z, and look for the values of the parameters

achieving L–stability, solving the system (4.2.12), with respect to q_1 and q_2. Fig. 4.6 shows some regions of L-stability in the parameter space (c_3, q_3), for some values of c_2, drawn using the Schur criterion. We notice that q_4 does not play any rule in the derivation of the methods, so it can be put equal to zero.

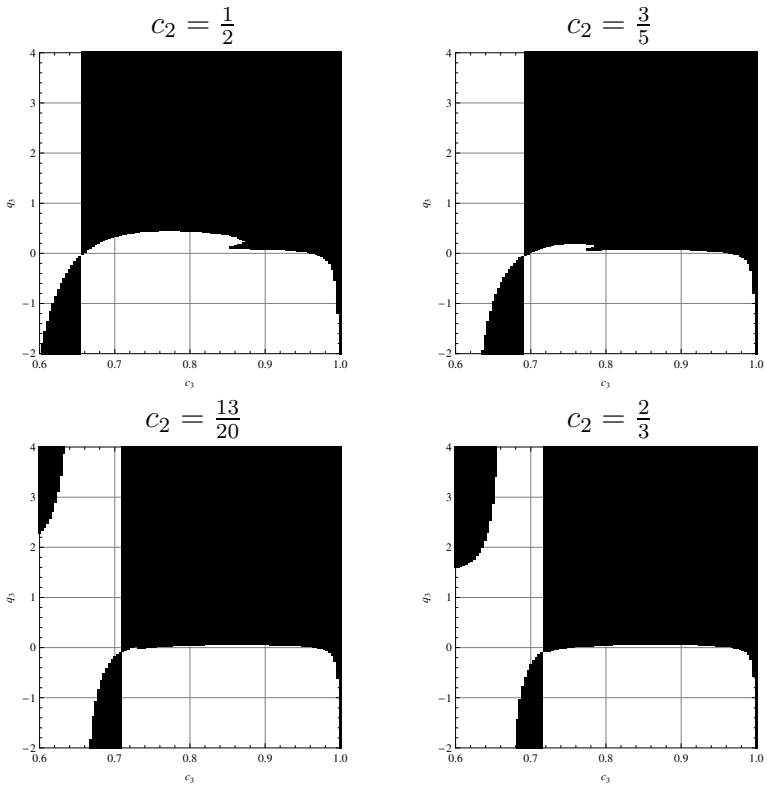

Figure 4.6: Regions of L-stability in the (c_3, q_3)-plane for two-step methods (3.2.1), for specific values of the abscissa c_2.

We now consider continuous four stage methods (3.2.1), obtained by imposing interpolation and collocation conditions (4.2.1), (4.2.2), and asking for $\theta = 0$, $u_j = 0$, $j = 1, 2, \ldots, m$. In line with assumptions (4.2.9), we impose $\varphi_0(s) = 0$ and

$$\chi_1(s) = \beta_1 \pi_4(s), \quad \chi_2(s) = \beta_2 \pi_4(s),$$
$$\chi_3(s) = \beta_3 \pi_4(s), \quad \chi_4(s) = \beta_3 \pi_4(s),$$

and, from the set of order conditions (3.2.15), we derive $\psi_1(s)$, $\psi_2(s)$, $\psi_3(s)$ and $\psi_4(s)$. We compute the stability function (4.2.11) and determine β_2 and β_3 as solutions of (4.2.12), in order to gain L-stability. Finally, using the Schur criterion, we draw some regions of L-stability in the parameters space (c_2, c_3), for different values of β_1, as shown in Fig. 4.7.

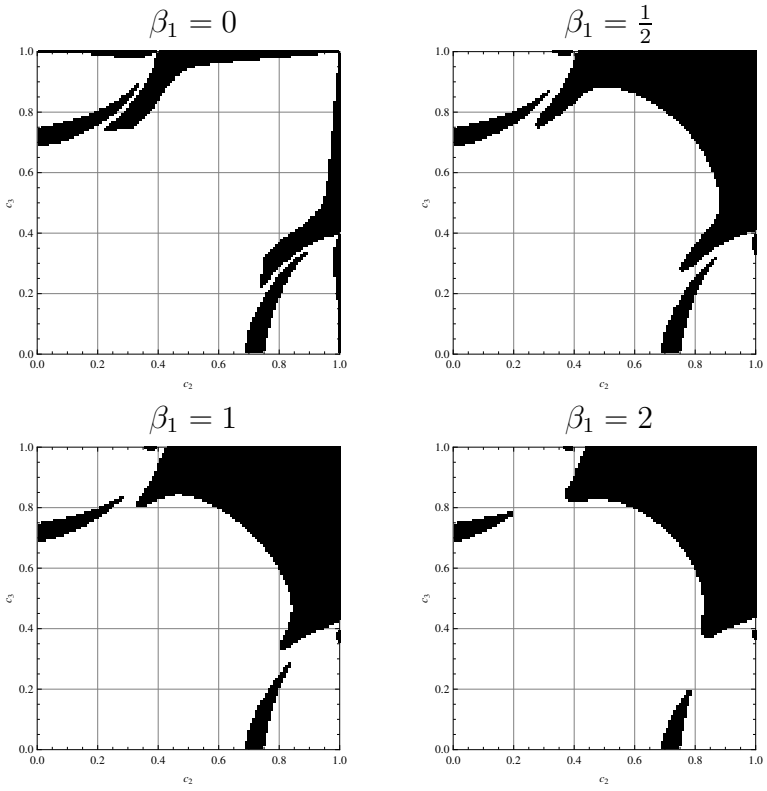

Figure 4.7: Regions of L-stability in the (c_2, c_3)-plane for two-step methods (3.2.1) with $p = m = 4$, for specific values of the parameter β_1.

4.2.6 Runge–Kutta stability

In this section we investigate the existence of continuous TSRK methods having the so-called Runge-Kutta stability, i.e. methods such that the stability matrix has one nonzero eigenvalue only, which is in general a very complicate requirement. We restrict our attention to one-stage continuous methods

derived in Section 4.2.2. We infer the following result.

Theorem 4.2.3 *For $c \in [0,1)$, the only interpolation based TSRK methods with $m = 1$ (compare Section 4.2.2) having Runge–Kutta stability are Runge–Kutta methods themselves. In the special case $c = 1$, $p_0 = 0$, all the corresponding methods posses Runge–Kutta stability, for any $q_0 \in \mathbb{R}$.*

Proof: It is sufficient to annihilate $p_0(z)$ in (4.2.17), in order to obtain a stability polynomial of the form

$$p(w,z) = w^2\Big(p_2(z)w + p_1(z)\Big), \tag{4.2.28}$$

having only one nonzero root. The conditions to impose in order to have $p_0(z)$ identically equal to 0 are

$$p_0 = 0, \quad (1-c)q_0 = 0.$$

The solution of this system is $p_0 = q_0 = 0$ and, correspondingly, the basis polynomials take the form

$$\varphi_0(s) = 0, \quad \varphi_1(s) = 1, \quad \chi(s) = 0, \quad \psi(s) = s,$$

and the continuous approximant (3.2.1) is

$$P(t_n + sh) = y_n + hsf\Big(P(t_n + ch)\Big). \tag{4.2.29}$$

The last part of the thesis is achieved simply setting $c = 1$ and $p_0 = 0$ in (4.2.17). □

With similar considerations, we can state an analogous result for interpolation-collocation TSRK methods with $m = 1$.

Theorem 4.2.4 *For $c \in [0,1]$ and $c \neq \frac{1}{2}$, the only interpolation-collocation methods with $m = 1$ (compare Section 4.2.2) having Runge-Kutta stability are Runge-Kutta methods themselves. For $c = \frac{1}{2}$ and $\alpha = \beta$, the corresponding TSRK methods have Runge–Kutta stability.*

4.2.7 Examples of methods

In this section we derive examples of A-stable and L-stable continuous TSRK methods (3.2.1) of order $p = m = 1, 2, 3, 4$, using the results derived in the previous sections. It is always assumed that $\theta = 0$ and $u_j = 0$, $j = 1, 2, \ldots, m$.

4.2.8 Examples of interpolation based methods

Example 1. Assuming $p = m = 1$, we set $c = \frac{3}{4}$ and obtain an A-stable method of type (3.2.1), with

$$\varphi_0(s) = 0, \quad \varphi_1(s) = 1, \quad \chi(s) = -s, \quad \psi(s) = 2s.$$

We notice that the L-stable method with $c = 1$, $q_0 = 0$ and $p_0 = 0$ is equivalent to the backward Euler method.

Example 2. Referring to the results derived in Section 4.2.3, the basis functions of the L-stable method of order $p = m = 2$ corresponding to the abscissa vector $c = [\frac{3}{4}, 1]^T$ are

$$\begin{aligned}&\varphi_0(s) = 0, \quad \varphi_1(s) = 1, \quad \chi_1(s) = s, \quad \chi_2(s) = -s,\\ &\psi_1(s) = s(9 - 2s), \quad \psi_2(s) = 2s(s - 1).\end{aligned}$$

We also show an example of A-stable FSAL method, setting $q_0 = 1$, $q_1 = 1$ and $q_2 = 0$, i.e.

$$\begin{aligned}&\varphi_0(s) = s(s - 1), \quad \varphi_1(s) = 1 + s - s^2, \quad \chi_1(s) = s, \quad \chi_2(s) = 0,\\ &\psi_1(s) = \tfrac{1}{2}s(-5 + 2s), \quad \psi_2(s) = \tfrac{9}{2}s.\end{aligned}$$

Example 3. Following the results contained in Section 4.2.4, we show the basis functions of a FSAL L-stable method with $p = q = m = 3$, corresponding to the abscissa vector $c = [0, \frac{3}{4}, 1]^T$:

$$\begin{aligned}&\varphi_0(s) = \tfrac{s}{4}(s - 1 + s)(4s - 3), \quad \varphi_1(s) = 1 - \tfrac{3}{4}s + \tfrac{7}{4}s^2 - s^3\\ &\chi_1(s) = -\tfrac{7s}{216}, \quad \chi_2(s) = -\tfrac{s}{27}, \quad \chi_3(s) = 0\\ &\psi_1(s) = s\Big(\tfrac{685}{216} - \tfrac{413}{72}s + \tfrac{55}{18}s^2\Big), \quad \psi_2(s) = -s\Big(\tfrac{101}{27} - \tfrac{94}{9}s + \tfrac{56}{9}s^2\Big)\\ &\psi_3(s) = s\Big(\tfrac{43}{18} - \tfrac{155}{24}s + \tfrac{25}{6}s^2\Big).\end{aligned}$$

Example 4. We next derive a FSAL L-stable method (3.2.1) with $p = m = 4$, corresponding to abscissa the vector $c = [0, \frac{1}{2}, \frac{3}{4}, 1]^T$. The basis functions of this method are

$$\begin{aligned}&\varphi_0(s) = \tfrac{s}{8}(s - 1)(2s - 1)(4s - 3), \quad \varphi_1(s) = 1 + \tfrac{3}{8}s - \tfrac{13}{8}s^2 + \tfrac{9}{4}s^3 - s^4,\\ &\chi_1(s) = -\tfrac{149}{264}s, \quad \chi_2(s) = \tfrac{233}{132}s, \quad \chi_3(s) = s, \quad \chi_4(s) = 0,\\ &\psi_1(s) = -s\Big(\tfrac{2435}{528} - \tfrac{117}{16}s + \tfrac{89}{8}s^2 - \tfrac{31}{6}s^3\Big)\\ &\psi_2(s) = s\Big(\tfrac{235}{33} - \tfrac{305}{12}s + \tfrac{205}{6}s^2 - \tfrac{46}{3}s^3\Big)\\ &\psi_3(s) = -s\Big(\tfrac{169}{33} - \tfrac{88}{3}s + \tfrac{112}{3}s^2 - 16s^3\Big)\\ &\psi_4(s) = s\Big(\tfrac{547}{528} - \tfrac{461}{48}s + \tfrac{289}{24}s^2 - \tfrac{29}{6}s^3\Big).\end{aligned}$$

4.2.9 Examples of interpolation-collocation based methods.

Example 5. Referring to the results derived in Section 4.2.2, we set $c = \frac{3}{4}$, obtaining a L-stable interpolation-collocation method with $p = m = 1$, whose basis functions are

$$\varphi_0(s) = s^2\Big(\tfrac{15}{7} - \tfrac{10}{7}s\Big), \quad \varphi_1(s) = 1 - \tfrac{15}{7}s^2 + \tfrac{10}{7}s^3,$$
$$\chi(s) = -s^2\Big(\tfrac{9}{7} - \tfrac{6}{7}s\Big), \quad \psi(s) = 1 + \tfrac{24}{7}s^2 - \tfrac{16}{7}s^3.$$

Example 6. Referring to the results derived in Section 4.2.3, we show the basis functions of the L-stable interpolation-collocation method (3.2.1) with $p = m = 2$, corresponding to the abscissa vector $c = [\frac{1}{2}, \frac{9}{10}]^T$, i.e.

$$\varphi_0(s) = 0, \quad \varphi_1(s) = 1,$$
$$\chi_1(s) = -s\Big(\tfrac{3}{4} - \tfrac{7}{60}s + \tfrac{1}{18}s^2\Big), \quad \chi_2(s) = -s\Big(\tfrac{3}{8} - \tfrac{7}{12}s + \tfrac{5}{18}s^2\Big),$$
$$\psi_1(s) = s\Big(\tfrac{69}{20} - \tfrac{187}{60}s + \tfrac{8}{9}s^2\Big), \quad \psi_2(s) = -s\Big(2 - \tfrac{29}{12}s + \tfrac{5}{9}s^2\Big).$$

Example 7. Following the results contained in Section 4.2.4 we show the basis functions of a FSAL L-stable interpolation-collocation method (3.2.1) with $p = m = 3$, corresponding to the abscissa vector $c = [0, \frac{3}{5}, 1]^T$:

$$\varphi_0(s) = 0, \quad \varphi_1(s) = 1, \quad \chi_1(s) = -s^3\Big(\tfrac{240}{337} - \tfrac{1280}{1011}s + \tfrac{200}{337}s^2\Big),$$
$$\chi_2(s) = s^3\Big(\tfrac{1125}{337} - \tfrac{2000}{337}s + \tfrac{1857}{674}s^2\Big), \quad \chi_3(s) = s^3\Big(\tfrac{3}{10} - \tfrac{8}{15}s + \tfrac{1}{4}s^2\Big),$$
$$\psi_1(s) = s\Big(1 - \tfrac{4}{3}s - \tfrac{113299}{30330}s^2 + \tfrac{115688}{15165}s^3 - \tfrac{14461}{4044}s^4\Big),$$
$$\psi_2(s) = 25s^2\Big(\tfrac{1}{12} + \tfrac{113}{6066}s - \tfrac{400}{3033}s^2 + \tfrac{125}{2022}s^3\Big)$$
$$\psi_3(s) = -s^2\Big(\tfrac{3}{4} - \tfrac{695}{2022}s - \tfrac{880}{1011}s^2 + \tfrac{275}{674}s^3\Big).$$

Example 8. We now consider the FSAL L-stable method (3.2.1) with $p = m = 4$, $\theta = 0$, $u_j = 0$, $j = 1, 2, \ldots, m$, corresponding to the abscissa vector $c = [0, \frac{7}{10}, \frac{9}{10}, 1]^T$, whose basis functions are

$$\varphi_0(s) = 0, \quad \varphi_1(s) = 1, \quad \chi_1(s) = -s^3\Big(\tfrac{63}{100} - \tfrac{223}{150}s + \tfrac{13}{10}s^2 - \tfrac{2}{5}s^3\Big),$$
$$\chi_2(s) = \tfrac{125840873}{10156165010}s^3\Big(189 - 446s + 390s^2 - 120s^3\Big),$$
$$\chi_3(s) = \tfrac{313000831}{6093699006}s^3(189 - 446s + 390s^2 - 120s^3), \quad \chi_4(s) = 0,$$
$$\psi_1(s) = s\Big(1 - \tfrac{223}{126}s - \tfrac{110596774973233}{9597575934450}s^2 + \tfrac{48055456715852}{1599595989075}s^3 - \tfrac{2838443145187}{106639732605}s^4 + \tfrac{873367121596}{106639732605}s^5\Big),$$
$$\psi_2(s) = s^2\Big(\tfrac{75}{7} - \tfrac{13154611771291}{639838395630}s + \tfrac{671254535668}{35546577535}s^2 - \tfrac{80390326549}{7109315507}s^3 + \tfrac{24735485092}{7109315507}s^4\Big),$$
$$\psi_3(s) = -s^2\Big(\tfrac{175}{9} - \tfrac{2867265551881}{54843291054}s + \tfrac{575594042414}{9140548509}s^2 - \tfrac{130770083795}{3046849503}s^3 + \tfrac{40236948860}{3046849503}s^4\Big),$$
$$\psi_4(s) = s^2\Big(\tfrac{21}{2} - \tfrac{28900702732187}{914054850900}s + \tfrac{2081690316751}{50780825050}s^2 - \tfrac{290054503193}{10156165010}s^3 + \tfrac{44623769722}{5078082505}s^4\Big).$$

4.3 Two-step modified collocation methods with reduced contribution of high order terms

This chapter concerns with the construction of two-step almost collocation methods (3.2.1) with narrowed contribution of the high order terms appearing in the local discretization error (3.2.34): more precisely, we aim to derive in the following sections methods such that the stage error constant $G_{p+1}(1)$ is equal to zero. This condition implies that terms of order $p+2$ only depend on the derivatives of the solution and not on the form of the equation. Moreover, this feature is of practical utility in the implementation of such methods in a variable stepsize-variable order, since it simplifies the order changing strategy. We present the construction of methods (3.2.1) of uniform order $p = m$ depending on up to four stages.

4.3.1 Construction of methods with $p = q = m = 1$

We now analyze two-step collocation methods (3.2.1) with $p = m = 1$, assuming that the collocation polynomial $P(t_n+sh)$ satisfies the interpolation condition

$$P(t_n) = y_n, \tag{4.3.1}$$

which implies that

$$\varphi_0(0) = 0, \quad \varphi_1(0) = 1, \quad \chi(0) = 0, \quad \psi(0) = 0.$$

Let us assume the following expression for the basis functions

$$\varphi_0(s) = p_0 + p_1 s, \quad \chi(s) = r_0 + r_1 s, \quad \varphi_1(s) = q_0 + q_1 s, \quad \psi(s) = s_0 + s_1 s.$$

Therefore, we have $p_0 = r_0 = s_0 = 0$ and $q_0 = 1$. We next impose the set of order conditions (3.2.15), obtaining

$$p_1 = -q_1, \qquad r_1 = 1 - q_1 - s_1.$$

Hence, the resulting family of one-stage methods (3.2.1) depends on q_1, s_1, and c, which must be determined in order to achieve high stability properties (e.g. A-stability and L-stability). We next consider the local truncation error (3.2.34) for one-stage methods of type (3.2.1) with $p = m = 1$, that is

$$\xi(t_n+sh) = h^2 E(s) y^{(2)}(t_n) + h^3 \Big(F(s) y^{(3)}(t_n) + G(s) \frac{\partial f}{\partial y}(y(t_n)) y^{(2)}(t_n) \Big) + O(h^4), \tag{4.3.2}$$

where $E(s) = C_1(s)$, $F(s) = C_2(s)$ and $G(s) = G_2(s)$ can be derived from formulae (3.2.35) and (3.2.36). In particular, the constant $G(s)$ takes the following form

$$G(s) = -\frac{c^2}{2}(-1 + q_1)(2 - c - q_1 + 2cq_1 - 2s_1)s.$$

Solving the equation $G(1) = 0$ with respect to s_1 we obtain

$$s_1 = \frac{2 - c - q_1 + 2cq_1}{2}.$$

As a consequence, the basis functions in (3.2.1) for $m = 1$, which now depend only on the parameter q_1 and the value of the abscissa c, take the following form

$$\begin{aligned} &\varphi_0(s) = -q_1 s, \qquad \chi(s) = -\frac{s}{2}(q_1 + 2cq_1 - c), \\ &\varphi_1(s) = 1 + q_1 s, \quad \psi(s) = -\frac{s}{2}(q_1 - 2cq_1 + c - 2). \end{aligned} \tag{4.3.3}$$

We next consider the linear stability analysis of this class of methods, first deriving the expression of the stability function (3.2.44) of these methods, which takes the form

$$p(\omega, z) = \omega\big(p_2(z)\omega^2 + p_1(z)\omega + p_0(z)\big), \tag{4.3.4}$$

where $p_0(z)$, $p_1(z)$ and $p_2(z)$ are polynomials of degree less than or equal 2 with respect to z. Applying the Schur criterion to the polynomial (4.3.4), we obtain the following result, which characterizes A-stable methods with $p = m = 1$.

Theorem 4.3.1 *Each one stage continuous method of type (3.2.1) which satisfies the restrictions discussed above is A-stable if and only if*

$$c > 1, \qquad \frac{c-1}{2c} \leq q_1 \leq 1. \tag{4.3.5}$$

Fig. 4.8 shows the corresponding region of A-stability in the parameter space (c, q_1). Let us provide an example of A-stable method. Setting $c = \frac{5}{4}$ and

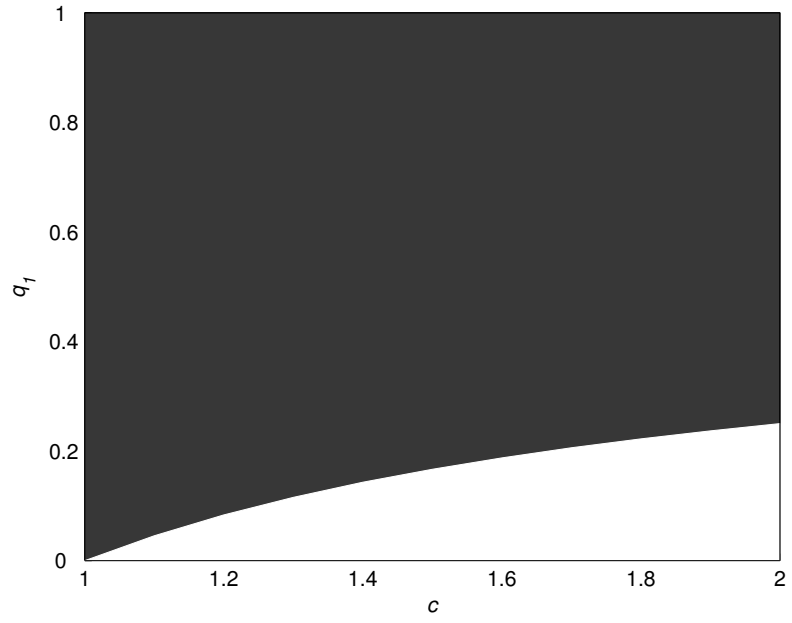

Figure 4.8: Region of A-stability in the (c, q_1)-plane for two-step methods (3.2.1) with $p = q = m = 1$ and satisfying the restrictions discussed above.

$q_1 = \frac{1}{2}$, the coefficients of the corresponding method (3.2.1) with $p = q = m = 1$ take the form

$$\varphi_0(s) = -\frac{s}{2}, \quad \chi(s) = -\frac{s}{4}, \quad \varphi_1(s) = \frac{2+s}{2}, \quad \psi(s) = \frac{3s}{4}. \tag{4.3.6}$$

The derived A-stable methods are also L-stable if

$$\lim_{z\to-\infty}\frac{p_0(z)}{p_2(z)}=0 \quad \text{and} \quad \lim_{z\to-\infty}\frac{p_1(z)}{p_2(z)}=0.$$

These conditions are equivalent to the system of equations

$$\begin{cases} (-1+c)(-c+q_1+2cq_1)=0, \\ 2-3c+2c^2-(1-2c+4c^2)q_1=0, \end{cases}$$

which has a unique solutions $(c,q_1)=(1,\frac{1}{3})$. Correspondingly, we obtain an L-stable method with

$$\varphi_0(s)=-\frac{s}{3},\quad \chi(s)=0,\quad \varphi_1(s)=\frac{3+s}{3},\quad \psi(s)=\frac{2s}{3}.$$

We also observe that the derived L-stable method achieve one order more of convergence since $E(1)=0$ and, therefore, its order is 2.

4.3.2 Construction of methods with $p=q=m=2$

We now consider two-stage continuous methods (3.2.1) with $p=q=m=2$. We always assume that $[c_1,c_2]=[\frac{1}{2},1]$. We next impose the interpolation condition (4.3.1), which leads to

$$\begin{array}{lll} \varphi_0(0)=0, & \chi_1(0)=0, & \chi_2(0)=0, \\ \varphi_1(0)=1, & \psi_1(0)=0, & \psi_2(0)=0. \end{array}$$

Correspondingly, we set

$$\varphi_0(s)=s(p_1+p_2s),\quad \chi_1(s)=s(q_1+q_2s),\quad \chi_2(s)=s(r_1+r_2s),$$

and derive $\varphi_1(s)$, $\psi_1(s)$ and $\psi_2(s)$ imposing the order conditions (3.2.15), obtaining

$$\begin{aligned} \varphi_1(s) &= 1-p_1s-p_2s^2, \\ \psi_1(s) &= s(2+3p_1-3q_1-2r_1-s+3p_2s-3q_2s-2r_2s), \\ \psi_2(s) &= -s(1+2p_1-2q_1-r_1-s+2p_2s-2q_2s-r_2s). \end{aligned}$$

This leads to a six-parameters family of methods depending on p_1, p_2, q_1, q_2, r_1 and r_2. These parameters will be chosen to obtain methods which are A-stable and L-stable. We next consider the linear stability analysis, deriving the stability polynomial (3.2.44)

$$p(\omega,z)=\omega(p_3(z)\omega^3+p_2(z)\omega^2+p_1(z)\omega+p_0(z)), \tag{4.3.7}$$

where

$$p_0(z) = (p_2 q_1 - p_1 q_2) z, \tag{4.3.8}$$

and $p_1(z)$, $p_2(z)$, $p_3(z)$ are polynomials of degree 2 with respect to z. However, imposing $q_1 = \frac{p_1 q_2}{p_2}$, the polynomial (4.3.8) annihilates and, correspondingly, the stability function (4.3.7) takes the form

$$p(\omega, z) = \omega^2(\tilde{p}_2(z)\omega^2 + \tilde{p}_1(z)\omega + \tilde{p}_0(z)),$$

where $\tilde{p}_0(z)$, $\tilde{p}_1(z)$, $\tilde{p}_2(z)$ are polynomials of degree 2 with respect to z. Therefore, the stability properties of the resulting methods depend on the quadratic function (see [72])

$$\tilde{p}(\omega, z) = \tilde{p}_2(z)\omega^2 + \tilde{p}_1(z)\omega + \tilde{p}_0(z). \tag{4.3.9}$$

We next impose the system of equations leading to L-stability, i.e.

$$\lim_{z\to-\infty} \frac{\tilde{p}_0(z)}{\tilde{p}_2(z)} = 0, \quad \lim_{z\to-\infty} \frac{\tilde{p}_1(z)}{\tilde{p}_2(z)} = 0.$$

This system takes the form

$$\begin{cases} p_2 r_1 - p_1 r_2 = 0, \\ p_2(-q_2 - r_1 + 3p_2 r_1 - 2q_2 r_1 - 2r_2) - p_1(q_2 + 3p_2 r_2 - 2q_2 r_2) = 0 \end{cases}$$

and has a unique solution given by

$$p_1 = -\frac{p_2(q_2 + 2r_2)}{q_2 + r_2}, \quad r_1 = -\frac{r_2(q_2 + 2r_2)}{q_2 + r_2}.$$

This leads to a three-parameter family of methods depending on $p_2,\ q_2,\ r_2$. We next apply the Schur criterion to determine the set of conditions involving these parameters, in order to be the corresponding methods L-stable. Let us fix, for example, q_2=2: we carry out a computer search of L-stable methods in the parameter space (p_2, r_2), using this criterion. The result is shown in Fig. 4.9.

We now consider the expression of the corresponding local truncation error (3.2.34)

$$\xi(t_n+sh) = h^3 E(s) y^{(3)}(t_n) + h^4 \Big(F(s) y^{(4)}(t_n) + G(s) \frac{\partial f}{\partial y}(y(t_n)) y^{(3)}(t_n) \Big) + O(h^5), \tag{4.3.10}$$

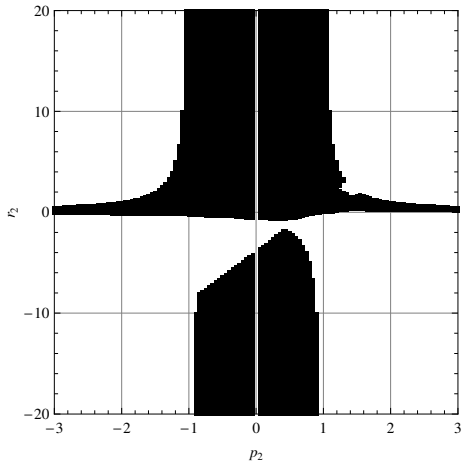

Figure 4.9: Regions of L-stability in the (p_2, r_2)-plane for two-step methods (3.2.1), with $q_2 = 2$ and $p = q = m = 2$.

where $E(s) = C_2(s)$, $F(s) = C_3(s)$ and $G(s) = G_3(s)$ can be derived from formulas (3.2.35) and (3.2.36). For $r_2 = 1$, we obtain

$$G(1) = 945 - 714p_2 + 133p_2^2,$$

and its roots are p_2=3, $p_2 = \frac{45}{19}$, but only the point ($\frac{45}{19}$,1) lies inside the shaded region of Fig. 4.9. The basis functions of the corresponding L-stable method take the form

$$\begin{aligned} \varphi_0(s) &= -\frac{60}{19}s + \frac{45}{19}s^2, \quad \varphi_1(s) = 1 + \frac{60}{19}s - \frac{45}{19}s^2, \\ \chi_1(s) &= -\frac{8}{3}s + 2s^2, \qquad \chi_2(s) = -\frac{4}{3}s + s^2, \\ \psi_1(s) &= \frac{182}{57}s - \frac{36}{19}s^2, \quad \psi_2(s) = -\frac{77}{57}s + \frac{24}{19}s^2. \end{aligned} \tag{4.3.11}$$

We observe that the error constant $E(1)$ is equal to 0 and, therefore, the above method has order 3 of convergence.

4.3.3 Construction of methods with $p = q = m = 3$

We now focus our attention on two-step continuous methods of order $p = q = m = 3$, assuming that $\varphi_0(s) = 0$ and imposing not only the interpolation

condition (4.3.1), but also the collocation condition

$$P'(t_n + c_i h) = f(t_n + c_i h, P(t_n + c_i h)), \tag{4.3.12}$$

for $i = 1, 2, 3$. We next assume $(c_1, c_2, c_3) = (\frac{1}{2}, \frac{3}{4}, 1)$ and

$$\begin{aligned} \chi_1(s) &= p_0 + p_1 s + p_2 s^2 + p_3 s^3 + p_4 s^4, \\ \chi_2(s) &= r_0 + r_1 s + r_2 s^2 + r_3 s^3 + r_4 s^4, \\ \chi_3(s) &= s_0 + s_1 s + s_2 s^2 + s_3 s^3 + s_4 s^4. \end{aligned}$$

Therefore, imposing the set of conditions (4.3.1) and (4.3.12), we obtain

$$\begin{aligned} \chi_1(s) &= \frac{s}{4}(-6p_4 + 13p_4 s - 12p_4 s^2 + 4p_4 s^3), \\ \chi_2(s) &= \frac{s}{4}(-6r_4 + 13r_4 s - 12r_4 s^2 + 4r_4 s^3), \\ \chi_3(s) &= \frac{s}{4}(-6s_4 + 13s_4 s - 12s_4 s^2 + 4s_4 s^3). \end{aligned}$$

We next compute the remaining basis functions by solving the system of order conditions (3.2.15) for $p = m = 3$, whose expressions are here omitted for brevity. We have now 3 free parameters (p_4, r_4 and s_4) to play with in order to achieve high stability properties. We next analyze the stability polynomial

$$p(\omega, z) = \omega^3\big(p_2(z)\omega^2 + p_1(z)\omega + p_0(z)\big),$$

where $p_0(z)$, $p_1(z)$ and $p_2(z)$ are polynomials of degree 3 with respect to z. The stability property of resulting methods now depend on the quadratic function

$$\tilde{p}(\omega, z) = p_2(z)\omega^2 + p_1(z)\omega + p_0(z). \tag{4.3.13}$$

We next solve the system of equations leading to L-stability

$$\lim_{z\to-\infty} \frac{\tilde{p}_0(z)}{\tilde{p}_2(z)} = 0, \quad \lim_{z\to-\infty} \frac{\tilde{p}_1(z)}{\tilde{p}_2(z)} = 0,$$

with respect to s_4, obtaining

$$s_4 = -\frac{1}{8}(4p_4 - 3r_4).$$

At this point, everything depends on the parameters p_4, r_4. We have next applied the Schur criterion to determine the set of conditions involving p_4, r_4,

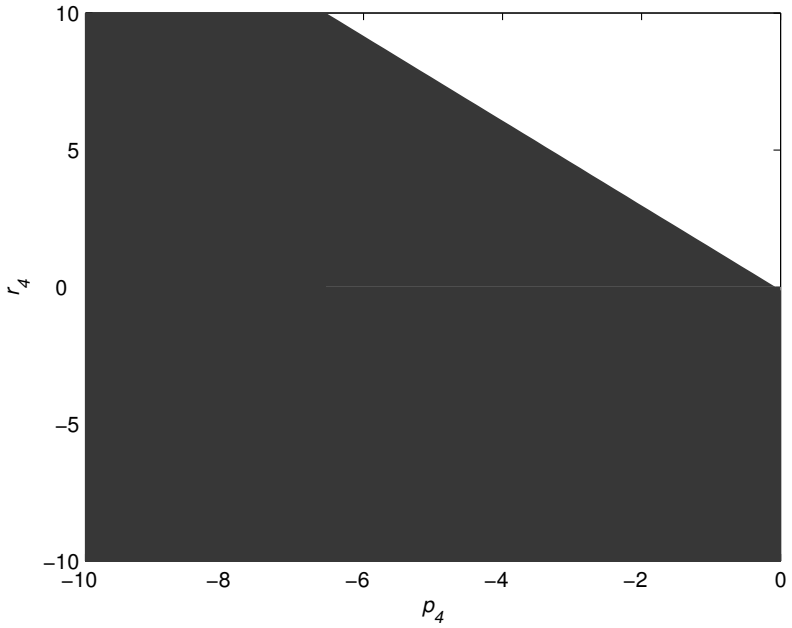

Figure 4.10: Regions of L-stability in the (p_4, r_4)-plane for two-step methods (3.2.1), with $p = q = m = 3$.

in order to be the corresponding methods L-stable and carried out a computer search of L-stable methods in the parameter space (p_4, r_4), according to these conditions. The results are given in fig. 4.10. We now consider the corresponding expression of the local truncation error (3.2.34) which is, in our case,

$$\xi(t_n+sh) = h^4 E(s) y^{(4)}(t_n) + h^5 \Big(F(s) y^{(5)}(t_n) + G(s) \frac{\partial f}{\partial y}(y(t_n)) y^{(4)}(t_n) \Big) + O(h^6), \tag{4.3.14}$$

where $E(s) = C_3(s)$, $F(s) = C_4(s)$ and $G(s) = G_4(s)$ are computed using formulas (3.2.35) and (3.2.36). In particular, the expression of $G_4(1)$ is

$$G(1) = -(784 + 60p_4 + 57r_4)(16 + 108p_4 + 69r4),$$

and it annihilates for

$$r_4 = -\frac{-4(196 + 15p_4)}{57}, \quad r_4 = -\frac{4(4 + 27p_4)}{69},$$

but only the first one is acceptable for us, because the line $r_4 = -\frac{4(4+27p_4)}{69}$ does not lie inside the L-stability region of fig. 4.10. Correspondingly, in order to achieve L-stability, we obtain from the Schur criterion that

$$\mu < p_4 < \frac{554}{21},$$

where μ is the negative root of the polynomial

$$\begin{aligned} q(x) &= -5045494028660092 - 341657542579860x - 5937877049931x^2 \\ &+ 80771782176x^3 + 136188864x^4. \end{aligned}$$

Moreover, we have

$$E(1) = \frac{554 - 21p_4}{912}, \quad F(1) = \frac{16841 - 315p_4}{54720},$$

and $E(1)$ annihilates in $p_4 = \frac{554}{21}$, which is not acceptable because it does not satisfy the L-stability constraint. However, we can choose a value of p_4 in order to have a small error constant. For example, taking $p_4 = 20$, we obtain the L-stable method with $E(1) = \frac{67}{456}$, $F(1) = \frac{10541}{54720}$, $G(1) = 0$, and

$$\begin{aligned} &\varphi_0(s) = 0, \quad \varphi_1(s) = 1, \quad \chi_1(s) = 5s(-6 + 13s - 12s^2 + 4s^3), \\ &\chi_2(s) = \frac{992}{19}s - \frac{6448}{57}s^2 + \frac{1984}{19}s^3 - \frac{1984}{57}s^4, \\ &\chi_3(s) = \frac{657}{19}s - \frac{2847}{38}s^2 + \frac{1314}{19}s^3 - \frac{438}{19}s^4, \\ &\psi_1(s) = -\frac{15594}{37}s + \frac{34129}{37}s^2 - \frac{31720}{37}s^3 + \frac{10624}{37}s^4, \\ &\psi_2(s) = \frac{6304}{19}s - \frac{13760}{19}s^2 + \frac{38432}{57}s^3 - \frac{4304}{19}s^4, \\ &\psi_3(s) = -114s + \frac{497}{2}s^2 - \frac{694}{3}s^3 + 78s^4. \end{aligned}$$

4.3.4 Construction of methods with $p = q = m = 4$

We now derive four-stage continuous methods (3.2.1) of order $p = q = m = 4$, with $\varphi_0(s) = 0$ and imposing the interpolation condition (4.3.1) and

the collocation condition (4.3.12) for $i = 1, 2, 3, 4$. We next set in advance the collocation abscissa $(c_1,\ c_2,\ c_3,\ c_4) = (0, \frac{1}{2},\ \frac{3}{4},\ 1)$ and

$$\begin{aligned}
\chi_1 &= s(p_1 + p_2 s + p_3 s^2 + p_4 s^3 + p_5 s^4);\\
\chi_2 &= s(q_1 + q_2 s + q_3 s^2 + q_4 s^3 + q_5 s^4);\\
\chi_3 &= s(r_1 + r_2 s + r_3 s^2 + r_4 s^3 + r_5 s^4);\\
\chi_4 &= s(s_1 + s_2 s + s_3 s^2 + s_4 s^3 + s_5 s^4);
\end{aligned}$$

where p_i, q_i, r_i, s_i, for $i = 1, 2, 3, 4$, are derived in order to satisfy the collocation conditions. We next derive $\psi_i(s)$, for $i = 1, 2, 3, 4$ imposing the set of order conditions (3.2.15), for $p = m = 4$. We omit their expressions for brevity. We have now 4 free parameters ($p_4,\ q_4,\ r_4$ and s_4) to compute in order to achieve high stability properties.

We next develop the linear stability analysis, studying the stability polynomial

$$p(\omega, z) = \omega^4\big(p_2(z)\omega^2 + p_1(z)\omega + p_0(z)\big),$$

where $p_0(z)$, $p_1(z)$, $p_2(z)$, $p_3(z)$ are polynomials of degree 4 with respect to z. Hence, the stability properties of the stability polynomial depend on the quadratic function

$$\tilde{p}(\omega, z) = p_2(z)\omega^2 + p_1(z)\omega + p_0(z).$$

We next solve the conditions for L-stability

$$\lim_{z\to-\infty} \frac{\tilde{p}_0(z)}{\tilde{p}_2(z)} = 0, \quad \lim_{z\to-\infty} \frac{\tilde{p}_1(z)}{\tilde{p}_2(z)} = 0,$$

with respect to p_5 and r_5, obtaining

$$p_5 = \frac{3072 + 17792q_5 - 21s_5}{99840}, \quad r_5 = -\frac{3072 + 9472q_5 - 21s_5}{3120}.$$

At this point, everything depends on the parameters $q_5,\ s_5$. We now consider the expression of the local truncation error (3.2.34)

$$\xi(t_n+sh) = h^5E(s)y^{(5)}(t_n)+h^6\Big(F(s)y^{(6)}(t_n)+G(s)\frac{\partial f}{\partial y}(y(t_n))y^{(5)}(t_n)\Big)+O(h^7),$$

where $E(s) = C_4(s)$, $F(s) = C_5(s)$ and $G(s) = G_5(s)$ have been computed using (3.2.35) and (3.2.36). In particular, $G(1)$ takes the form

$$G(1) = -s_5(5768192 - 258048q_5 - 265631s_5),$$

and annihilates for $q_5 = \frac{5768192-265631s_5}{258048}$. With this position, it is possible to prove using the Schur criterion that, for any $s_5 < 0$, the corresponding method is A-stable and L-stable. If we choose $s_5 = -1$, we obtain

$$
\begin{aligned}
&\varphi_0(s) = 0, \quad \varphi_1(s) = 1,\\
&\chi_1(s) = -\frac{64995145}{16515072}s^2 + \frac{844936885}{74317824}s^3 - \frac{64995145}{5505024}s^4 + \frac{12999029}{3096576}s^5,\\
&\chi_2(s) = -\frac{30169115}{1376256}s^2 + \frac{392198495}{6193152}s^3 - \frac{30169115}{458752}s^4 + \frac{6033823}{258048}s^5,\\
&\chi_3(s) = \frac{17413015}{258048}s^2 - \frac{226369195}{1161216}s^3 + \frac{17413015}{86016}s^4 - \frac{3482603}{48384}s^5,\\
&\chi_4(s) = \frac{15}{16}s^2 - \frac{65}{24}s^3 + \frac{45}{16}s^4 - s^5,\\
&\psi_1(s) = s - \frac{58553863}{8257536}s^2 + \frac{602930779}{37158912}s^3 - \frac{42497543}{2752512}s^4 + \frac{8132507}{1548288}s^5,\\
&\psi_2(s) = -\frac{40474777}{196608}s^2 + \frac{533249989}{884736}s^3 - \frac{41392281}{65536}s^4 + \frac{8330885}{36864}s^5,\\
&\psi_3(s) = \frac{67081319}{258048}s^2 - \frac{877562171}{1161216}s^3 + \frac{67998823}{86016}s^4 - \frac{13691515}{48384}s^5,\\
&\psi_4(s) = -\frac{164345193}{1835008}s^2 + \frac{238305005}{917504}s^3 - \frac{497623099}{1835008}s^4 + \frac{11139847}{114688}s^5,
\end{aligned}
$$

with $E(1) = \frac{1}{36864}$, $F(1) = \frac{69411889}{1486356480}$, $G(1) = 0$.

4.4 Two-step modified collocation methods with structured coefficient matrices

It is well known (compare, for instance, [36, 42, 119, 122, 147]) that highly stable multistage integration methods are necessarily implicit. Because of the implicitness of such methods, the numerical solution of nonlinear systems of equations is strongly involved in the integration process and the computational cost of an implicit numerical method then strictly depends on the computational cost required to solve such nonlinear systems: for this reason, we focus our attention on the development and the analysis of highly stable continuous formulae within the class (3.2.1), depending on structured coefficient matrices. In fact, 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 a fast computation (e.g. in a parallel environment) can be provided.

In order to derive numerical methods having the mentioned features, we consider a modification of the two-step algebraic collocation technique described in the previous sections. Since two-step (almost) collocation methods (3.2.1) are equivalent to a suitable class of TSRK methods (2.2.1), it is opportune to spend few words concerning TSRK methods depending on structured coefficient matrices. We have already mentioned the TSRK methods can be represented according to the following tensor notation

$$\begin{cases} y_{n+1} = \theta y_{n-1} + \tilde{\theta} y_n + h\big((v^T \otimes I_d)F^{[n-1]} + (w^T \otimes I_d)F^{[n]}\big), \\ Y^{[n]} = (u \otimes I_d)y_{n-1} + \big((e-u) \otimes I_d\big)y_n + h\big((A \otimes I_d)F^{[n-1]} + (B \otimes I_d)F^{[n]}\big), \end{cases} \tag{4.4.1}$$

where $F^{[n]} = [f_1(Y_1^{[n]}), \ldots, f_1(Y_m^{[n]}), \ldots, f_d(Y_1^{[n]}), \ldots, f_d(Y_m^{[n]})]^T$, I_d is the identity matrix of dimension d, $e = [1, \ldots, 1]^T \in \mathbb{R}^m$ and $\otimes$ is the usual Kronecker tensor product. The computational cost of a TSRK method (4.4.1) is strongly related to the solution of the nonlinear system for the computation of $Y^{[n]}$, whose coefficient matrix depends on the matrix B. An efficient solution of such system could be provided if B takes a special structure (e.g. lower triangular or diagonal). Jackiewicz and Tracogna identified in [140] four different types of TSRK methods according to the structure of B:

- *methods of type 1 and 2*, with

$$B = \begin{bmatrix} \lambda & & & \\ b_{21} & \lambda & & \\ \vdots & & \ddots & \\ b_{m1} & b_{m2} & \ldots & \lambda \end{bmatrix},$$

  with $\lambda = 0$ and $\lambda \neq 0$ respectively, suitable to integrate nonstiff and stiff systems respectively in a serial computing environment;

- *methods of type 3 and 4*, with

$$B = \begin{bmatrix} \lambda & & \\ & \ddots & \\ & & \lambda \end{bmatrix},$$

with $\lambda = 0$ and $\lambda \neq 0$ respectively, suitable to integrate nonstiff and stiff systems respectively in a parallel computing environment.

In particular, if B is a full matrix, (4.4.1) requires the solution of a nonlinear system of dimension $md \times md$; if B is lower triangular, m successive nonlinear systems of dimension d must be solved while, in case of B diagonal, the solution of m independent nonlinear systems of dimension d must be provided. Moreover,

- in the case of type 2 methods, i.e. B lower triangular and one-point spectrum, if the nonlinear system in (4.4.1) is solved by means of Newton-type iterations, the stored LU-factorization of the coefficient matrix $I_d - h\lambda J_i^{[n]}$ can be repeatedly used for a certain number of iterations, where $J_i^{[n]}$ is the i-th column of the jacobian of $F^{[n]}$;
- if B is diagonal, a fast resolution of the nonlinear system in a parallel environment can be provided.

The purpose of this section is the derivation of highly-stable two-step almost collocation methods (3.2.1) equivalent to TSRK methods (4.4.1) of type 2 and 4 (see [140]), therefore developing families of diagonally implicit continuous methods, following the lines drawn in the discrete case [10, 28, 29, 35, 34, 42, 43, 44, 45, 122, 138, 163] which led to the classes of DIRK, SDIRK, SIRK and DIMSIMs methods.

We will next consider the analysis of methods with B triangular and diagonal respectively, together with the requirements to fulfill in order to gain the desired structure; we will then describe the procedure to follow in order to derive highly stable structured formulae (i.e. A-stable and L-stable) within the class (3.2.1) and provide some examples of such methods.

4.4.1 Two-step almost collocation methods with triangular coefficient matrix

We analyze in this section the class of two-step almost collocation methods (3.2.1) such that the matrix B results to be lower triangular. Since $B = \left(\psi_j(c_i)\right)_{i,j=1}^{m}$, the conditions to impose in order to enforce a special structure on B strongly involve the basis polynomials $\psi_j(s)$, $j = 1, 2, \ldots, m$. The following result holds.

Proposition 4.4.1 *The matrix B is lower triangular if and only if*

$$\psi_j(s) = \omega_j(s) \prod_{k=1}^{j-1} (s - c_k), \tag{4.4.2}$$

where $\omega_j(s)$ is a polynomial of degree less or equal than $p-j+1$, $j = 2, \ldots, m$ and p is the order of the method.

Proof: We suppose B lower triangular: as a consequence, $b_{ij} = 0$ for $i < j$ $(j = 2, \ldots, m)$ and, therefore, $\psi_j(c_i) = 0$ for $i < j$ $(j = 2, \ldots, m)$. This implies that $c_1, c_2, \ldots, c_{j-1}$ are roots of $\psi_j(s)$, $j = 2, \ldots, m$. Hence, $\psi_j(s)$ can be factorized in the form (4.4.2), where $\omega_j(s)$ is a polynomial of degree $\deg(\psi_j(s)) - j + 1$, $j = 2, \ldots, m$. However, we infer from the system of order conditions (3.2.15), that $\deg(\psi_j(s)) \leq p$, where p is the order of the method. This completes the if part. The only if part is trivial. □

We next analyze the order of convergence of the resulting methods. In accordance with Proposition 4.4.1, the system of order conditions (3.2.15) can be specialized to the case of methods with B lower triangular, as reported in the following result.

Theorem 4.4.1 *A two-step collocation method (3.2.1) equivalent to a TSRK method (4.4.1) with B lower triangular has order p if and only if*

$$\begin{cases} \varphi_0(s) + \varphi_1(s) = 1, \\ \dfrac{(-1)^k}{k!}\varphi_0(s) + \displaystyle\sum_{j=1}^{m} \left(\chi_j(s) \frac{(c_j - 1)^{k-1}}{(k-1)!} + \frac{c_j^{k-1}}{(k-1)!} \sum_{\ell=0}^{p-j+1} \alpha_\ell^{(j)} s^\ell \prod_{r=1}^{j-1} (s - c_r) \right) = \dfrac{s^k}{k!}, \end{cases} \tag{4.4.3}$$

with $s \in [0, 1]$, $k = 1, 2, \ldots, p$ and $\alpha_\ell^{(j)} \in \mathbb{R}$, $\ell = 0, \ldots, p-j+1$, $j = 1, \ldots, m$.

Proof: Replace the expression (4.4.2) for $\psi_j(s)$ in (3.2.15). □

The real parameters $\alpha_\ell^{(j)} \in \mathbb{R}$, $\ell = 0, \ldots, p - j + 1$, $j = 1, \ldots, m$, can be regarded as free parameters which add degrees of freedom and can be used in order to enforce the corresponding methods to be highly stable (i.e. A-stable or L-stable), as it will be discussed in Section 4.4.3, where the practical construction of highly stable formulae within the discussed classes of methods is pointed out.

As a consequence of the above results, we can state the following corollary regarding the order of convergence of the derived formulae.

Corollary 4.4.1 *A two-step collocation method (3.2.1) equivalent to a TSRK method (4.4.1) with B lower triangular has uniform order of convergence at most equal to $m+2$.*

Proof: The order conditions (4.4.3) form a system of $p+1$ equations in $m+3$ unknowns which is compatible if $p = m+2$. □

The result contained in Corollary 4.4.1 provides an improvement with respect to diagonally-implicit Runge–Kutta methods, whose effective order is m, where m is the number of stages: this is due to the fact that Runge–Kutta methods usually do not have high stage order and, therefore, they suffer from the order reduction phenomenon (see, for instance, [42]) in the integration of stiff systems. Two-step almost collocation methods (3.2.1), instead, have high stage order $q = p$ overall the integration interval and, for this reason, they do not suffer from order reduction in the integration of stiff systems. Numerical evidences confirming this theoretical expectation are provided in Chapter 7.

4.4.2 Two-step almost collocation methods with diagonal coefficient matrix

We now consider the properties of two-step almost collocation methods (3.2.1) equivalent to TSRK methods (4.4.1) with B diagonal, presenting the main results following the lines drawn in the previous section. We first provide the analytical expression of the polynomials $\psi_j(s)$, $j = 1, 2, \ldots, m$, which enforces the diagonal structure of B.

Proposition 4.4.2 *The matrix B is diagonal if and only if*

$$\psi_j(s) = \omega_j(s) \prod_{\substack{k=1 \\ k \neq j}}^{m} (s - c_k), \tag{4.4.4}$$

where $\omega_j(s)$ is a polynomial of degree less or equal than $p - m + 1$, $j = 1, 2, \ldots, m$ and p is the order of the method.

When the matrix B is diagonal, i.e. when the functions $\psi_j(s)$ assume the expression (4.4.4), the set of order conditions (3.2.15) takes the following form.

Theorem 4.4.2 *A two-step collocation method (3.2.1) equivalent to a TSRK method (4.4.1) with B diagonal has order p if and only if*

$$\begin{cases} \varphi_0(s) + \varphi_1(s) = 1, \\ \dfrac{(-1)^k}{k!}\varphi_0(s) + \displaystyle\sum_{j=1}^{m} \left(\chi_j(s) \frac{(c_j - 1)^{k-1}}{(k-1)!} + \frac{c_j^{k-1}}{(k-1)!} \sum_{\ell=0}^{p-m+1} \mu_\ell^{(j)} s^\ell \prod_{r=1}^{j-1} (s - c_r) \right) = \dfrac{s^k}{k!}, \end{cases} \tag{4.4.5}$$

with $s \in [0,1]$, $k = 1, 2, \ldots, p$ and $\mu_\ell^{(j)} \in \mathbb{R}$, $\ell = 0, \ldots, p-m+1$, $j = 1, \ldots, m$.

Also in this case, the real parameters $\mu_\ell^{(j)} \in \mathbb{R}$, $\ell = 0, \ldots, p - m + 1$, $j = 1, \ldots, m$, can be regarded as degrees of freedom to use in order to obtain highly stable methods: this is the object of investigation in Section 4.4.3. We conclude with the following result concerning the order of convergence of the considered methods, which is a direct consequence of Theorem 4.4.2.

Corollary 4.4.2 *A two-step collocation method (3.2.1) equivalent to a TSRK method (4.4.1) with B diagonal has uniform order of convergence at most equal to $m + 1$.*

4.4.3 Construction of highly stable formulae

We now focus our attention on the procedures to follow in order to construct highly stable two-step almost collocation methods (3.2.1) of order p corresponding to TSRK methods (4.4.1) with B lower triangular or diagonal, according to the considerations reported in the previous sections. The derivation of highly stable methods is a nontrivial task, especially if we ask to create a reasonable balance between high effective order and strong stability properties. For this reason, in order to obtain highly-stable methods, we neglect some order conditions, obtaining some free parameters to be used to gain A-stability and L-stability. The number r of refused order conditions is what we call *relaxation index*. In the remainder of this section, we address the aspects regarding the construction of A-stable and L-stable methods of order $p = m + 2 - r$, with $r = 0, 1, 2$ (i.e. methods of order $m + 2$, $m + 1$ or m), within the class (3.2.1) corresponding to TSRK methods (4.4.1) with B lower triangular, and of order $p = m + 1 - r$, with $r = 0, 1$ (i.e. methods of order $m + 1$ or m) and B diagonal.

- *Construction of methods with B lower triangular*

First of all, we distinguish the following cases:

i. if $r = 0$, we assume $\psi_j(s)$, $j = 2, \dots, m$, of the form (4.4.2) with

$$\omega_j(s) = \alpha_0^{(j)} + \alpha_1^{(j)} s + \ldots + \alpha_{p-j+1}^{(j)} s^{p-j+1}; \tag{4.4.6}$$

ii. if $r = 1$, we consider $\omega_j(s)$, $j = 2, \dots, m$, of the form (4.4.6) and set

$$\varphi_0(s) = \beta_0 + \beta_1 s + \ldots + \beta_p s^p; \tag{4.4.7}$$

iii. if $r = 2$, we consider $\omega_j(s)$ and $\varphi_0(s)$ of the form (4.4.6) and (4.4.7) and set

$$\psi_1(s) = \gamma_0 + \gamma_1 s + \ldots + \gamma_p s^p. \tag{4.4.8}$$

We impose some interpolation and/or collocation conditions, chosen from the following sets:

$$\begin{array}{ll} \varphi_0(-1) = 1, & \varphi_0'(c_{i-1}) = 0, \\ \varphi_0(0) = 0, & \varphi_0'(c_i) = 0, \end{array} \tag{4.4.9}$$

$$\begin{array}{ll} \varphi_1(-1) = 0, & \varphi_1'(c_{i-1}) = 0, \\ \varphi_1(0) = 1, & \varphi_1'(c_i) = 0, \end{array} \tag{4.4.10}$$

$$\begin{array}{ll} \chi_j(-1) = 0, & \chi_j'(c_{i-1}) = \delta_{ij}, \\ \chi_j(0) = 0, & \chi_j'(c_i) = 0, \end{array} \tag{4.4.11}$$

$$\begin{array}{ll} \psi_j(-1) = 0, & \psi_j'(c_{i-1}) = 0, \\ \psi_j(0) = 0, & \psi_j'(c_i) = \delta_{ij}, \end{array} \tag{4.4.12}$$

where δ_{ij} is the usual Kronecker delta, $i, j = 1, 2, \dots, m$. Then, we derive the values of some $\alpha_\ell^{(j)}$, β_i, γ_i, for $\ell = 0, 1, \dots, p - j + 1$ and $i = 0, 1, \dots, p$, in such a way that the chosen conditions are satisfied on the fixed functions $\omega_j(s), \varphi_0(s)$ and/or $\varphi_1(s)$, according to the value of the relaxation index r. We next solve the system of order conditions (3.2.15) up to p, with respect to the remaining basis functions: they automatically inherit the same interpolation/collocation conditions imposed, as proved in Theorem 3.2.9

If in addition to the triangular structure for B we also require it to be one point spectrum, i.e. we ask for two-step almost collocation formulae (3.2.1) equivalent to type 2 TSRK methods (4.4.1), we spend some of the remaining

free parameters within the set of $\alpha_\ell^{(j)}$ and γ_i, for $\ell = 0, 1, \ldots, p-j+1$ and $i = 0, \ldots, p$, in order to equal all the values on the diagonal of B, i.e. all the $\psi_j(c_j)$, $j = 1, \ldots, m$, equal to a correspond to a real common value λ.

We next compute the stability polynomial

$$p(\eta, z) = \sum_{k=0}^{m+2} p_k(z)\eta^k, \tag{4.4.13}$$

of degree $m+2$ with respect to η, where each $p_k(z)$ is a rational function in z, $k = 0, 1, \ldots, m+2$. If possible, we spend some of the free parameters in order to reduce the degree of $p(\eta, z)$ with respect to η (e.g. to obtain a quadratic stability polynomial, compare [72]). Let us suppose that the resulting degree of $p(\eta, z)$ with respect to η is ν. Using the Schür criterion, we determine the values of the remaining parameters corresponding to A-stable methods. If the corresponding set of A-stable methods is nonempty, we search for the related subset of L-stable methods, solving the nonlinear system

$$\begin{cases} \lim\limits_{z\to-\infty} \dfrac{p_0(z)}{p_\nu(z)} = 0, \\ \qquad\vdots \\ \lim\limits_{z\to-\infty} \dfrac{p_{\nu-1}(z)}{p_\nu(z)} = 0. \end{cases} \tag{4.4.14}$$

- *Construction of methods with B diagonal*

In order to obtain a diagonal shape for the matrix B, together with strong stability properties, we proceed as follows. We distinguish the following cases:

i. if $r = 0$, we assume $\psi_j(s)$, $j = 1, \ldots, m$, of the form (4.4.4) with

$$\omega_j(s) = \mu_0^{(j)} + \mu_1^{(j)} s + \ldots + \mu_{p-m+1}^{(j)} s^{p-m+1}; \tag{4.4.15}$$

ii. if $r = 1$, we consider $\omega_j(s)$, $j = 1, \ldots, m$, of the form (4.4.15) and set

$$\varphi_0(s) = \sigma_0 + \sigma_1 s + \ldots + \sigma_p s^p. \tag{4.4.16}$$

We next impose some interpolation and/or collocation conditions on the functions fixed above, chosen from the sets (4.4.9) and (4.4.12), i.e. we derive the values of some $\mu_\ell^{(j)}$, σ_i, for $\ell = 0, 1, \ldots, p-m+1$ and $i = 0, 1, \ldots, p$,

in such a way that these conditions are satisfied. Then, we solve the system of order conditions (3.2.15) up to p, with respect to the remaining basis functions: the computed polynomials automatically inherit the same interpolation/collocation conditions imposed.

If in addition to the diagonal structure for B we also require it to be one point spectrum, i.e. we ask for two-step almost collocation formulae (3.2.1) equivalent to type 4 TSRK methods (4.4.1), we spend some of the remaining free parameters within the set of $\mu_\ell^{(j)}$, for $\ell = 0, 1, \ldots, p-m+1$, in order to equal all the values on the diagonal of B, i.e. all the $\psi_j(c_j)$, $j = 1, \ldots, m$, equal to a correspond to a real common value λ.

We next compute the stability polynomial (4.4.13) of degree $m+2$ with respect to η, where $p_k(z)$ is a rational function in z, $k = 0, 1, \ldots, m+2$. If possible, we spend some of the free parameters in order to reduce the degree of $p(\eta, z)$ with respect to η. Let us suppose that the resulting degree of $p(\eta, z)$ with respect to η is ρ. Using the Schür criterion, we determine the values of the remaining parameters corresponding to A-stable methods. If the corresponding set of A-stable methods is nonempty, we search for the related subset of L-stable methods, solving the nonlinear system

$$\begin{cases} \lim\limits_{z\to-\infty} \dfrac{p_0(z)}{p_\rho(z)} = 0, \\ \qquad\vdots \\ \lim\limits_{z\to-\infty} \dfrac{p_{\rho-1}(z)}{p_\rho(z)} = 0. \end{cases} \tag{4.4.17}$$

4.4.4 Analysis of methods with $m = 2$ and B lower triangular

We first aim to derive two-stage highly stable methods (3.2.1) having the maximum attainable continuous order $p = m + 2 = 4$, with B lower triangular. Since $r = 0$, we assume $\omega_2(s)$ of type (4.4.6), i.e.

$$\omega_2(s) = \alpha_0 + \alpha_1 s + \alpha_2 s^2 + \alpha_3 s^3 \tag{4.4.18}$$

and, therefore, $\psi_2(s)$ assumes the form

$$\psi_2(s) = (\alpha_0 + \alpha_1 s + \alpha_2 s^2 + \alpha_3 s^3)(s - c_1). \tag{4.4.19}$$

We next impose the interpolation condition $\psi_2(0) = 0$ and the collocation conditions $\psi_2'(c_1) = 0$ and $\psi_2'(c_2) = 1$, obtaining

$$\alpha_0 = 0,\ \alpha_1 = \frac{c_1(-1 + 2c_1^2c_2\alpha_3 - 6c_1c_2^2\alpha_3 + 4c_2^3\alpha_3)}{c_1^2 - 4c_1c_2 + 3c_2^2},$$
$$\alpha_2 = \frac{1 - c_1^3\alpha_3 + 2c_1^2c_2\alpha_3 + 3c_1c_2^2\alpha_3 - 4c_2^3\alpha_3}{c_1^2 - 4c_1c_2 + 3c_2^2}.$$

At this point, three free parameters are left, i.e. c_1, c_2 and α_3. We next compute the stability polynomial (4.4.13), which assumes the form

$$p(\eta, z) = \eta\big(p_0(z) + p_1(z)\eta + p_2(z)\eta^2 + p_3(z)\eta^3\big).$$

Applying the Schür criterion, it is possible to prove that no A-stable methods with $m = 2$ and $p = 4$ exist and, therefore, we relax one order condition, in order to find highly stable methods within the class (3.2.1) with $m = 2$ and $p = m + 1 = 3$, corresponding to TSRK methods (4.4.1) with B lower triangular. Since $r = 1$, we assume $\omega_2(s)$ of the form

$$\omega_2(s) = \alpha_0 + \alpha_1 s + \alpha_2 s^2, \tag{4.4.20}$$

and $\varphi_0(s)$ of the type

$$\varphi_0(s) = \beta_0 + \beta_1 s + \beta_2 s^2 + \beta_3 s^3. \tag{4.4.21}$$

We next impose the interpolation and collocation conditions

$$\varphi_0(0) = 0,\ \varphi_0'(c_2) = 0,\ \psi_2(0) = 0,\ \psi_2'(c_2) = 1,$$

obtaining

$$\alpha_0 = 0, \quad \alpha_2 = \frac{1 + c_1\alpha_1 - 2c_2\alpha_1}{c_2(-2c_1 + 3c_2)}, \quad \beta_0 = 0, \quad \beta_3 = -\frac{\beta_1 + 2c_2\beta_2}{3c_2^2}.$$

As a consequence, a five-parameter family of methods (3.2.1) arises: the degrees of freedom are $c_1, c_2, \alpha_1, \beta_1, \beta_2$. We next compute the stability polynomial (4.4.13), which assumes the form

$$p(\eta, z) = \eta\big(p_0(z) + p_1(z)\eta + p_2(z)\eta^2 + p_3(z)\eta^3\big),$$

and compute the values of β_2 and c_2 annihilating $p_0(z)$, in such a way that the stability properties of the related methods depend on the quadratic stability function

$$\tilde{p}(\eta, z) = p_1(z) + p_2(z)\eta + p_3(z)\eta^2. \tag{4.4.22}$$

These values are

$$\beta_2 = \frac{-\beta_1(c_1 + 6\alpha_1 - 6c_1\alpha_1 - 2c_1^2\alpha_1 + 2c_1^3\alpha_1)}{2c_1\alpha_1(3 - 5c_1 + 2c_1^2)}, \quad c_2 = 1.$$

We next apply the Schür criterion on the polynomial (4.4.22) in correspondence of

$$\beta_1 = \frac{6c_1\alpha_1(-1 + c_1)}{2 + 3c_1}$$

achieving L-stability, i.e. solving the system (4.4.23) which, for the polynomial (4.4.22), takes the form

$$\begin{cases} \lim\limits_{z\to-\infty} \dfrac{p_1(z)}{p_3(z)} = 0, \\ \lim\limits_{z\to-\infty} \dfrac{p_1(z)}{p_3(z)} = 0. \end{cases} \tag{4.4.23}$$

Applying the Schür criterion we are able to find the values of the remaining free parameters c_1 and α_1 corresponding to A-stable and, in particular, L-stable methods. The results are given in Figure 4.1. The region provided in Figure 4.1 arises from the union of the following sets:

$$\begin{aligned}
\Sigma_1 &= \left\{(c_1,\alpha_1) \in \mathbb{R}^2 : \frac{1+\sqrt{17}}{4} < c_1 < \frac{3}{2},\ \frac{1}{5 - 9c_1 + 4c_1^2} < \alpha_1 < \frac{1 + c_1 - c_1^2}{(-1+c_1)^3}\right\}, \\
\Sigma_2 &= \left\{(c_1,\alpha_1) \in \mathbb{R}^2 : \frac{3}{2} < c_1 \leq \frac{1+\sqrt{5}}{2},\ \frac{1 + c_1 - c_1^2}{(c_1-1)^3} < \alpha_1 < \frac{1}{4c_1^2 - 9c_1 + 5}\right\}, \\
\Sigma_3 &= \left\{(c_1,\alpha_1) \in \mathbb{R}^2 : c_1 > \frac{1+\sqrt{5}}{2},\ \frac{-c_1^2 + c_1 + 1}{(c_1-1)^3} < \alpha_1 < 0\right\}, \\
\Sigma_4 &= \left\{(c_1,\alpha_1) \in \mathbb{R}^2 : c_1 > \frac{1+\sqrt{5}}{2},\ 0 < \alpha_1 < \frac{1}{4c_1^2 - 9c_1 + 5}\right\}.
\end{aligned}$$

If we aim for two-step almost collocation methods (3.2.1) equivalent to type 2 TSRK methods (4.4.1), we apply the same procedure above described but, instead of spending some parameters to reduce the degree of the stability polynomial, we use them to obtain equal values on the diagonal of the matrix B. For this reason, after computing $\alpha_0,\ \alpha_2,\ \beta_0,\ \beta_3$ as above, we determine the value of β_2 such that $b_{11} = b_{22}$, obtaining

$$\beta_2 = \frac{-3(2c_1^3 - 6(1+\alpha_1) - 3c_1^2(3 + 2\alpha_1) + 3c_1(5 + 4\alpha_1))}{c_1(-1+3c_1)(3-2c_1)^2},$$

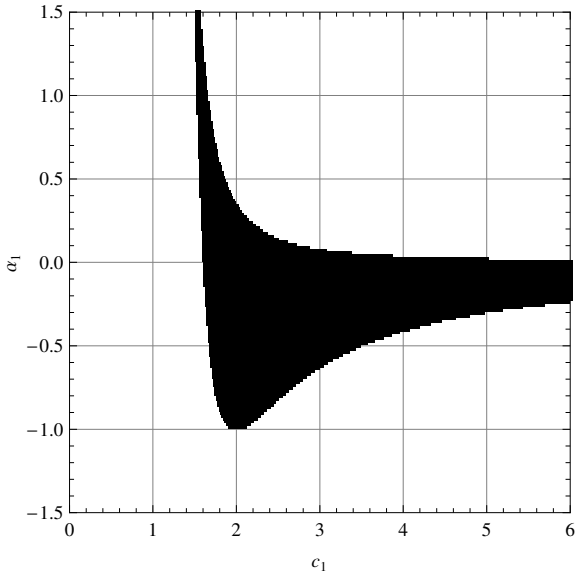

Figure 4.1: Region of L-stability in the (c_1, α_1)-plane for diagonally implicit two-step almost collocation methods (3.2.1), with $m = 2$ and $p = 3$

in correspondence of the values $c_2 = 1$ and $\beta_1 = 0$. This values are chosen in order to simplify the structure of the stability polynomial, whose stability properties are analyzed using the Schür criterion. The results of this analysis are reported in Figure 4.2.

4.4.5 Analysis of methods with $m = 2$ and B diagonal

We conclude this section deriving highly stable two-step almost collocation methods (3.2.1) with $m = 2$, equivalent to type 4 TSRK methods (4.4.1). It possible to prove, applying the Schür criterion, that no A-stable type 4 almost collocation methods (3.2.1) of order $m = 2$ and $p = 3$ exist and, therefore, we relax one order condition, searching for methods of order $p = m = 2$. We assume that

$$\omega_1(s) = \mu_0^{(1)} + \mu_1^{(1)} s, \quad \omega_2(s) = \mu_0^{(2)} + \mu_1^{(2)} s, \quad \varphi_0(s) = \sigma_0 + \sigma_1 s + \sigma_2 s^2,$$

and impose the collocation conditions

$$\varphi_0(0) = 0, \quad \psi_1(0) = 0, \quad \psi_2(0) = 0,$$

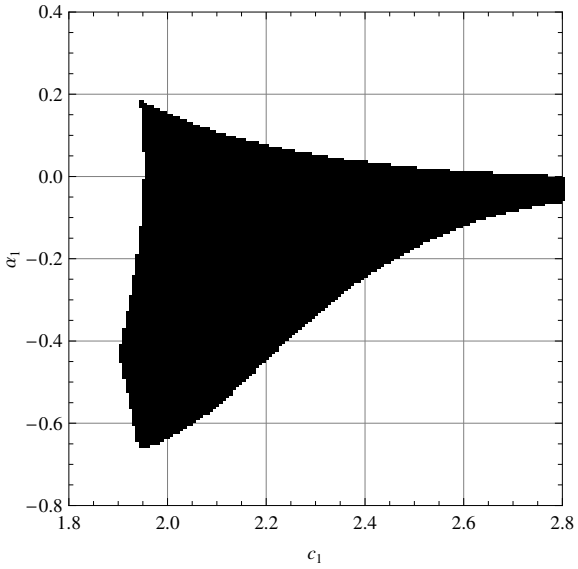

Figure 4.2: Region of A-stability in the (c_1, α_1)-plane for type 2 two-step almost collocation methods (3.2.1), with $m = 2$ and $p = 3$, for $c_2 = 1$ and $\beta_1 = 0$

obtaining $\mu_0^{(1)} = \mu_0^{(2)} = \sigma_0$. Six free parameters are left, i.e. $\mu_1^{(1)}, \mu_2^{(1)}, \sigma_1, \sigma_2, c_1$ and c_2. We set $\mu_1^{(1)} = \mu_2^{(1)} = 1$ and derive the values of σ_1, σ_2 and c_2 solving the system (4.4.17) for L-stability, obtaining

$$\sigma_1 = -\frac{2(7c_1^3 - 8c_1^2 - 9c_1 + 2)}{4c_1^2 - 3c_1 + 3}, \quad \sigma_2 = \frac{4c_1^3 + 4c_1^2 - 23c_1 + 3}{4c_1^2 - 3c_1 + 3}, \quad c_2 = 1.$$

Applying the Schür criterion we obtain that the resulting methods are L-stable if and only if $c_1 \in (-\frac{2}{5}, \frac{270}{619})$.

Chapter 5

Stability features for discrete TSRK methods

This chapter concerns with some new contributions we have provided to the analysis of the stability properties of TSRK methods (2.2.1). In particular, we have focused our attention on the following two issues:

- the recent literature concerning with the numerical solution of ODEs (1.1.1) has given particular emphasis to the special requirement of "practicality" of the numerical methods: in fact, together with the classical properties of consistency, high order of convergence, strong stability properties and so on, some authors are considering also the necessity to develop *practical* methods. The adjective "practical" has been used for the first time by J. Butcher and W. Wright in [50] and underlines the importance to deal with methods having good properties which also result, at the same time, realistic in terms of their derivation. For this purpose, it is also important to provide, together with the theoretical analysis of the features we aim to satisfy, also an algorithmic technique for the derivation of methods having those features. This work has been carried out in the context of General Linear Methods in the thesis of W. Wright [209], which has paved the way to a series of papers dealing with GLMs with *inherent Runge–Kutta stability* (IRKS), which is an algebraic property on the coefficients guaranteeing that the corresponding stability matrix has one nonzero eigenvalue only, which is a property typical of Runge–Kutta methods. As a consequence, the

stability function (2.1.18) of the resulting GLMs with IRKS take the form

$$p(\eta, z) = \eta^{r-1}(\eta - R(z)),$$

where $R(z)$ is the nonzero eigenvalue of the stability matrix. This property provides a remarkable improvement in the derivation of highly stable methods, because of the very easy expression of the stability polynomial. The purpose we aim to carry out in Section 5.1 is the derivation of an analogous property for TSRK methods, focusing our attention on the derivation of highly stable TSRK methods whose coefficient matrix has a one-point spectrum. We will see that the natural extension of Inherent Runge–Kutta Stability for TSRK methods will be connected to the analysis of a quadratic stability function, instead of a linear one;

- we have presented in Section 2.1.4 the importance of nonlinear stability, collocating things in the perspective of General Linear Methods. However nothing in the literature has appeared up to now concerning the analysis of the nonlinear stability properties of TSRK methods (2.2.1). Therefore, it is interesting to analyze the behaviour of TSRK methods with respect to more general test equations and consider the possibility to derive algebraically stable TSRK methods. The analysis, reported in Section 5.2 will be carried out by using the tools presented by A. Hill in [127], which is based on the analysis of the Nyquist function (2.1.31) associated to TSRK methods. This approach will also require some optimization tools, since we aim for the minimization of a certain objective function depending on the coefficient of the methods, as it is described in Section 5.2.

The results concerning the above features have been presented for the first time in our papers [72, 138, 93].

5.1 Two-Step Runge–Kutta Methods with Inherent Quadratic Stability

In this section we will describe the construction of TSRK methods (2.2.1) of order p and stage order $q = p$ whose stability properties are determined

by quadratic stability functions, also assuming that the coefficient matrix B has a one point spectrum

$$\sigma(B) = \{\lambda\}, \quad \lambda > 0. \tag{5.1.1}$$

This feature would allow the efficient implementation of such methods similarly as in the case of singly implicit Runge-Kutta (SIRK) methods considered by Burrage [28], Butcher [34], and Burrage, Butcher and Chipman [29], see also [36], [42]. For TSRK methods whose coefficient matrix B has a one point spectrum (5.1.1), it is more convenient to work with the function $p(\omega, z)$ defined by

$$p(\omega, z) = (1 - \lambda z)^s \tilde{p}(\omega, z), \tag{5.1.2}$$

since the coefficients of ω^i, $i = 0, 1, \ldots, s+2$, are polynomials of degree s with respect to z.

In this section we investigate TSRK methods (2.2.1) of order $p = s$ and stage order $q = p$ for which the stability polynomial $p(\omega, z)$ takes the simple form

$$p(\omega, z) = \omega^s \Big((1 - \lambda z)^s \omega^2 - p_1(z)\omega + p_0(z) \Big), \tag{5.1.3}$$

where $p_1(z)$ and $p_0(z)$ are polynomials of degree s with respect to z. Methods for which this is the case are said to possess quadratic stability (QS). The interest in TSRK methods having QS property lies in many reasons. The construction of high-order A-stable and L-stable TSRK methods is a challenging problem, because we have to deal with stability polynomials of degree $s + 2$, where s is the number of stages. This is often not manageable when s is high using the approach based on Schur criterion (see [147], [184]). The machinery we intend to develop ensures that the stability polynomial is *a priori* quadratic, which gives us the chance to succeed in finding highly stable methods also for a large number of stages. It is also important to observe that quadratic stability is the most natural requirement for TSRK methods, as it will be afterward explained. The characterization of such methods, inspired by the recent work on GLMs with IRKS [49], [50], [51], [209], [210], will be given in the following sections, together with the construction of A-stable and L-stable TSRK methods with $p = q = s$ for the number of stages $s = 1$, 2, 3, and 4.

5.1.1 Characterization of TSRK methods with quadratic stability

To investigate the form of the stability function of the method (2.2.1) it is convenient to introduce some equivalence relation between matrices of the same dimensions. We say that the two matrices D and E are equivalent, which will be denoted by $D \equiv E$, if they are equal except for the first two rows. This relation has several useful properties which will aid in the derivation of TSRK methods with appropriate stability properties. It can be verified that if $F \in \mathbb{R}^{(\nu+2)\times(\nu+2)}$ is a matrix partitioned as follows

$$F = \left[\begin{array}{c|c} F_{11} & F_{12} \\ \hline F_{21} & F_{22} \end{array}\right],$$

where $F_{11} \in \mathbb{R}^{2\times 2}$, $F_{12} \in \mathbb{R}^{2\times\nu}$, $F_{21} \in \mathbb{R}^{\nu\times 2}$, $F_{22} \in \mathbb{R}^{\nu\times\nu}$, and if $F_{21} = 0$, then $D \equiv E$ implies $FD \equiv FE$. Moreover, for any matrix F we have also $D \equiv E$ implies $DF \equiv EF$.

In general, it is a very complicated task to construct TSRK methods (2.2.1) which possess QS, especially for methods with large number of stages s, since this requires the solution of large systems of polynomial equations of high degree, for the unknown coefficients of the methods. However, if we are willing to restrict the class of methods, it is possible to find interrelations between the coefficients matrices which ensure that this is the case, i.e. that the TSRK method (2.2.1) possesses QS. Such conditions in the case of GLMs with Runge-Kutta stability were discovered recently by Butcher and Wright [50], [209]. They take a similar form for TSRK methods with QS. This is formalized in the following definition.

Definition 5.1.1 *The TSRK method (2.2.1), regarded as GLM with coefficients* $\mathbf{A}$, $\mathbf{U}$, $\mathbf{B}$, *and* $\mathbf{V}$ *defined by (2.2.3) has inherent quadratic stability (IQS) if there exists a matrix* $\mathbf{X} \in \mathbb{R}^{(s+2)\times(s+2)}$ *such that*

$$\mathbf{BA} \equiv \mathbf{XB}, \tag{5.1.4}$$

and

$$\mathbf{BU} \equiv \mathbf{XV} - \mathbf{VX}. \tag{5.1.5}$$

The significance of this definition follows from the following theorem.

Theorem 5.1.1 *Assume that the TSRK method (2.2.1) has IQS. Then its stability function $\tilde{p}(\omega, z)$ defined in (5.1.2) assumes the form*

$$\tilde{p}(\omega, z) = \omega^s \Big(\omega^2 - \tilde{p}_1(z)\omega + \tilde{p}_0(z)\Big), \tag{5.1.6}$$

where $\tilde{p}_1(z)$ and $\tilde{p}_0(z)$ are rational functions with respect to z.

Proof: The proof of this theorem follows along the lines of the corresponding result for GLMs with IRKS [50], [209]. Assuming $\mathbf{I_s} - z\mathbf{A}$ nonsingular, the IQS relation (5.1.4) is equivalent to

$$\mathbf{B} \equiv (\mathbf{I}_{s+2} - z\mathbf{X})\mathbf{B}(\mathbf{I}_s - z\mathbf{A})^{-1}. \tag{5.1.7}$$

To investigate the characteristic polynomial of the corresponding matrix $\mathbf{M}(z)$, as defined in (2.2.6), it is more convenient to consider the matrix related to $\mathbf{M}(z)$ by similarity transformation. Using (5.1.5) and (5.1.7) and assuming that $\mathbf{I}_{s+2} - z\mathbf{X}$ is nonsingular, we obtain

$$(\mathbf{I}_{s+2} - z\mathbf{X})\mathbf{M}(z)(\mathbf{I}_{s+2} - z\mathbf{X})^{-1} \equiv \mathbf{V}. \tag{5.1.8}$$

It follows from the structure of the matrix $\mathbf{V}$ in (2.2.3) and the relation (5.1.8) that the matrix $(\mathbf{I}_{s+2} - z\mathbf{X})\mathbf{M}(z)(\mathbf{I}_{s+2} - z\mathbf{X})^{-1}$ can be partitioned as follows

$$(\mathbf{I}_{s+2} - z\mathbf{X})\mathbf{M}(z)(\mathbf{I}_{s+2} - z\mathbf{X})^{-1} = \left[\begin{array}{c|c} \tilde{M}_{11}(z) & \tilde{M}_{12}(z) \\ \hline 0 & 0 \end{array}\right], \tag{5.1.9}$$

where $\tilde{M}_{11}(z) \in \mathbb{R}^{2\times 2}$, $\tilde{M}_{12}(z) \in \mathbb{R}^{2\times s}$, and 0 stands for zero matrix of dimension $s \times 2$ and $s \times s$, respectively. This relation implies that the characteristic polynomial $\tilde{p}(\omega, z)$ of the matrix

$$(\mathbf{I}_{s+2} - z\mathbf{X})\mathbf{M}(z)(\mathbf{I}_{s+2} - z\mathbf{X})^{-1}$$

and $\mathbf{M}(z)$ assumes the form (5.1.6). □

The proof of Theorem 5.1.1 also explains the reason why it is natural in the context of TSRK methods to investigate quadratic stability. In fact, it follows that the stability matrix $\mathbf{M}(z)$ and the coefficient matrix $\mathbf{V}$ are related by the equation (5.1.8), where the matrix $\mathbf{M}(z)$ satisfies (5.1.9). Moreover, in the case of TSRK methods, the matrix $\mathbf{V}$ has a very precise structure given by the representation in (2.2.3), and its eigenvalues are 1, $-\theta$ and 0

(with multiplicity s). Therefore, for $\theta \neq 0$, looking for a stability function of the type (5.1.3) is quite natural choice.

To express the IQS conditions (5.1.4) and (5.1.5) in terms of the coefficients θ, u, v, w, A, and B of TSRK method (2.2.1) we partition the matrix $\mathbf{X}$ as follows

$$\mathbf{X} = \left[\begin{array}{c|c} X_{11} & X_{12} \\ \hline X_{21} & X_{22} \end{array}\right], \tag{5.1.10}$$

where $X_{11} \in \mathbb{R}^{2\times 2}$, $X_{12} \in \mathbb{R}^{2\times s}$, $X_{21} \in \mathbb{R}^{s\times 2}$, $X_{22} \in \mathbb{R}^{s\times s}$. We also partition accordingly the matrices $\mathbf{B}$, $\mathbf{U}$, and $\mathbf{V}$ (see (2.2.3))

$$\mathbf{B} = \left[\begin{array}{c} B_{11} \\ \hline I_s \end{array}\right], \quad \mathbf{U} = \left[\begin{array}{c|c} U_{11} & A \end{array}\right], \quad \mathbf{V} = \left[\begin{array}{c|c} V_{11} & V_{12} \\ \hline 0 & 0 \end{array}\right],$$

where $B_{11} \in \mathbb{R}^{2\times s}$, $U_{11} \in \mathbb{R}^{s\times 2}$, $V_{11} \in \mathbb{R}^{2\times 2}$, $V_{12} \in \mathbb{R}^{2\times s}$ are given by

$$B_{11} = \left[\begin{array}{c} w^T \\ 0 \end{array}\right],\ U_{11} = \left[\begin{array}{cc} e-u & u \end{array}\right],\ V_{11} = \left[\begin{array}{cc} 1-\theta & \theta \\ 1 & 0 \end{array}\right],\ V_{12} = \left[\begin{array}{c} v^T \\ 0 \end{array}\right],$$

and 0 in $\mathbf{V}$ stands for zero matrices of dimension $s\times 2$ and $s\times s$, respectively.

Theorem 5.1.2 *A TSRK method (2.2.1) has IQS if there exist vectors $\alpha, \beta \in \mathbb{R}^s$ and a matrix $X \in \mathbb{R}^{s\times s}$ such that the following conditions are satisfied*

$$B = \alpha w^T + X, \quad e = \alpha + \beta, \quad u = \theta\alpha, \quad A = \alpha v^T. \tag{5.1.11}$$

Proof: According to the way we have partitioned the above matrices, IQS conditions (5.1.4) and (5.1.5) are equivalent to

$$B = X_{21}B_{11} + X_{22},$$

and

$$U_{11} = X_{21}V_{11}, \quad A = X_{21}V_{12},$$

respectively. By setting

$$\left[\begin{array}{cc} \alpha & \beta \end{array}\right] = X_{21} \in \mathbb{R}^{s\times 2}, \quad X = X_{22}, \tag{5.1.12}$$

with $\alpha, \beta \in \mathbb{R}^s$, the theorem follows. □

5.1.2 Construction of TSRK methods with IQS properties

We first compute the coefficient matrix A and the vector v from stage order and order conditions (2.2.4) and (2.2.5). Introducing the notation

$$C = \begin{bmatrix} c & \frac{c^2}{2!} & \cdots & \frac{c^s}{s!} \end{bmatrix}, \quad \tilde{C} = \begin{bmatrix} e & \frac{c}{1!} & \cdots & \frac{c^{s-1}}{(s-1)!} \end{bmatrix},$$

$$d = \begin{bmatrix} -1 & \frac{1}{2!} & \cdots & \frac{(-1)^s}{s!} \end{bmatrix}^T, \quad g = \begin{bmatrix} 1 & \frac{1}{2!} & \cdots & \frac{1}{s!} \end{bmatrix}^T,$$

$$E = \begin{bmatrix} e & \frac{c-e}{1!} & \cdots & \frac{(c-e)^{s-1}}{(s-1)!} \end{bmatrix},$$

the conditions (2.2.4) and (2.2.5) are equivalent to $AE = C - ud^T - B\tilde{C}$, and $v^T E = g^T - \theta d^T - w^T \tilde{C}$, respectively. We note that, by assuming distinct abscissas, the matrices C, $\tilde{C}$, and E are nonsingular because of Vandermonde type. Hence we obtain

$$A = (C - ud^T - B\tilde{C})E^{-1}, \tag{5.1.13}$$

and

$$v^T = (g^T - \theta d^T - w^T \tilde{C})E^{-1}. \tag{5.1.14}$$

To obtain TSRK methods with IQS we compute the matrix X from the first condition in (5.1.11), i.e. $X = B - \alpha w^T$, and the vectors β and u from the second and third condition of (5.1.11), i.e.

$$\beta = e - \alpha, \quad u = \theta\alpha. \tag{5.1.15}$$

Then we enforce the last condition in (5.1.11) using the representations of A and v given by (5.1.13) and (5.1.14). This leads to

$$C - ud^T - B\tilde{C} = \alpha g^T - \theta\alpha d^T - \alpha w^T \tilde{C}$$

and, since $\tilde{C}$ is nonsingular, using the condition $u = \theta\alpha$ we obtain

$$B = (C - \alpha(g^T - w^T \tilde{C}))\tilde{C}^{-1}. \tag{5.1.16}$$

Computing the matrix B from (5.1.16) and then the matrix A from (5.1.13), where $u = \theta\alpha$, and the vector v from (5.1.14) we obtain a family of TSRK

methods (2.2.1) of order $p = s$ and stage order $q = p$ which depends on the parameters θ, α, c and w. By construction these methods satisfy IQS conditions (5.1.11). We impose next the condition (5.1.1) that the matrix B has a one point spectrum $\sigma(B) = \{\lambda\}$, where λ will be chosen in such a way that the resulting method has some desirable stability properties. This is equivalent to the requirement that the characteristic polynomial of B assumes the simple form $\det(\omega I_s - B) = (\omega - \lambda)^s$. Since

$$\det(\omega I_s - B) = \sum_{k=0}^{s} b_k \omega^{s-k},$$

where $b_0 = 1$, $b_k = b_k(\theta, \alpha, c, w)$, $k = 1, 2, \ldots, s$, and

$$(\omega - \lambda)^s = \sum_{k=0}^{s} \left(\begin{array}{c} s \\ k \end{array} \right) (-1)^k \lambda^k \omega^{s-k}$$

this is equivalent to the system of equations

$$b_k(\theta, \alpha, c, w) = \left(\begin{array}{c} s \\ k \end{array} \right) (-1)^k \lambda^k, \quad k = 1, 2, \ldots, s. \tag{5.1.17}$$

Since it follows from (5.1.16) that

$$B = (C - \alpha g^T)\tilde{C}^{-1} + \alpha w^T,$$

the system (5.1.17) is linear with respect to w, and its solution leads, by virtue of Theorem 5.1.1, to methods for which stability polynomial $p(\omega, z)$ takes the form (5.1.3), i.e.

$$p(\omega, z) = \omega^s \Big((1 - \lambda z)^s \omega^2 - p_1(z)\omega + p_0(z) \Big).$$

The polynomials $p_1(z)$ and $p_0(z)$ appearing in $p(\omega, z)$ take the form

$$p_1(z) = p_{10} + p_{11} z + \cdots + p_{1,s-1} z^{s-1} + p_{1s} z^s,$$

$$p_0(z) = p_{00} + p_{01} z + \cdots + p_{0,s-1} z^{s-1} + p_{0s} z^s.$$

Since, by (2.2.6)–(5.1.3),

$$p(\omega, 0) = \det(\omega I_{s+2} - \mathbf{V}) = \omega^s(\omega - 1)(\omega + \theta) = \omega^s(\omega^2 - p_1(0)\omega + p_0(0)),$$

it follows that

$$p_1(0) = p_{10} = 1 - \theta, \quad p_0(0) = p_{00} = -\theta. \tag{5.1.18}$$

Moreover we observe that, assuming A-stability, the L-stability requirement is equivalent to

$$\lim_{z\to\infty} \frac{p_1(z)}{(1-\lambda z)^s} = 0, \quad \lim_{z\to\infty} \frac{p_0(z)}{(1-\lambda z)^s} = 0,$$

which leads to the conditions

$$p_{1s} = 0, \quad p_{0s} = 0. \tag{5.1.19}$$

Therefore, the polynomials $p_1(z)$ and $p_2(z)$ now take the form

$$p_1(z) = 1 - \theta + p_{11}z + \cdots + p_{1,s-1}z^{s-1},$$
$$p_0(z) = -\theta + p_{01}z + \cdots + p_{0,s-1}z^{s-1}.$$

For the method of order $p = s$ the stability polynomial $p(\omega, z)$ satisfies the condition

$$p(e^z, z) = O(z^{s+1}), \quad z \to 0. \tag{5.1.20}$$

Expanding (5.1.20) into power series around $z = 0$ it follows from (5.1.18) that the constant term vanishes, and comparing to zero terms of order z^k, $k = 1, 2, \ldots, s$, we obtain a system of s linear equations for the $2(s-1)$ coefficients p_{1j}, p_{0j}, $j = 1, 2 \ldots, s-1$, of the polynomials $p_1(z)$ and $p_0(z)$. This system has a family of solutions depending on λ, θ, and $s-2$ additional parameters which may be chosen from p_{1j} and p_{0j}.

The last point is now the computation of the vector α, which can be carried out comparing the expression of the stability polynomial now computed with the one coming from (5.1.9), i.e.

$$\tilde{p}(\omega, z) = \omega^s \det(\omega I_2 - \tilde{M}_{11}(z)). \tag{5.1.21}$$

Since the IQS conditions (5.1.11) do not depend on the blocks X_{11} and X_{12} of the matrix $\mathbf{X}$ in (5.1.10) we can assume without loss of generality that $X_{11} = 0$ and $X_{12} = 0$. Therefore, it follows from (5.1.9) that

$$\left[\begin{array}{c|c} I_2 & 0 \\ \hline -zX_{21} & I_2 - zX_{22} \end{array}\right] \left[\begin{array}{c|c} M_{11}(z) & M_{12}(z) \\ \hline M_{21}(z) & M_{22}(z) \end{array}\right]$$
$$= \left[\begin{array}{c|c} \tilde{M}_{11}(z) & \tilde{M}_{12}(z) \\ \hline 0 & 0 \end{array}\right] \left[\begin{array}{c|c} I_2 & 0 \\ \hline -zX_{21} & I_2 - zX_{22} \end{array}\right].$$

Hence,

$$M_{11}(z) = \tilde{M}_{11}(z) - z\tilde{M}_{12}(z)X_{21}, \quad M_{12}(z) = \tilde{M}_{12}(z)(I_2 - zX_{22}),$$

which, taking into account (5.1.12), leads to the following formula for the matrix $\tilde{M}_{11}(z)$

$$\tilde{M}_{11}(z) = M_{11}(z) + zM_{12}(z)(I_2 - zX)^{-1}[\ \alpha \quad \beta\ ], \tag{5.1.22}$$

where we recall $X = X_{22}$.

The construction of highly stable TSRK methods (2.2.1) with IQS properties and coefficient matrix B with one point spectrum $\sigma(B) = \{\lambda\}$ can be summarized in the following algorithm.

1. Choose the abscissa vector c with distinct components, such that the matrices $\tilde{C}$ and E defined at the beginning of this section are nonsingular.

2. Choose the parameters θ and $\lambda > 0$ so that the stability polynomial $p(\omega, z)$ is A-stable and also L-stable.

3. Compute the coefficient matrix B from the formula (5.1.16). This matrix depends on the vectors α and w.

4. Compute the vectors β and u from the second and third condition of (5.1.11), i.e. $\beta = e - \alpha$ and $u = \theta\alpha$.

5. Compute the coefficient matrix A from (5.1.13) and the vector v from (5.1.14). They depend on α and w.

6. Solve the system (5.1.17) with respect to w. This leads to a family of methods with IQS for which the matrix B has a one point spectrum $\sigma(B) = \{\lambda\}$.

7. Compute the matrix $\tilde{M}_{11}(z)$ from the relation (5.1.22) and the stability polynomial $\tilde{p}(\omega, z)$ from (5.1.22). Everything now depends on p_{1j} and p_{0j}, which must be computed from (5.1.20).

8. Compare the expression of $\tilde{p}(\omega, z)$ obtained in point 7 with (5.1.3), and solve with respect to the parameter vector α. Then the stability polynomial of resulting TSRK method (2.2.1) corresponds to the polynomial $p(\omega, z)$ in point 2.

5.1.3 Construction of highly stable quadratic stability polynomials

In this section we derive methods of order $p = q = s$ with quadratic stability polynomials which are A-stable and L-stable for $s = 1$, 2, 3, and 4.

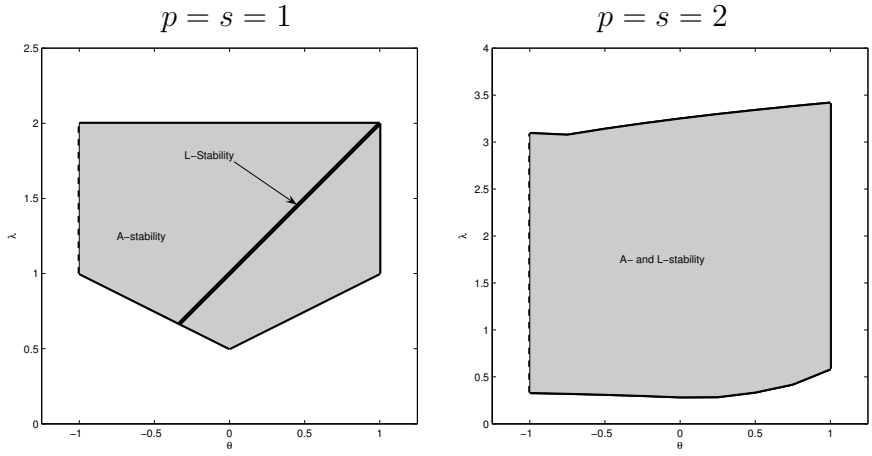

Figure 5.1: Regions of A-stability and L-stability in the (θ, λ)-plane for $p(\omega, z)$ with $p = s = 1$ and $p = s = 2$. The figure on the left corresponds to $p_{01} = 0$.

For $s = 1$ the stability polynomial (5.1.3) takes the form

$$p(\omega, z) = \omega\Big((1 - \lambda z)\omega^2 - p_1(z)\omega + p_0(z)\Big),$$

with $p_1(z) = 1 - \theta + p_{11}z$, $p_0(z) = -\theta + p_{01}z$. The solution of the equation corresponding to (5.1.20) with $s = 1$ is $p_{11} = 1 - \lambda - p_{01} + \theta$. Assuming that $p_{01} = 0$ it can be verified using the Schur criterion that $p(\omega, z)$ is A-stable if and only if $2\lambda + \theta \geq 1$, $2\lambda - \theta \geq 1$ and $\lambda \leq 2$. Moreover, $p_{11} = 0$ leads to $\theta = \lambda - 1$ and the resulting polynomial is L-stable if and only if $\frac{2}{3} \leq \lambda \leq 2$. This is illustrated in Fig. 5.1, where the range of parameters (θ, λ) for which $p(\omega, z)$ is A-stable corresponds to the shaded region and the range of (θ, λ) for which $p(\omega, z)$ is L-stable is plotted by a thick line.

For $s = 2, 3, 4$, we are looking for A-stable methods which are also L-stable. This is the case if the degrees of the polynomials $p_0(z)$ and $p_1(z)$

in (5.1.3) are equal to $s-1$, according to (5.1.19). For $s=2$ the stability polynomial (5.1.3) takes the form

$$p(\omega,z)=\omega^2\Big((1-\lambda z)^2\omega^2-p_1(z)\omega+p_0(z)\Big),$$

with $p_1(z)=1-\theta+p_{11}z$ and $p_0(z)=-\theta+p_{01}z$. The system of equations corresponding to (5.1.20) with $s=2$ takes the form

$$p_{11}-p_{01}=1-2\lambda+\theta,\quad 2p_{11}=3-8\lambda+2\lambda^2+\theta.$$

and the unique solution to this system is given by

$$p_{11}=\frac{3-8\lambda+2\lambda^2+\theta}{2},\quad p_{01}=\frac{1-4\lambda+2\lambda^2-\theta}{2}.$$

The range of parameters (θ,λ) for which the $p(\omega,z)$ is A-stable and also L-stable is plotted in Fig. 5.1 by the shaded region.

For $s=3$ the stability polynomial (5.1.3) takes the form

$$p(\omega,z)=\omega^3\Big((1-\lambda z)^3\omega^2-p_1(z)\omega+p_0(z)\Big),$$

with $p_1(z)=1-\theta+p_{11}z+p_{12}z^2$ and $p_0(z)=-\theta+p_{01}z+p_{02}z^2$. The system of equations corresponding to (5.1.20) with $s=3$ takes the form

$$p_{11}-p_{01}=1-3\lambda+\theta,\quad 2p_{11}-2p_{02}+2p_{12}=3-12\lambda+6\lambda^2+\theta,$$

$$3p_{11}+6p_{12}=7-36\lambda+36\lambda^2-6\lambda^3+\theta,$$

and assuming that $p_{02}=0$, the unique solution to this system is given by

$$p_{11}=\frac{2(1-9\lambda^2+3\lambda^3+\theta)}{3},\quad p_{12}=\frac{5-36\lambda+54\lambda^2-12\lambda^3-\theta}{6},$$

$$p_{01}=\frac{1-9\lambda+18\lambda^2-6\lambda^3+\theta}{3}.$$

The range of parameters (θ,λ) for which the $p(\omega,z)$ is A-stable and also L-stable is plotted in Fig. 5.2 by the shaded region.

Finally, for $s=4$ the stability polynomial (5.1.3) takes the form

$$p(\omega,z)=\omega^4\Big((1-\lambda z)^4\omega^2-p_1(z)\omega+p_0(z)\Big),$$

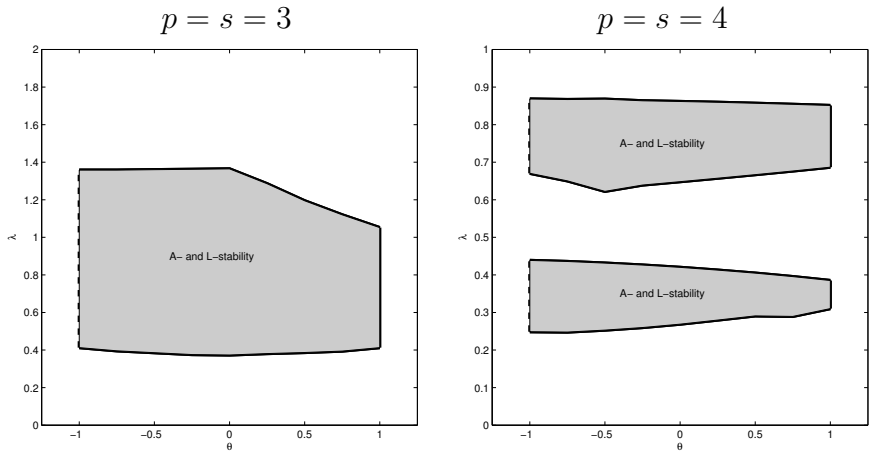

Figure 5.2: Regions of A-stability and L-stability in the (θ, λ)-plane for $p(\omega, z)$ with $p = s = 3$ and $p = s = 4$. The figure on the left corresponds to $p_{02} = 0$ and the figure on the right to $p_{03} = p_{13} = 0$.

with $p_1(z) = 1-\theta+p_{11}z+p_{12}z^2+p_{13}z^3$ and $p_0(z) = -\theta+p_{01}z+p_{02}z^2+p_{03}z^3$. The system of equations corresponding to (5.1.20) with $s = 4$ takes the form

$$p_{11} - p_{01} = 1 - 4\lambda + \theta, \quad 2p_{11} + 2p_{12} - 2p_{02} = 3 - 16\lambda + 12\lambda^2 + \theta,$$

$$3p_{11} + 6p_{12} + 6p_{13} - 6p_{03} = 7 - 48\lambda + 72\lambda^2 - 24\lambda^3 + \theta,$$

$$4p_{11} + 12p_{12} + 24p_{13} = 15 - 128\lambda + 288\lambda^2 - 192\lambda^3 + 24\lambda^4 + \theta,$$

and assuming that $p_{13} = 0$ and $p_{03} = 0$ the unique solution to this system is given by

$$p_{11} = \frac{1 - 32\lambda + 144\lambda^2 - 144\lambda^3 + 24\lambda^4 - \theta}{2},$$

$$p_{12} = \frac{17 - 192\lambda + 576\lambda^2 - 480\lambda^3 + 72\lambda^4 - \theta}{12},$$

$$p_{01} = \frac{3 - 40\lambda + 144\lambda^2 - 144\lambda^3 + 24\lambda^4 + \theta}{2},$$

$$p_{02} = \frac{7 - 96\lambda + 360\lambda^2 - 384\lambda^3 + 72\lambda^4 + \theta}{12}.$$

The range of parameters (θ, λ) for which the $p(\omega, z)$ is A-stable and also L-stable is plotted in Fig. 5.2 by the shaded region. The regions for $s = 2$,

3, and 4 were obtained by computer searches in the parameter space (θ, λ) using the Schur criterion.

5.1.4 Examples of TSRK methods with IQS

In this section we will follow the issues described in the previous pages to derive examples of A-stable and L-stable TSRK methods (2.2.1) with IQS and for which the coefficient matrix B has a one point spectrum $\sigma(B) = \{\lambda\}$. These examples correspond to $p = q = s$ for $s = 1$, 2, 3, and 4. It is always assumed that $\theta = 0$ which implies that $u = 0$, see (5.1.15).

Example 1. TSRK methods with $p = q = s = 1$. The coefficients of the method corresponding to $\lambda = 1$ and arbitrary abscissa c are given by

$$\begin{array}{c|c|c} u & A & B \\ \hline \theta & v & w \end{array} = \begin{array}{c|c|c} 0 & c-1 & 1 \\ \hline 0 & c-1 & 2-c \end{array}.$$

The stability polynomial $p(\omega, z)$ of this family of methods is

$$p(\omega, z) = \omega\big((1-z)\omega - 1\big)$$

for any c. In particular, for $c = 1$ this method is equivalent to the backward Euler method.

Example 2. TSRK methods with $p = q = s = 2$. The coefficients of the method corresponding to $\lambda = \frac{5}{4}$ and abscissa vector $c = [0, 1]^T$ are given by

$$\begin{array}{c|c|c} u & A & B \\ \hline \theta & v^T & w^T \end{array} = \begin{array}{c|cc|cc} 0 & -\frac{25}{32} & -\frac{25}{32} & \frac{75}{32} & -\frac{25}{32} \\ 0 & -\frac{11}{32} & -\frac{11}{32} & \frac{49}{32} & \frac{5}{32} \\ \hline 0 & -\frac{11}{32} & -\frac{11}{32} & \frac{49}{32} & \frac{5}{32} \end{array}.$$

The stability polynomial $p(\omega, z)$ of this method is

$$p(\omega, z) = \omega^2\Big(\Big(1 - \frac{5}{4}z\Big)^2\omega^2 - \Big(1 - \frac{31}{16}z\Big)\omega - \frac{7}{16}z\Big).$$

Example 3. TSRK methods with $p = q = s = 3$. The coefficients of the method corresponding to $\lambda = \frac{3}{4}$ and abscissa the vector $c = [0, \frac{1}{2}, 1]^T$ are

given by $u = 0$, $\theta = 0$,

$$A = \begin{bmatrix} \frac{1371718}{2008359} & -\frac{1349029}{610487} & -\frac{598537}{334774} \\ \frac{1996151}{1120476} & -\frac{3899713}{676582} & -\frac{4599017}{986185} \\ \frac{2289675}{1145977} & -\frac{2640065}{408409} & -\frac{4106281}{785118} \end{bmatrix},$$

$$B = \begin{bmatrix} \frac{3955778}{915873} & -\frac{573724}{492365} & \frac{253229}{1575340} \\ \frac{4717083}{411104} & -\frac{3938351}{1455396} & \frac{307583}{814540} \\ \frac{6683188}{522061} & -\frac{3272705}{1193527} & \frac{472108}{741259} \end{bmatrix},$$

$$v = \begin{bmatrix} \frac{2289675}{1145977} & -\frac{2640065}{408409} & -\frac{4106281}{785118} \end{bmatrix}^T,$$

$$w = \begin{bmatrix} \frac{6683188}{522061} & -\frac{3272705}{1193527} & \frac{472108}{741259} \end{bmatrix}^T.$$

The stability polynomial $p(\omega, z)$ of this method is

$$p(\omega, z) = \omega^3\Big(\Big(1 - \frac{3}{4}z\Big)^3\omega^2 - \Big(1 - \frac{179}{96}z + \frac{53}{96}z^2\Big)\omega - \frac{59}{96}z\Big).$$

Example 4. TSRK methods with $p = q = s = 4$. The coefficients of the method corresponding to $\lambda = \frac{1}{3}$ and the abscissa vector $c = [0, \frac{1}{3}, \frac{2}{3}, 1]^T$ are given by

$$A = \begin{bmatrix} -\frac{73571}{418565} & \frac{316790}{450193} & -\frac{383309}{370547} & -\frac{1102057}{1459404} \\ -\frac{324116}{495273} & \frac{3108022}{1186313} & -\frac{2008351}{521461} & -\frac{1905671}{677809} \\ -\frac{813738}{787901} & \frac{4021146}{972541} & -\frac{6409321}{1054477} & -\frac{6349415}{1430988} \\ -\frac{426460}{370257} & \frac{4154204}{900915} & -\frac{12185608}{1797671} & -\frac{6621076}{1338039} \end{bmatrix},$$

$$B = \begin{bmatrix} \frac{1082275}{789096} & -\frac{47158}{1102905} & -\frac{20658}{230377} & \frac{16548}{733283} \\ \frac{2053468}{392523} & \frac{173881}{1660851} & -\frac{337517}{836884} & \frac{86197}{880374} \\ \frac{13765224}{1684843} & \frac{119918}{620675} & -\frac{387828}{932779} & \frac{214966}{1621163} \\ \frac{8694859}{954168} & \frac{68987}{727614} & -\frac{198815}{935168} & \frac{90358}{331129} \end{bmatrix},$$

$$v = \begin{bmatrix} -\frac{426460}{370257} & \frac{4154204}{900915} & -\frac{12185608}{1797671} & -\frac{6621076}{1338039} \end{bmatrix}^T,$$

$$w = \left[\begin{array}{cccc} \frac{8694859}{954168} & \frac{68987}{727614} & -\frac{198815}{935168} & \frac{90358}{331129} \end{array} \right]^T.$$

The stability polynomial $p(\omega, z)$ of this method is

$$p(\omega, z) = \omega^4 \Big(\Big(1 - \frac{1}{3} z \Big)^4 \omega^2 - p_1(z)\omega + p_0(z) \Big)$$

with

$$p_1(z) = 1 - \frac{744347}{1148421} z + \frac{2965}{320219} z^2, \quad p_0(z) = -\frac{241021}{765596} z - \frac{198226}{1427227} z^2.$$

5.2 Algebraically Stable TSRK methods

The numerical search for algebraically stable TSRK methods with $\vartheta = 0$, $u = 0$ and $\vartheta \neq 0$, $u \neq 0$, which is based on the criterion consisting of the conditions 1–4 presented in Section 2.1.4 is described in the next pages.

5.2.1 Analysis of TSRK methods with $\vartheta = 0$ and $u = 0$.

In this section we consider TSRK methods with $\vartheta = 0$ and $u = 0$. Then, the corresponding TSRK method (2.2.1) can be represented as GLM with coefficient matrices $\mathbf{A}$, $\mathbf{U}$, $\mathbf{B}$ and $\mathbf{V}$ defined by

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{U} \\ \hline \mathbf{B} & \mathbf{V} \end{array} \right] = \left[\begin{array}{c|cc} B & e & A \\ \hline w^T & 1 & v^T \\ I & 0 & 0 \end{array} \right]. \tag{5.2.1}$$

It can be verified that the stability function $p(\eta, z)$ of this method takes the form

$$\begin{array}{rcl} p(\eta, z) & = & \eta^{s+1} - R_1(z)\eta^s + R_2(z)\eta^{s-1} + \cdots \\ & + & (-1)^s R_s(z)\eta + (-1)^{s+1} R_{s+1}(z), \end{array} \tag{5.2.2}$$

where $R_i(z)$ are rational functions

$$R_i(z) = \frac{p_i(z)}{p_0(z)}, \quad i = 1, 2, \ldots, s+1,$$

with

$$\begin{aligned} p_0(z) &= 1 + p_{01}z + \cdots + p_{0s}z^s, \\ p_1(z) &= 1 + p_{11}z + \cdots + p_{1s}z^s, \\ p_2(z) &= p_{21}z + \cdots + p_{2s}z^s, \\ &\vdots \\ p_s(z) &= p_{s,s-1}z^{s-1} + p_{ss}z^s, \\ p_{s+1}(z) &= p_{s+1,s}z^s. \end{aligned}$$

To investigate stability properties of GLMs (5.2.1) it is more convenient to work with the polynomial

$$\tilde{p}(\eta, z) = p_0(z)p(\eta, z) \tag{5.2.3}$$

instead of the rational function $p(\eta, z)$ and we will always adopt this approach. It can be verified that for TSRK method (5.2.1) the preconsistency vector $\mathbf{q}_0$ takes the form

$$\mathbf{q}_0 = \left[\begin{array}{c} 1 \\ \hline \mathbf{0} \end{array}\right] \in \mathbb{R}^{s+1}$$

and the vector $\tilde{\mathbf{w}}$ satisfying (2.1.32) is

$$\tilde{\mathbf{w}} = \left[\begin{array}{c} 1 \\ \hline v \end{array}\right] \in \mathbb{R}^{s+1}.$$

Hence, the matrix $\tilde{\mathbf{D}}$ defined by (2.1.33) takes the form

$$\tilde{\mathbf{D}} = \operatorname{diag}\left(\left[\begin{array}{c|c} w & I \end{array}\right] \left[\begin{array}{c} 1 \\ \hline v \end{array}\right] \right) = \operatorname{diag}(w + v).$$

We compute next the Nyquist stability function $\mathbf{N}(\xi)$ corresponding to TSRK method (5.2.1) and the Hermitian part of $\tilde{\mathbf{D}}\mathbf{N}(\xi)$. Using the formula

$$\left[\begin{array}{c|c} B & A \\ \hline 0 & D \end{array}\right]^{-1} = \left[\begin{array}{c|c} B^{-1} & -B^{-1}AD^{-1} \\ \hline 0 & D^{-1} \end{array}\right] \tag{5.2.4}$$

where B and D are square and nonsingular matrices, we have

$$\begin{aligned} \mathbf{N}(\xi) &= B + \left[\begin{array}{c|c} e & A \end{array}\right] \left[\begin{array}{c|c} \xi - 1 & -v^T \\ \hline 0 & \xi I \end{array}\right]^{-1} \left[\begin{array}{c} w^T \\ \hline I \end{array}\right] \\ &= B + \frac{1}{1-\xi}\, e\, w^T + \frac{1}{\xi(\xi-1)}\, e\, v^T + \frac{1}{\xi}A. \end{aligned}.$$

We have also

$$\begin{aligned}\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big) &= \frac{1}{2}\bigg(\tilde{\mathbf{D}}\Big(B+\frac{1}{\xi}B\Big)+\Big(B^T+\frac{1}{\bar{\xi}}A^T\Big)\tilde{\mathbf{D}}\\ &\quad+\frac{1}{\xi-1}\tilde{\mathbf{D}}ew^T+\frac{1}{\bar{\xi}-1}we^T\tilde{\mathbf{D}}+\frac{1}{\xi(\xi-1)}\tilde{\mathbf{D}}ev^T+\frac{1}{\bar{\xi}(\bar{\xi}-1)}ve^T\tilde{\mathbf{D}}\bigg),\end{aligned}$$

where $\bar{\xi}$ stands for conjugate of ξ. Taking into account that

$$\tilde{\mathbf{D}}e=\mathrm{diag}(w+v)e=w+v,\quad e^T\tilde{\mathbf{D}}=e^T\mathrm{diag}(w+v)=(w+v)^T,$$

it follows that

$$\begin{aligned}\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big) &= \frac{1}{2}\bigg(\tilde{\mathbf{D}}\Big(B+\frac{1}{\xi}A\Big)+\Big(B^T+\frac{1}{\bar{\xi}}A^T\Big)\tilde{\mathbf{D}}\\ &\quad+\Big(\frac{1}{\xi-1}+\frac{1}{\bar{\xi}-1}\Big)ww^T+\Big(\frac{1}{\bar{\xi}-1}+\frac{1}{\xi(\xi-1)}\Big)wv^T\\ &\quad+\Big(\frac{1}{\xi-1}+\frac{1}{\bar{\xi}(\bar{\xi}-1)}\Big)vw^T+\Big(\frac{1}{\xi(\xi-1)}+\frac{1}{\bar{\xi}(\bar{\xi}-1)}\Big)vv^T\bigg).\end{aligned}$$

We compute next the limit

$$\lim_{t\to 0}\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big)\Big|_{\xi=e^{it}}.$$

Since

$$\lim_{t\to 0}\Big(\frac{1}{\xi-1}+\frac{1}{\bar{\xi}-1}\Big)\Big|_{\xi=e^{it}}=-1,\quad \lim_{t\to 0}\Big(\frac{1}{\bar{\xi}-1}+\frac{1}{\xi(\xi-1)}\Big)\Big|_{\xi=e^{it}}=-2,$$

$$\lim_{t\to 0}\Big(\frac{1}{\xi-1}+\frac{1}{\bar{\xi}(\bar{\xi}-1)}\Big)\Big|_{\xi=e^{it}}=-2,\quad \lim_{t\to 0}\Big(\frac{1}{\xi(\xi-1)}+\frac{1}{\bar{\xi}(\bar{\xi}-1)}\Big)\Big|_{\xi=e^{it}}=-3,$$

it follows that

$$\begin{aligned}\lim_{t\to 0}\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big)\Big|_{\xi=e^{it}} &= \frac{1}{2}\Big(\mathrm{diag}(w+v)(B+A)\\ &\quad+(B+A)^T\mathrm{diag}(w+v)-ww^T-2(wv^T+vw^T)-3vv^T\Big).\end{aligned}\tag{5.2.5}$$

Observe also that $\mathrm{He}(\tilde{\mathbf{D}}\mathbf{N}(\xi))$ does not have a limit as $\xi\to 1$. For example, as $\xi=x\to 1$ along the real axis we have

$$\lim_{x\to 1}\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big)\Big|_{\xi=x}=\infty.$$

5.2.2 Analysis of TSRK methods with $\vartheta \neq 0$ and $u \neq 0$.

We now consider the general case where $\vartheta \neq 0$ and $u \neq 0$: in this case, the TSRK method (2.2.1) can be represented as GLM with coefficient matrices $\mathbf{A}$, $\mathbf{U}$, $\mathbf{B}$ and $\mathbf{V}$ defined by

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{U} \\ \hline \mathbf{B} & \mathbf{V} \end{array}\right] = \left[\begin{array}{c|ccc} B & e-u & u & A \\ \hline w^T & 1-\vartheta & \vartheta & v^T \\ 0 & 1 & 0 & 0 \\ I & 0 & 0 & 0 \end{array}\right]. \tag{5.2.6}$$

Similarly as in Section 5.2.1 it can be verified that the stability function $p(\eta, z)$ of this method takes the form

$$\begin{array}{rcl} p(\eta,z) & = & \eta^{s+2} - R_1(z)\eta^{s+1} + R_2(z)\eta^s + \cdots \\ & + & (-1)^{s+1}R_{s+1}(z)\eta + (-1)^{s+2}R_{s+2}(z), \end{array} \tag{5.2.7}$$

where $R_i(z)$ are rational functions

$$R_i(z) = \frac{p_i(z)}{p_0(z)}, \quad i = 1, 2, \ldots, s+2,$$

with

$$\begin{array}{l} p_0(z) = 1 + p_{01}z + \cdots + p_{0s}z^s, \\ p_1(z) = 1 - \vartheta + p_{11}z + \cdots + p_{1s}z^s, \\ p_2(z) = -\vartheta + p_{21}z + \cdots + p_{2s}z^s, \\ \qquad\qquad \vdots \\ p_{s+1}(z) = p_{s+1,s-1}z^{s-1} + p_{s+1,s}z^s, \\ p_{s+2}(z) = p_{s+2,s}z^s. \end{array}$$

As before, to investigate stability properties of GLMs (5.2.6) it is more convenient to work with the polynomial $\tilde{p}(\eta, z) = p_0(z)p(\eta, z)$ instead of the rational function $p(\eta, z)$ and we will again always adopt this approach.

Similarly as in Section 5.2.1 it can be verified that for TSRK method (5.2.6) the preconsistency vector $\mathbf{q}_0$ takes the form

$$\mathbf{q}_0 = \left[\begin{array}{c} 1 \\ 1 \\ \mathbf{0} \end{array}\right] \in \mathbb{R}^{s+2},$$

the vector $\tilde{\mathbf{w}}$ satisfying (2.1.32) is

$$\tilde{w} = \frac{1}{1+\vartheta}\begin{bmatrix} 1 \\ \vartheta \\ \mathbf{w} \end{bmatrix} \in \mathbb{R}^{s+2},$$

and the matrix $\tilde{\mathbf{D}}$ defined by (2.1.33) is

$$\tilde{\mathbf{D}} = \frac{1}{1+\vartheta}\mathrm{diag}(w+v).$$

We compute next the Nyquist stability function $\mathbf{N}(\xi)$ corresponding to TSRK method (5.2.1) and the Hermitian part of $\tilde{\mathbf{D}}\mathbf{N}(\xi)$. Using the formula (5.2.4) we obtain

$$(\xi\mathbf{I}-\mathbf{V})^{-1} = \left[\begin{array}{cc|c} \frac{\xi}{\Delta} & \frac{\vartheta}{\Delta} & \frac{v^T}{\Delta} \\ \frac{1}{\Delta} & \frac{\xi-1+\vartheta}{\Delta} & \frac{v^T}{\xi\Delta} \\ \hline 0 & 0 & \frac{1}{\xi}I \end{array}\right],$$

where

$$\Delta = \xi(\xi-1+\vartheta)-\vartheta = (\xi-1)(\xi+\vartheta).$$

This leads to

$$\begin{aligned} \mathbf{N}(\xi) &= \mathbf{A}+\mathbf{U}(\xi\mathbf{I}-\mathbf{V})^{-1}\mathbf{B} \\ &= B + \frac{\xi}{(\xi-1)(\xi+\vartheta)}ew^T + \frac{1}{(\xi-1)(\xi+\vartheta)}ev^T \\ &- \frac{1}{\xi+\vartheta}uw^T - \frac{1}{\xi(\xi+\vartheta)}uv^T + \frac{1}{\xi}A. \end{aligned}$$

We have also

$$\begin{aligned} \mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big) &= \frac{1}{2}\Bigg(\tilde{\mathbf{D}}\Big(B+\frac{1}{\xi}A\Big) + \Big(B^T+\frac{1}{\bar{\xi}}A^T\Big)\tilde{\mathbf{D}} \\ &+ \frac{\xi}{(\xi-1)(\xi+\vartheta)}\tilde{\mathbf{D}}ew^T + \frac{\bar{\xi}}{(\bar{\xi}-1)(\bar{\xi}+\vartheta)}we^T\tilde{\mathbf{D}} \\ &+ \frac{1}{(\xi-1)(\xi+\vartheta)}\tilde{\mathbf{D}}ev^T + \frac{1}{(\bar{\xi}-1)(\bar{\xi}+\vartheta)}ve^T\tilde{\mathbf{D}} \\ &- \frac{1}{\xi+\vartheta}\tilde{\mathbf{D}}uw^T - \frac{1}{\bar{\xi}+\vartheta}wu^T\tilde{\mathbf{D}} \\ &- \frac{1}{\xi(\xi+\vartheta)}\tilde{\mathbf{D}}uv^T - \frac{1}{\bar{\xi}(\bar{\xi}+\vartheta)}vu^T\tilde{\mathbf{D}}\Bigg). \end{aligned}$$

Using the relations

$$\tilde{\mathbf{D}}e = \frac{1}{1+\vartheta}(w+v), \quad e^T\tilde{\mathbf{D}} = \frac{1}{1+\vartheta}(w+v)^T,$$
$$\tilde{\mathbf{D}}uw^T = \frac{1}{1+\vartheta}\big((w+v)\cdot u\big)w^T, \quad wu^T\tilde{\mathbf{D}} = \frac{1}{1+\vartheta}w\big((w+v)\cdot u\big)^T,$$
$$\tilde{\mathbf{D}}uv^T = \frac{1}{1+\vartheta}\big((w+v)\cdot u\big)v^T, \quad vu^T\tilde{\mathbf{D}} = \frac{1}{1+\vartheta}v\big((w+v)\cdot u\big)^T.$$

where, $u\cdot w$ denotes componentwise multiplication of vectors, this can be written as

$$\begin{aligned}
\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big) = \frac{1}{2(1+\vartheta)}\Bigg(&\mathrm{diag}(w+v)\Big(B+\frac{1}{\xi}A\Big)\\
&+\Big(B^T+\frac{1}{\bar\xi}A^T\Big)\mathrm{diag}(w+v)\\
&+\Big(\frac{\xi}{(\xi-1)(\xi+\vartheta)}+\frac{\bar\xi}{(\bar\xi-1)(\bar\xi+\vartheta)}\Big)ww^T\\
&+\Big(\frac{1}{(\xi-1)(\xi+\vartheta)}+\frac{\bar\xi}{(\bar\xi-1)(\bar\xi+\vartheta)}\Big)wv^T\\
&+\Big(\frac{\xi}{(\xi-1)(\xi+\vartheta)}+\frac{1}{(\bar\xi-1)(\bar\xi+\vartheta)}\Big)vw^T\\
&+\Big(\frac{1}{(\xi-1)(\xi+\vartheta)}+\frac{1}{(\bar\xi-1)(\bar\xi+\vartheta)}\Big)vv^T\\
&-\frac{1}{\xi+\vartheta}\big((w+v)\cdot u\big)w^T-\frac{1}{\bar\xi+\vartheta}w\big((w+v)\cdot u\big)^T\\
&-\frac{1}{\xi(\xi+\vartheta)}\big((w+v)\cdot u\big)v^T-\frac{1}{\bar\xi(\bar\xi+\vartheta)}v\big((w+v)\cdot u\big)^T\Bigg).
\end{aligned}$$

We compute next the limit

$$\lim_{t\to 0}\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big)\Big|_{\xi=e^{it}}.$$

Since

$$\lim_{t\to 0}\Big(\frac{\xi}{(\xi-1)(\xi+\vartheta)}+\frac{\bar\xi}{(\bar\xi-1)(\bar\xi+\vartheta)}\Big)\Big|_{\xi=e^{it}} = -\frac{1-\vartheta}{(1+\vartheta)^2},$$
$$\lim_{t\to 0}\Big(\frac{1}{(\xi-1)(\xi+\vartheta)}+\frac{\bar\xi}{(\bar\xi-1)(\bar\xi+\vartheta)}\Big)\Big|_{\xi=e^{it}} = -\frac{2}{(1+\vartheta)^2},$$

$$\lim_{t\to 0}\Big(\frac{\xi}{(\xi-1)(\xi+\vartheta)}+\frac{1}{(\bar{\xi}-1)(\bar{\xi}+\vartheta)}\Big)\Big|_{\xi=e^{it}}=-\frac{2}{(1+\vartheta)^2},$$
$$\lim_{t\to 0}\Big(\frac{1}{(\xi-1)(\xi+\vartheta)}+\frac{1}{(\bar{\xi}-1)(\bar{\xi}+\vartheta)}\Big)\Big|_{\xi=e^{it}}=-\frac{3+\vartheta}{(1+\vartheta)^2},$$

it follows that

$$\begin{aligned}\lim_{t\to 0}\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big)\Big|_{\xi=e^{it}}=\frac{1}{2(1+\vartheta)}\Bigg(&\mathrm{diag}(w+v)\Big(B+A\Big)\\&+(B+A)^T\mathrm{diag}(w+v)\\&-\frac{1-\vartheta}{(1+\vartheta)^2}ww^T-\frac{2}{(1+\vartheta)^2}(wv^T+vw^T)-\frac{3+\vartheta}{(1+\vartheta)^2}vv^T\\&-\frac{1}{1+\vartheta}\Big(\big((w+v)\cdot u\big)(w+v)^T+(w+v)\big((w+v)\cdot u\big)^T\Big)\Bigg).\end{aligned}$$

Observe that for $\vartheta=0$ and $u=0$ this formula for $\lim_{t\to 0}\mathrm{He}(\tilde{\mathbf{D}}\mathbf{N}(\xi))$ reduces to the formula obtained in Section 5.2.1.

5.2.3 Examples of TSRK methods with $\vartheta=0$ and $u=0$.

We have implemented an algorithm for numerical search for algebraically stable TSRK methods written as GLMs (5.2.1). This algorithm is based on minimizing the objective function which computes the negative value of the minimum of the eigenvalues of the matrix $\mathrm{He}(\tilde{\mathbf{D}}\mathbf{N}(\xi))$ for ξ such that $|\xi|=1$ and $\xi\in\mathbb{C}-\sigma(\mathbf{V})$. This objective function is a numerical realization of the necessary condition 4 for algebraic stability, which is listed at the end of Section 2.1.4. Once the methods for which $\mathrm{He}(\tilde{\mathbf{D}}\mathbf{N}(\xi))\geq 0$ for ξ such that $|\xi|=1$ and $\xi\in\mathbb{C}-\sigma(\mathbf{V})$ are found, the remaining necessary conditions 1-3 for algebraic stability are verified on the case by case basis.

In what follows we will present the results of our search for A-stable and algebraically stable methods (5.2.1) with the number of stages $s=1$, $s=2$, and $s=3$.

1. Methods with $s=1$, $p=2$, and $q=2$. Solving stage order and order conditions

$$C_k=0,\quad \tilde{C}_k=0,\quad k=1,2,$$

we obtain a one-parameter family of methods of order $p = 2$ and stage order $q = 2$ with coefficients given by

$$\begin{array}{c|c|c} u & B & A \\ \hline \vartheta & w & v \end{array} = \begin{array}{c|c|c} 0 & \frac{c(2-c)}{2} & \frac{c^2}{2} \\ \hline 0 & \frac{3-2c}{2} & \frac{2c-1}{2} \end{array},$$

where c is the abscissa. These methods are not algebraically stable for any c and are A-stable only if $c = 1$, for which the resulting method is equivalent to the trapezoidal rule

$$y_n = y_{n-1} + \frac{h}{2}\Big(f(y_{n-1}) + f(y_n)\Big).$$

2. Methods with $s = 1$, $p = 2$, and $q = 1$. Solving stage order and order conditions

$$C_1 = 0, \quad \tilde{C}_1 = 0, \quad \tilde{C}_2 = 0,$$

we obtain a two-parameter family of methods of order $p = 2$ and stage order $q = 1$ depending on a and c. The coefficients of these methods are given by

$$\begin{array}{c|c|c} u & B & A \\ \hline \vartheta & w & v \end{array} = \begin{array}{c|c|c} 0 & a & c-a \\ \hline 0 & \frac{3-2c}{2} & \frac{2c-1}{2} \end{array}.$$

It can be verified using the Schur criterion discussed in Section 5.2.1 that these methods are A-stable if $a \geq 1/2$. It can be also verified using the approach based on Albert theorem described in Section 2.1.4 that the conditions (2.1.27), (2.1.28) and (2.1.30) are satisfied if

$$g_{22} > 0 \quad \text{and} \quad 0 < \frac{g_{11}}{g_{22}} < \frac{4}{1-4c+4c^2} \quad \text{and} \quad a = \frac{1+4c-4c^2}{4} + \frac{g_{22}}{g_{11}}.$$

This implies that these methods are algebraically stable if $a > 1/2$. Putting, for example, $c = 3/4$ and $a = 1$ we obtain the method

$$\begin{array}{c|c|c} u & B & A \\ \hline \vartheta & w & v \end{array} = \begin{array}{c|c|c} 0 & 1 & -\frac{1}{4} \\ \hline 0 & \frac{3}{4} & \frac{1}{4} \end{array}$$

for which the matrices $\mathbf{M}$ and $\tilde{\mathbf{M}}$ defined in Section 2.1.4 are nonnegative definite if we choose

$$\mathbf{G} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{4} \\ \frac{1}{4} & \frac{9}{16} \end{bmatrix}, \quad \mathbf{D} = 1.$$

This confirm again that this particular TSRK method is algebraically stable.

3. Methods with $s = 2$, $p = 4$, and $q = 4$. Solving stage order and order conditions

$$C_k = 0, \quad \tilde{C}_k = 0, \quad k = 1, 2, 3, 4,$$

we obtain a two-parameter family of methods of order $p = 4$ and stage order $q = 4$ depending on the components of the abscissa vector c_1 and c_2. The coefficients of these methods are not listed here. However, we were not able to find the methods which are algebraically stable or A-stable and we conjecture that such methods do not exist in this class.

4. Methods with $s = 2$, $p = 4$, and $q = 3$. Solving stage order and order conditions

$$C_k = 0, \quad k = 1, 2, 3, \quad \tilde{C}_k = 0, \quad k = 1, 2, 3, 4,$$

we obtain a four-parameter family of methods of order $p = 4$ and stage order $q = 3$ depending on c_1, c_2, b_{12} and b_{22}. The stability polynomial $\tilde{p}(\eta, z)$ defined by (5.2.3) takes the form

$$\tilde{p}(\eta, z) = p_0(z)\eta^3 - p_1(z)\eta^2 + p_2(z)\eta - p_3(z),$$

where

$$p_0(z) = 1 + p_{01}z + p_{02}z^2, \quad p_1(z) = 1 + p_{11}z + p_{12}z^2,$$

$$p_2(z) = p_{21}z + p_{22}z^2, \quad p_3(z) = p_{31}z^2.$$

In our search for A-stable methods we compute first the parameter b_{22} from the algebraic equation

$$p_{31} = 0$$

and then apply the Schur criterion discussed in Section 5.2.1 to the quadratic polynomial

$$p_0(z)\eta^2 - p_1(z)\eta + p_2(z).$$

The results of this search are presented in Fig. 5.1 in the parameter space (c_1, c_2) for selected values of the parameter b_{12}, and in Fig. 5.2 in the parameter space (c_1, b_{12}) for $c_2 = 1$.

We will search next for algebraically stable methods using the conditions 1–4 listed at the end of Section 2.1.4. This search is based on minimizing the negative value of the objective function which computes the minimum of the

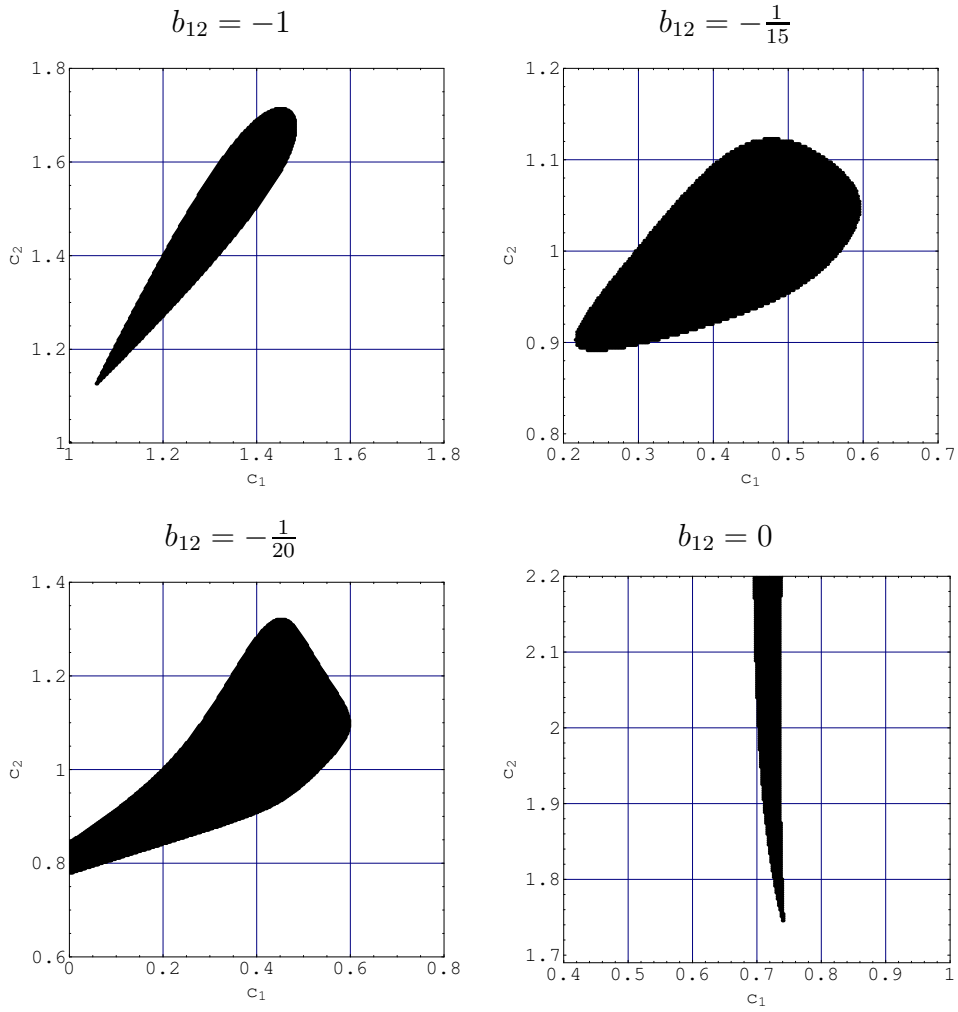

Figure 5.1: Regions of A-stability in the parameter space (c_1, c_2) for TSRK methods with $p = 4$ and $q = 3$, for specific values of b_{12}.

eigenvalues of the matrix $\mathrm{He}(\tilde{\mathbf{D}}\mathbf{N}(\xi))|_{\xi=e^{it}}$ for $t \in [0, 2\pi]$. It can be verified using formula (5.2.5) that

$$\lim_{t\to 0} \det\left(\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big)\Big|_{\xi=e^{it}}\right) = -\frac{F(c_1, c_2, b_{12}, b_{22})^2}{(c_1 - c_2)^2(c_1 - c_2 - 1)^4(c_1 - c_2 + 1)^4},$$

where $F(c_1, c_2, b_{12}, b_{22})$ is a polynomial with respect to c_1, c_2, b_{12}, b_{22}. To

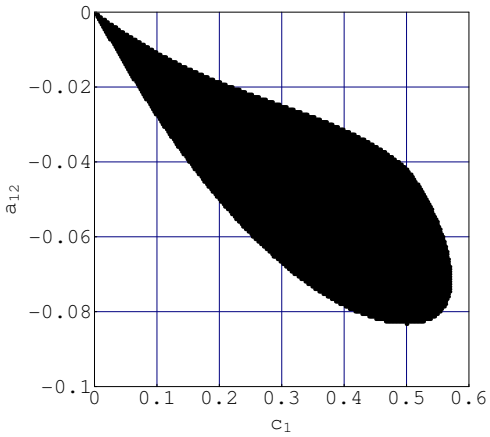

Figure 5.2: Region of A-stability in the parameter space (c_1, b_{12}) for TSRK methods methods with $p = 4$ and $q = 3$ for $c_2 = 1$.

satisfy the condition 4 at the end of Section 2.1.4, i.e.,

$$\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big) \geq 0, \quad |\xi| = 1, \quad \xi \in \mathbb{C} - \sigma(\mathbf{V}),$$

for $\xi = 1$ or $t = 0$, $t = 2\pi$, we compute the parameter b_{12} from the equation

$$F(c_1, c_2, b_{12}, b_{22}) = 0. \tag{5.2.8}$$

The search in the parameter space (c_1, c_2, b_{22}) did not lead to any methods which are algebraically stable. We were only able to find methods for which

$$\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big)\Big|_{\xi=e^{it}} \geq -2.68 \cdot 10^{-3}, \tag{5.2.9}$$

$t \in [0, 2\pi]$. Not imposing the condition (5.2.8) and searching in the parameter space $(c_1, c_2, b_{12}, b_{22})$ we were able to find some algebraically stable methods but, unfortunately, with unrealistically large values of some parameters c_1, c_2, b_{12}, or b_{22}. Restricting this search to $0 \leq c_1, c_2 \leq 1$, $-1 \leq b_{12}, b_{22} \leq 1$ we found methods for which

$$\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big)\Big|_{\xi=e^{it}} \geq -5.28 \cdot 10^{-4}, \tag{5.2.10}$$

$t \in [0, 2\pi]$.

5. Methods with $s = 3$, $p = 4$, and $q = 4$. Solving stage order and order conditions

$$C_k = 0, \quad \tilde{C}_k = 0, \quad k = 1, 2, 3, 4,$$

we obtain an eleven-parameter family of methods of order $p = 4$ and stage order $q = 4$ depending on c_1, c_2, c_3, b_{ij}, $i = 1, 2, 3$, $j = 1, 2$, w_3, and v_3. Searching for A-stable methods we assume that the abscissa vector $c = [0, 1/2, 1]^T$. The stability polynomial (5.2.3) for this family of methods takes the form

$$\tilde{p}(\eta, z) = \eta\big(p_0(z)\eta^4 - p_1(z)\eta^3 + p_2(z)\eta^2 - p_3(z)\eta + p_4(z)\big).$$

where $p_i(z)$ are polynomials of degree 3 with respect to z. We compute next the parameters b_{11} b_{12}, and b_{13} to annihilate polynomials $p_3(z)$ and $p_4(z)$. This leads to a five-parameter family of methods depending on b_{22}, b_{31}, b_{32}, w_3, and v_3 whose stability properties are determined by quadratic polynomial

$$p_0(z)\eta^2 - p_1(z)\eta + p_0(z).$$

The results of computer search based on the Schur criterion are presented in Fig. 5.3 in the parameter space (w_3, v_3) for selected values of the parameters b_{22}, b_{31}, b_{32}. We also searched for methods which are algebraically stable with general abscissa vector c. Although we did not find such methods, we found formulas for which

$$\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big)\Big|_{\xi=e^{it}} \geq -3.50 \cdot 10^{-11}, \tag{5.2.11}$$

$t \in [0, 2\pi]$. This bound was obtained by dividing the interval $[0, 2\pi]$ into $n = 10000$ subintervals. Dividing $[0, 2\pi]$ into $n = 1000$ and $n = 100$ subintervals, these bounds are equal to 0.

The coefficients of a method satisfying (5.2.11) are

$$c = \begin{bmatrix} 0.748023646320140 & -0.088623514454709 & 1.356515696201252 \end{bmatrix}^T,$$

$$B = \begin{bmatrix} 0.421393024773032 & 0.363279074448260 & -0.048601648229138 \\ -0.136821530809582 & 0.352101387625363 & 0.033470857866822 \\ 0.730130053789655 & 0.254440972752177 & 0.213275751785994 \end{bmatrix},$$

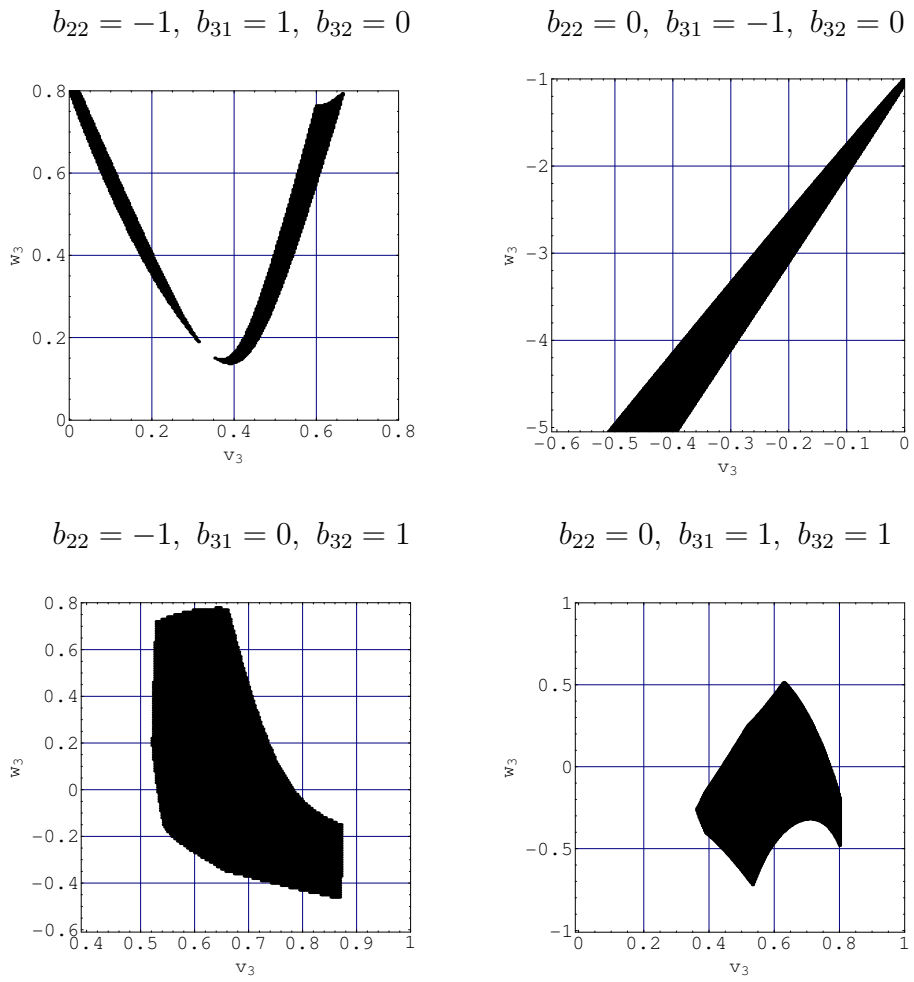

Figure 5.3: Regions of A-stability in the (w_3, v_3)-plane, for TSRK methods with $s = 3$ and $p = q = 4$, for specific values of the parameters $b_{22},\ b_{31},\ b_{32}$.

$$A = \begin{bmatrix} -0.061994904923431 & -0.014321664726926 & 0.088269764978343 \\ -0.413117314149065 & 0.027004921378105 & 0.048738163633648 \\ -0.090220513163391 & 0.002986566608366 & 0.245902864428450 \end{bmatrix},$$

$$w = \begin{bmatrix} 0.622394316996030 & 0.313242750536090 & -0.011784503142076 \end{bmatrix}^T,$$

$$v = \begin{bmatrix} -0.062831671181596 & -0.008857653267082 & 0.147836760058631 \end{bmatrix}^T.$$

6. Methods with $s = 3$, $p = 4$, and $q = 3$. Solving stage order and order conditions

$$C_k = 0, \quad k = 1, 2, 3, \quad \tilde{C}_k = 0, \quad k = 1, 2, 3, 4,$$

we obtain an eleven-parameter family of methods of order $p = 4$ and stage order $q = 3$ depending on b_{ij}, $i, j = 1, 2, 3$, w_3, and v_3. In our search for A-stable methods we assume again that $c = [0, 1/2, 1]^T$. We compute the parameters b_{11}, b_{12}, and b_{13} to reduce the degree of stability polynomial to 3. As a result we obtain an eight-parameter family of methods depending on b_{21}, b_{22}, b_{23}, b_{31}, b_{32}, b_{33}, w_3, and v_3. The result of this search are produced on Fig. 5.4. The four cases in Fig. 5.4 correspond to:

- Case 1: $b_{21} = 2/3, b_{22} = 1, b_{23} = 1, b_{31} = -1/3, b_{32} = -1, b_{33} = 1/2$.
- Case 2: $b_{21} = 2, b_{22} = 1/2, b_{23} = 1, b_{31} = 1, b_{32} = -1, b_{33} = 1$.
- Case 3: $b_{21} = 2, b_{22} = 1, b_{23} = 1, b_{31} = 0, b_{32} = 1, b_{33} = 1$.
- Case 4: $b_{21} = 2, b_{22} = 1/2, b_{23} = 1/2, b_{31} = 1, b_{32} = 1/4, b_{33} = 1$.

In our search for algebraically stable methods, we found formulas for which

$$\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big)\Big|_{\xi=e^{it}} \geq -3.77 \cdot 10^{-11}, \tag{5.2.12}$$

$t \in [0, 2\pi]$. As before, this bound was obtained by dividing the interval $[0, 2\pi]$ into $n = 10000$ subintervals. Dividing $[0, 2\pi]$ into $n = 1000$ and $n = 100$ subintervals, these bounds are equal to 0.

The coefficients of a method satisfying (5.2.12) are

$$c = \begin{bmatrix} 0 & \frac{9}{10} & \frac{1}{5} \end{bmatrix}^T,$$

$$B = \begin{bmatrix} 1.923612711387117 & 0.332510317035363 & -2.361485324294705 \\ -0.042296580010526 & 0.264934173014572 & 0.753491461692750 \\ -0.754264567991751 & -0.285410192009791 & 1.590352848308200 \end{bmatrix},$$

$$A = \begin{bmatrix} -0.059505325459301 & 0.350609358761479 & -0.185741737429953 \\ 0.035823298712475 & -0.054110902952294 & -0.057841450456976 \\ 0.583885022389948 & -0.292809804314522 & -0.641753306382084 \end{bmatrix},$$

$$w = \begin{bmatrix} 0.085993246200685 & 0.374327773490317 & 0.659375459065431 \end{bmatrix}^T,$$

$$v = \begin{bmatrix} 0.123800116578393 & -0.056810353776912 & -0.186686241557912 \end{bmatrix}^T.$$

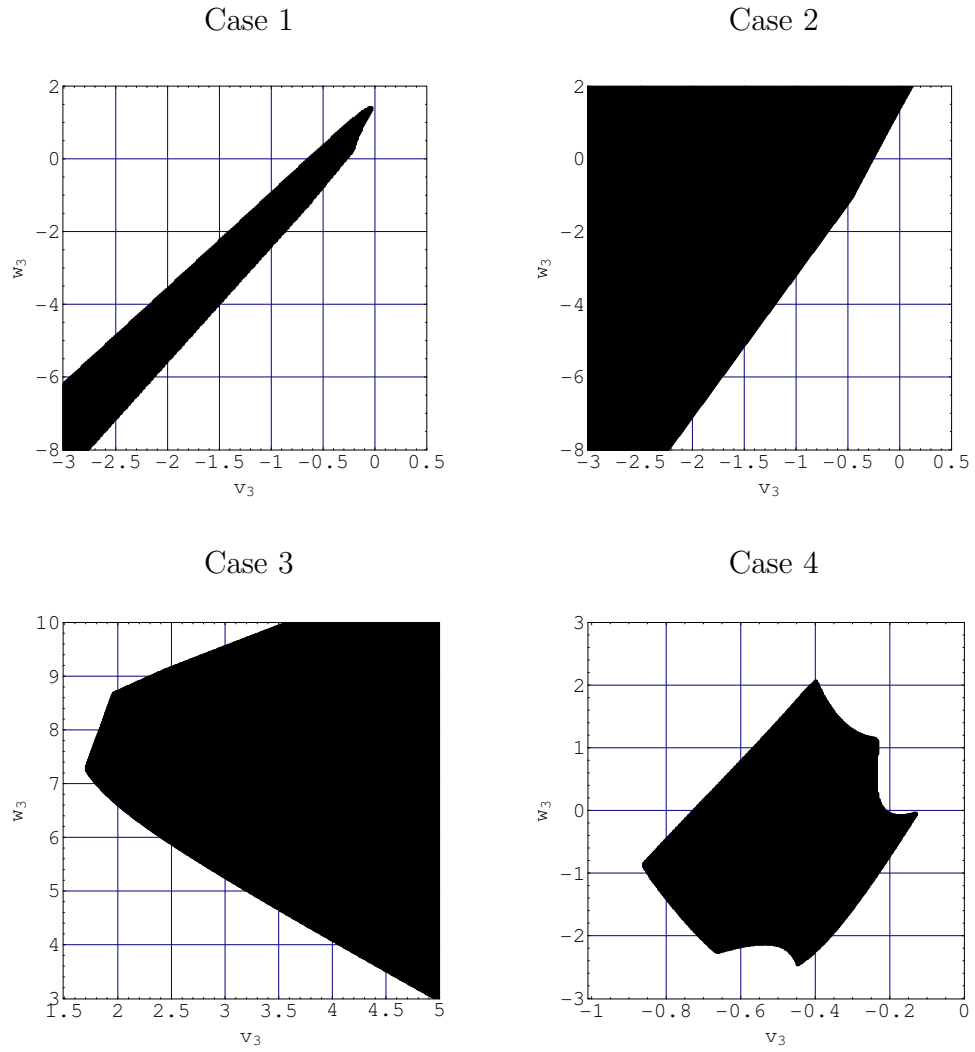

Figure 5.4: Regions of A-stability in the (w_3, v_3)-plane, for TSRK methods with $s = 3$ and $p = q = 4$ for specific values of the parameters

5.2.4 Examples of TSRK methods with $\vartheta \neq 0$ and $u \neq 0$.

As in Section 5.2.3, we will use Schur criterion to search for A-stable TSRK methods, and the criterion based on conditions 1-4 listed at the end of Sec-

tion 2.1.4 to search for TSRK methods which are algebraically stable. In the remainder of this section we will present the results of our search for such methods (5.2.6) with the number of stages $s = 1$, $s = 2$, and $s = 3$.

1. Methods with $s = 1$, $p = 2$, and $q = 2$. Assuming that $\mathbf{c} = c_1 = 1$ and solving stage order and order conditions

$$C_k = 0, \quad \tilde{C}_k = 0, \quad k = 1, 2,$$

we obtain a two-parameter family of methods of order $p = 2$ and stage order $q = 2$ with coefficients given by

$$\begin{array}{c|c|c} u & B & A \\ \hline \vartheta & w & v \end{array} = \begin{array}{c|c|c} u & \frac{1-u}{2} & \frac{1+3u}{2} \\ \hline \vartheta & \frac{1-\vartheta}{2} & \frac{1+3\vartheta}{2} \end{array}.$$

It can be verified using Schur criterion that these methods are A-stable if and only if

$$-1 < \vartheta < 0 \quad \text{and} \quad \frac{\vartheta(\vartheta+3)}{2(\vartheta+1)} \le u \le \frac{\vartheta}{2}$$

or $u = \vartheta = 0$. The last case corresponds again to the trapezoidal rule obtained in Section 5.2.3.

To search for methods which are algebraically stable we use the criteria (2.1.29) or (2.1.30) based on Albert theorem discussed in Section 2.1.4. The example of such a method is given by

$$\begin{array}{c|c|c} u & B & A \\ \hline \vartheta & w & v \end{array} = \begin{array}{c|c|c} -\frac{3}{4} & \frac{7}{8} & -\frac{5}{8} \\ \hline -\frac{1}{2} & \frac{3}{4} & -\frac{1}{4} \end{array}.$$

We can verify that choosing positive definite matrices $\mathbf{G}$ and $\mathbf{D}$,

$$\mathbf{G} = \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{bmatrix} = \begin{bmatrix} \frac{14}{13} & -\frac{9}{13} & -\frac{21}{52} \\ -\frac{9}{13} & \frac{1}{2} & \frac{4}{13} \\ -\frac{21}{52} & \frac{4}{13} & \frac{27}{104} \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} \frac{5}{26} \end{bmatrix},$$

the matrices $\mathbf{M}$ and $\tilde{\mathbf{M}}$ defined in Section 2.1.4 are nonnegative definite.

2. Methods with $s = 2$, $p = 4$, and $q = 4$. Solving stage order and order conditions

$$C_k = 0, \quad \tilde{C}_k = 0, \quad k = 1, 2, 3, 4,$$

we obtain a five-parameter family of methods of order $p = 4$ and stage order $q = 4$ depending on c_1, c_2, u_1, u_2, and ϑ. In our search for A-stable methods we computed first ϑ, u_1, and u_2 to reduce the degree of stability polynomial from 4 to 2. As a consequence we obtain a two-parameter family of methods depending on c_1 and c_2. The results of computer search are presented on Fig. 5.5.

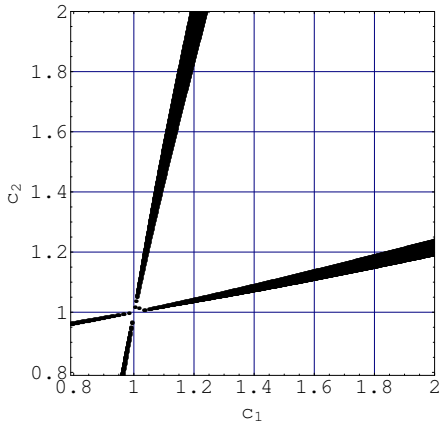

Figure 5.5: Regions of A-stability and in the (c_1, c_2)-plane, for TSRK methods with $s = 2$ and $p = q = 4$

We were not able to find algebraically stable methods in this class, and the best bound we were able to satisfy is

$$\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big)\Big|_{\xi=e^{it}} \geq -2.89 \cdot 10^{-3},$$

$t \in [0, 2\pi]$.

3. Methods with $s = 2$, $p = 4$, and $q = 3$. Solving stage order and order conditions

$$C_k = 0, \quad k = 1, 2, 3, \quad \tilde{C}_k = 0, \quad k = 1, 2, 3, 4,$$

we obtain a seven-parameter family of methods of order $p = 4$ and stage order $q = 3$ depending on c_1, c_2, b_{12}, b_{22}, u_1, u_2, and ϑ. To search for A-stable methods we assume that $c = [1/2, 1]^T$. We next determine ϑ to reduce the degree of $\tilde{p}(\eta, z)$ from 4 to 3. This leads to a four-parameter family of

methods depending on b_{12}, b_{22}, u_1, and u_2. The results of the search are presented in Fig 5.6 for selected values of the parameters u_1 and u_2.

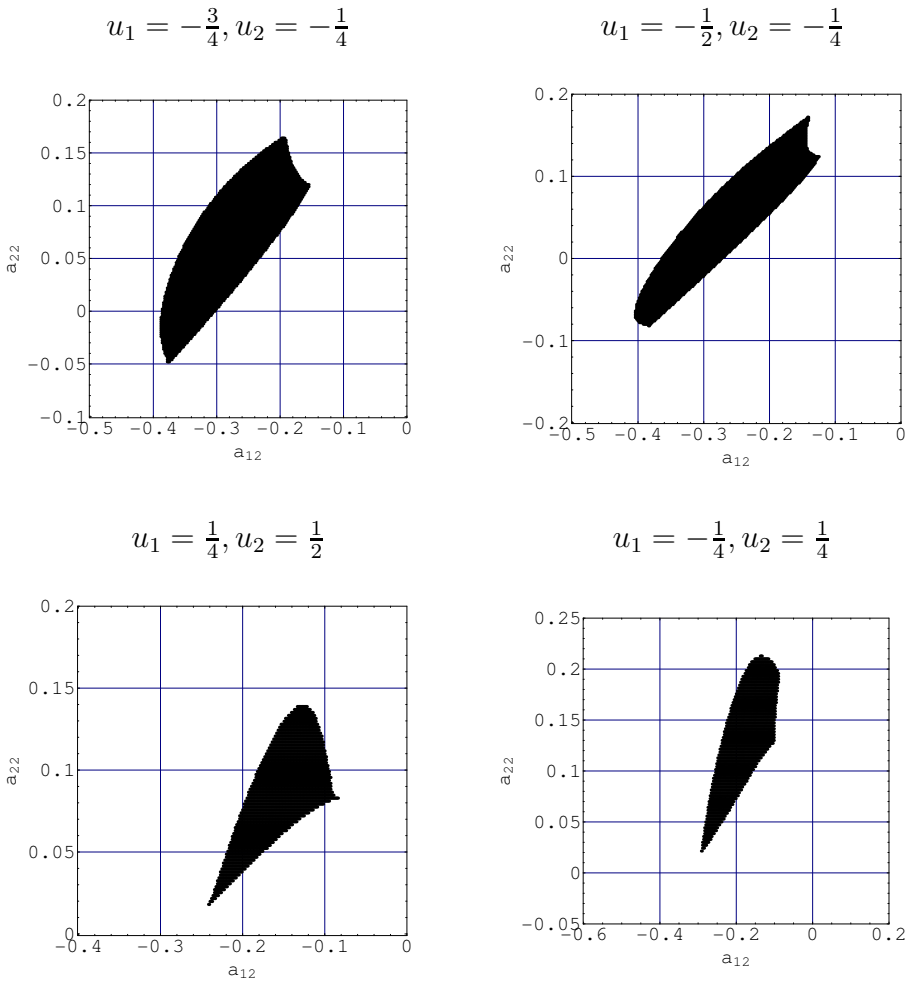

Figure 5.6: Regions of A-stability in the (b_{12}, b_{22})-plane, for TSRK methods with $s = 2$, $p = 4$, and $q = 3$, for specific values of the parameters u_1 and u_2.

As in the previous case, we were not able to find algebraically stable methods in this class, and the best bound we were able to satisfy is

$$\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big)\Big|_{\xi=e^{it}} \geq -7.70 \cdot 10^{-8},$$

$t \in [0, 2\pi]$. The coefficients of a method corresponding to this bound are $\vartheta = 0.045477710128446$,

$$c = \begin{bmatrix} 2.336580469857886 & 1.243897612851233 \end{bmatrix}^T,$$

$$u = \begin{bmatrix} -0.009140162241697 & -0.004971562691777 \end{bmatrix}^T,$$

$$B = \begin{bmatrix} 0.382519266813101 & 1.231791538880037 \\ -0.067916000101827 & 0.439694807309414 \end{bmatrix},$$

$$A = \begin{bmatrix} 0.123031161533802 & 0.590098340389249 \\ 0.161425752706173 & 0.705721490245697 \end{bmatrix},$$

$$w = \begin{bmatrix} -0.033116900308262 & -0.046491551194616 \end{bmatrix}^T,$$

$$v = \begin{bmatrix} 0.309391097083087 & 0.815695064548217 \end{bmatrix}^T.$$

We have also tried to find methods with c_1 and c_2 in the interval $[0, 1]$. We found methods for which

$$\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big)\Big|_{\xi=e^{it}} \geq -4.42 \cdot 10^{-5},$$

$t \in [0, 2\pi]$. The coefficients of a method corresponding to this bound are $\vartheta = 0.334852100666355$,

$$c = \begin{bmatrix} 0.817535264370424 & 0.111499700613041 \end{bmatrix}^T,$$

$$u = \begin{bmatrix} 0.381412710198958 & 0.549351922349542 \end{bmatrix}^T,$$

$$B = \begin{bmatrix} 0.261809682531266 & 0.716578941913569 \\ -0.055221137652820 & 0.331382626202203 \end{bmatrix},$$

$$A = \begin{bmatrix} 0.064452365556917 & 0.156106984567630 \\ 0.115480859585862 & 0.269209274827339 \end{bmatrix},$$

$$w = \begin{bmatrix} 0.484238639043203 & 0.651111475466369 \end{bmatrix}^T,$$

$$v = \begin{bmatrix} 0.058567341150044 & 0.140934645006739 \end{bmatrix}^T.$$

4. Methods with $s = 3$, $p = 4$, and $q = 4$. Solving stage order and order

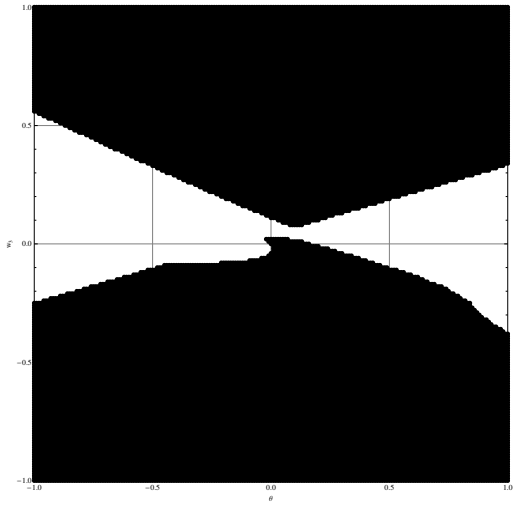

Figure 5.7: Regions of A- and L-stability in the (ϑ, v_3)-plane, for TSRK methods with $s = 3$, $p = 4$, and $q = 4$

conditions

$$C_k = 0, \quad \tilde{C}_k = 0, \quad k = 1, 2, 3, 4,$$

we obtain a fifteen-parameter family of methods of order $p = 4$ and stage order $q = 4$ depending on c_1, c_2, c_3, b_{ij}, $i = 1, 2, 3$, $j = 1, 2$, w_3, v_3, u_1, u_2, u_3, and ϑ. We used most of these parameters to reduce the degree of stability polynomial from 5 to 3 and to achieve L-stability. In Fig. 5.7 we present the A- and L-stable methods in the parameter space (ϑ, v_3).

We next investigate algebraic stability. In this family we have found methods for which

$$\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big)\Big|_{\xi=e^{it}} \geq 0,$$

$t \in [0, 2\pi]$, with small negative c_1 and with c_2 and c_3 in the interval $[0, 1]$. The coefficients of such a method are $\vartheta = -0.848157324846955$,

$$c = \begin{bmatrix} -0.001034365439338 & 0.460202200222726 & 0.904412078001496 \end{bmatrix}^T,$$

$$u = \begin{bmatrix} -4.095771377315513 & -8.224398492425298 & -1.491121282659948 \end{bmatrix}^T,$$

$$B = \begin{bmatrix} 0.313393425882524 & -0.230606333014524 & 0.057318614979789 \\ 0.145030029219694 & 0.433646707347174 & -0.082047057840003 \\ 0.298592649650861 & 0.485659580434091 & 0.175944782009046 \end{bmatrix},$$

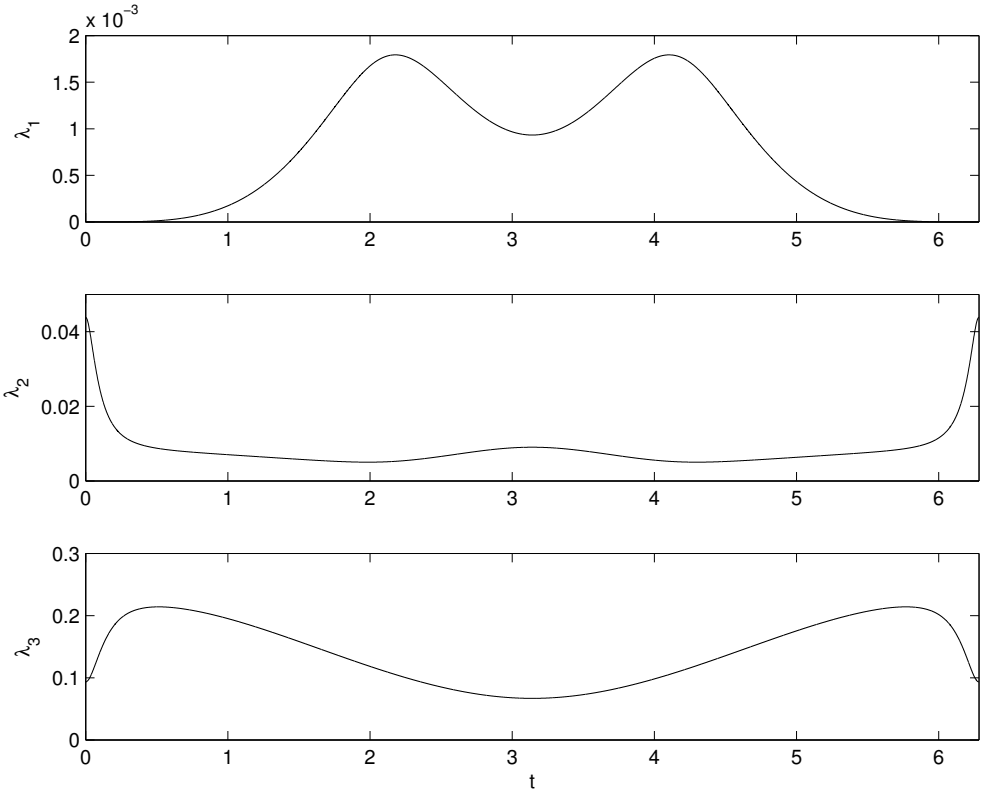

Figure 5.8: Eigenvalues $\lambda_1(t)$, $\lambda_2(t)$, and $\lambda_3(t)$ of the matrix $\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big)$ for $\xi = e^{it}$, $t \in [0, 2\pi]$

$$A = \begin{bmatrix} -0.589101054212143 & -2.620606437887970 & -1.027203958502527 \\ -1.339460229044111 & -4.705966457621601 & -2.215399284263721 \\ -0.206836462942642 & -0.983942583478110 & -0.356127170331698 \end{bmatrix},$$

$$w = \begin{bmatrix} 0.174157205530582 & 0.575265278532646 & 0.258760729183023 \end{bmatrix}^T,$$

$$v = \begin{bmatrix} -0.135787037506238 & -0.496815665848241 & -0.223737834738728 \end{bmatrix}^T.$$

The eigenvalues $\lambda_1(t)$, $\lambda_2(t)$, and $\lambda_3(t)$ of the matrix $\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big)$ for $\xi = e^{it}$, $t \in [0, 2\pi]$, are plotted in Fig. 5.8.

5. Methods with $s = 3$, $p = 4$, and $q = 3$. Solving stage order and order conditions

$$C_k = 0, \quad k = 1, 2, 3, \quad \tilde{C}_k = 0, \quad k = 1, 2, 3, 4,$$

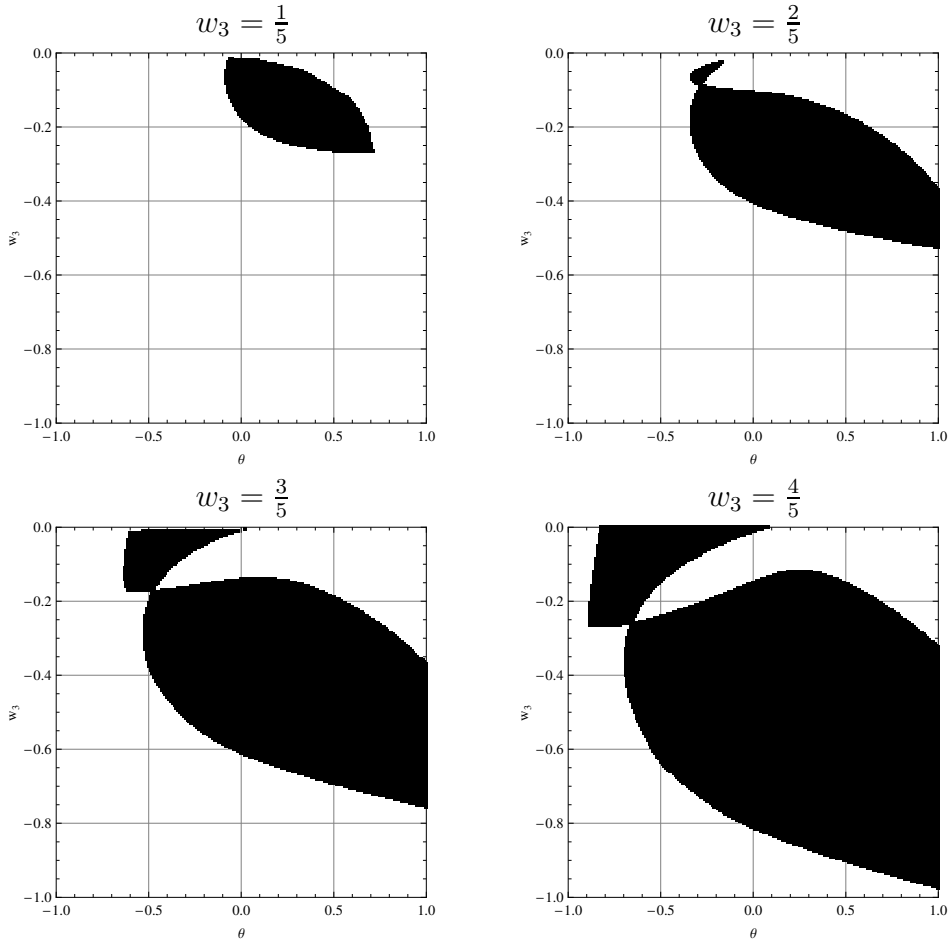

Figure 5.9: Regions of L-stability in the (ϑ, v_3)-plane, for TSRK methods with $s = 3$, $p = 4$, and $q = 3$, for selected values of w_3

we obtain an eighteen-parameter family of methods of order $p = 4$ and stage order $q = 4$ depending on c_1, c_2, c_3, b_{ij}, $i, j = 1, 2, 3$, w_3, v_3, u_1, u_2, u_3, and ϑ. Similarly as before we use most of these parameters to reduce the degree of stability polynomial from 5 to 3 and to obtain L-stability. The results of this search are presented on Fig 5.9 in the parameter space (ϑ, v_3) for selected values of the parameter w_3.

Concerning algebraic stability, we were looking for methods with the abscissas c_1, c_2, and c_3 in the interval $[0, 1]$. We have found such methods for

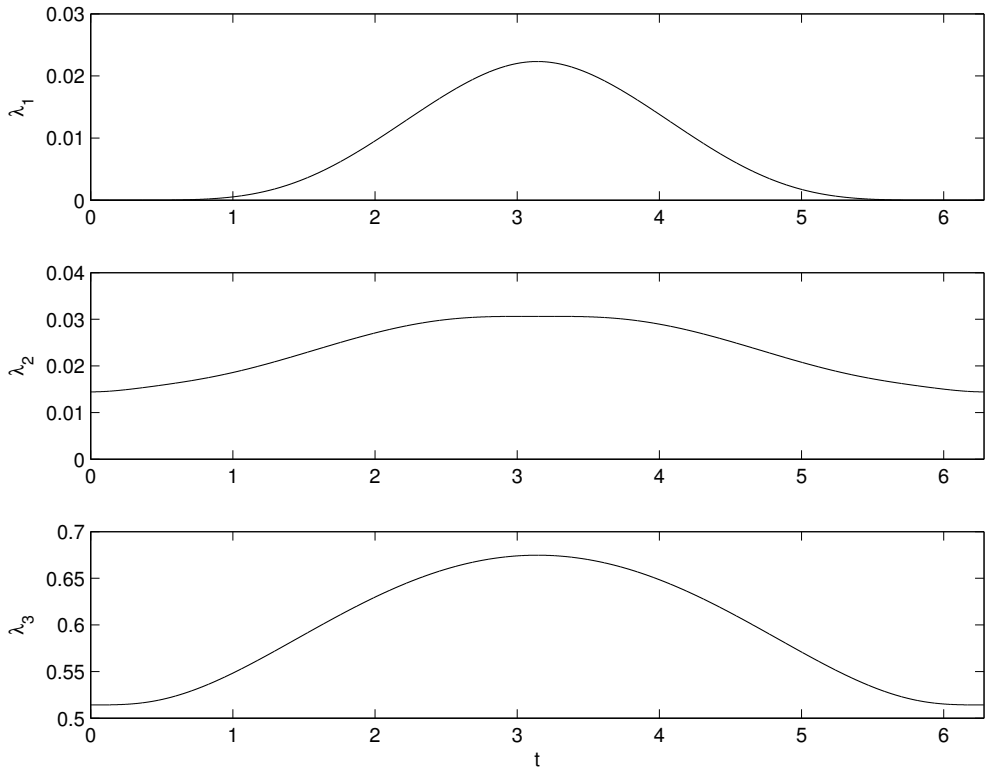

Figure 5.10: Eigenvalues $\lambda_1(t)$, $\lambda_2(t)$, and $\lambda_3(t)$ of the matrix $\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big)$ for $\xi = e^{it}$, $t \in [0, 2\pi]$

which

$$\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big)\Big|_{\xi=e^{it}} \geq 0,$$

$t \in [0, 2\pi]$. The coefficients of such a method are $\vartheta = -0.564128015497646$,

$$c = \begin{bmatrix} 0.070458343197336 & 0.681445056919784 & 0.952288029979521 \end{bmatrix}^T,$$

$$u = \begin{bmatrix} -1.270302241246827 & -1.227105131502048 & -1.431550985893559 \end{bmatrix}^T,$$

$$B = \begin{bmatrix} 0.421381377377544 & -0.267169299260385 & 0.109795290241695 \\ 0.043388505707860 & 1.529849094557059 & -0.802719654220106 \\ 1.009683545410309 & -0.770029595044248 & 0.856095443003834 \end{bmatrix},$$

$$A = \begin{bmatrix} -0.319430622728631 & -1.232896302162007 & 0.088475658482292 \\ -0.184576227521167 & -1.251755340576103 & 0.120153547470190 \\ -0.633408516424409 & -0.641844792716881 & -0.299759040142642 \end{bmatrix},$$

$$w = \begin{bmatrix} 0.400771327681170 & 0.573865935770521 & 0.071538532173898 \end{bmatrix}^T,$$

$$v = \begin{bmatrix} -0.165958972572252 & -0.440020959651672 & -0.004323878899313 \end{bmatrix}^T.$$

The eigenvalues $\lambda_1(t)$, $\lambda_2(t)$, and $\lambda_3(t)$ of the matrix $\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big)$ for $\xi = e^{it}$, $t \in [0, 2\pi]$, are plotted in Fig. 5.10.

5.2.5 An example of algebraically stable method by Hewitt and Hill

To put things in some perspective we analyze also the example of algebraically stable GLM of with $s = 2$, $p = 4$, and $q = 3$ constructed by Hewitt and Hill [125]. For this method the abscissa vector $\mathbf{c}$ and the coefficient matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{U}$, and $\mathbf{V}$ are given by

$$\mathbf{c} = \begin{bmatrix} 0 & \frac{193y^2-129y^4-297y^6-243y^8}{8} - \frac{1}{6} \end{bmatrix}^T,$$

$$\mathbf{A} = \begin{bmatrix} \frac{265}{864} + \frac{793y^2}{576} - \frac{5y^4}{6} - \frac{123y^6}{64} - \frac{27y^8}{16} & \frac{215}{864} - \frac{5299y^2}{576} + \frac{623y^4}{96} + \frac{915y^6}{64} + \frac{189y^8}{16} \\ \frac{101}{432} + \frac{3821y^2}{288} - \frac{463y^4}{48} - \frac{669y^6}{32} - \frac{135y^8}{8} & \frac{67}{432} + \frac{793y^2}{288} - \frac{5y^4}{3} - \frac{123y^6}{32} - \frac{27y^8}{8} \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{17y-1125y^3+828y^5+1783y^7+1458y^9}{24} & \frac{-11y+1125y^3-828y^5-1782y^7-1458y^9}{24} \end{bmatrix},$$

$$\mathbf{U} = \begin{bmatrix} 1 & -\frac{7y+9y^3}{16} \\ 0 & \frac{y+9y^3}{8} \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}.$$

Here $y = \pm\sqrt{z/3}$ and z is one of the two positive roots of the equation

$$9z^5 + 33z^4 + 46z^3 - 186z^2 + 9z + 1 = 0.$$

Choosing the root $z = 0.1032814360$ and $y = \sqrt{z/3}$, the decimal representation of the resulting method is

$$\mathbf{c} = [0, 0.6432188884]^T,$$

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{U} \\ \hline \mathbf{B} & \mathbf{V} \end{array}\right] = \left[\begin{array}{cc|cc} 0.3530415762 & -0.0595835887 & 1 & -0.08476931053 \\ 0.6782443859 & 0.2477498188 & 1 & 0.03037947026 \\ \hline 0.6666666667 & 0.3333333333 & 1 & 0 \\ -0.1598351741 & 0.2062215576 & 0 & 0.5 \end{array}\right].$$

It can be verified that for this method

$$\mathrm{He}\big(\tilde{\mathbf{D}}\mathbf{N}(\xi)\big)\Big|_{\xi=e^{it}} \geq -5.21 \cdot 10^{-11},$$

$t \in [0, 2\pi]$. This bound was obtained by dividing the interval $[0, 2\pi]$ into $n = 10000$ subintervals. Dividing $[0, 2\pi]$ into $n = 1000$ and $n = 100$ subintervals, the bounds are $-1.84 \cdot 10^{-13}$ and $-1.32 \cdot 10^{-15}$, respectively.

Chapter 6

Implementation of two-step collocation methods

In this chapter we discuss some practical issues concerning the variable stepsize implementation of two-step (almost) collocation methods (3.2.1) for the numerical solution of stiff differential systems (1.1.1). Such issues include the estimation of the principal part of the local discretization error and its assessment for large values of the step size, the comparison of possible strategies of step size control, the high order terms estimation in view of a variable order implementation, the computation of the stage values via Newton iterations. It is important to underline that the implementation of two-step collocation methods can benefit from their special structure, i.e. from the continuous approximant $P(t_n + sh)$ in (3.2.1), especially in order to perform a suitable step changing strategy. Many of these issues have been treated in our paper [95] (where we have also reported some of the numerical evidences included in Chapter 7) and are strongly linked to existing literature concerning the variable stepsize implementation of numerical methods for stiff ODEs (see, for instance, the monographes [42, 122, 138, 147, 185, 186, 188] and the references therein contained).

6.1 Starting procedure

Two-step (almost) collocation methods (3.2.1) and, more in general, TSRK methods (2.2.1) require a starting procedure for the computation of the miss-

ing starting values $y_1 \approx y(t_0 + h_0)$ and $Y^{[0]} \approx \big(y(t_0 + c_j h_0)\big)_{j=1}^{m}$, where h_0 is the initial stepsize. Hairer-Wanner [121] and Tracogna-Welfert [198] observed that for TSRK methods of order p and stage order $q = p$ or $q = p-1$, it is possible to choose as a starting procedure any continuous Runge-Kutta method (see, for instance, [167, 168, 169]) of uniform order p or to use a discrete Runge-Kutta method of order p with suitable stepsize. Of course, since the methods we aim to implement are A-stable and L-stable, the starting procedure must respect the same requirement: therefore, we base our starting procedure on the m-stage Runge–Kutta formulae

$$\begin{cases} y_1 = y_0 + h_0(b^T \otimes I)F^{[0]}, \\ Y^{[0]} = (e \otimes I)y_0 + h_0(A \otimes I)F^{[0]}, \end{cases} \tag{6.1.1}$$

based on the Gaussian nodes (see [36]), which are A-stable and of order $p = 2m$ and stage order $q = m$. In the integration of stiff systems, such methods suffer from order reduction (for instance, see [42]) and, therefore, their effective order is equal to m. These are implicit methods and, therefore, the stage vector $Y^{[0]}$ has to be determined by solving the nonlinear system in (6.1.1): its solution is computed by using Newton iterations, in the following way. We set

$$\Phi(Y^{[0]}) = Y^{[0]} - (e \otimes I)y_0 - h_0(A \otimes I)F^{[0]}$$

and aim to solve the system $\Phi(Y^{[0]}) = 0$, of dimension $md \times md$. We take as initial guess the vector

$$Y^{[0],0} = [y_0, \ldots, y_0]^T \in \mathbb{R}^{md},$$

and start the following Newton-type iterative procedure

$$Y^{[0],i+1} = Y^{[0],i} - \big(\partial\Phi(Y^{[0],i})\big)^{-1}\Phi(Y^{[0],i}), \tag{6.1.2}$$

for $i = 0, 1, \ldots, \nu - 1$, where

$$\partial\Phi(Y^{[0]}) = I_{md} - h(A \otimes I_d)J \in \mathbb{R}^{md \times md},$$

and J is the jacobian matrix of $F^{[0]}$, i.e. the block diagonal matrix

$$J = \begin{bmatrix} \partial f(Y_1^{[0]}) & & \\ & \ddots & \\ & & \partial f(Y_m^{[0]}) \end{bmatrix},$$

where $\partial f(Y_j^{[0]})$ is the jacobian matrix of f evaluated in $Y_j^{[0]}$, for $j = 1, 2, \ldots, m$. The expression (6.1.2) is equivalent to the linear system

$$-\partial\Phi(Y^{[0],i})\delta Y^{[0]} = \Phi(Y^{[0],i}), \tag{6.1.3}$$

where $\delta Y^{[0]} = Y^{[0],i+1} - Y^{[0],i}$. We next solve the system (6.1.3) with respect to $\delta Y^{[0]}$, for example by Gaussian elimination, and derive

$$Y^{[0],i+1} = Y^{[0],i} + \delta Y^{[0]}.$$

We stop the iterative scheme at the ν-th step, when $\|\delta Y^{[0]}\|_\infty < tol$ and $\|\Phi Y^{[0],\nu}\|_\infty < tol$, and take $Y^{[0]} = Y^{[0],\nu}$. We next compute the value $\hat{y}_1$, applying the Runge–Kutta methods twice, i.e. with two steps of stepsize $h_0/2$, in order to estimate the local error by means of Richardson extrapolation (see [119])

$$est(t_1) = \frac{2^{2m}(y_1 - \hat{y}_1)}{1 - 2^{2m}}.$$

It is well known that Richardson extrapolation is accurate, but also expensive. However, its usage in our implementation is only restrict to the very first step of the integration, so its contribution to the overall cost of the algorithm is not significant. Finally, the stepsize h_0 is adjusted until $\|est(t_1)\| < tol$. The strategy followed in order to compute a new stepsize at each step is described in details in Section 6.5.

6.2 Local error estimation

In this section we will look for estimates of the term $h^{p+1}\tilde{y}^{(p+1)}(t_n)$, as discussed in Section 3.2.5, in order to reliably estimate the local discretization error (3.2.29). Following the lines drawn in previous papers on discrete TSRK methods (for instance [16]), we aim to provide an estimation to $h^{p+1}\tilde{y}^{(p+1)}(t_n)$ having a similar shape with respect to the form of the method itself, i.e.

$$\begin{aligned} h^{p+1}\tilde{y}^{(p+1)}(t_n) &= \alpha_0 y_{n-1} + \alpha_1 y_n + \\ &+ h\sum_{j=1}^{m}\Big[\beta_j f(P(t_{n-1} + c_j h)) + \gamma_j f(P(t_n + c_j h))\Big], \end{aligned} \tag{6.2.1}$$

where α_0, α_1, β_j and γ_j, $j = 1, 2, \ldots, m$, are real parameters which can be derived using the following result.

Theorem 6.2.1 *Assume that the solution $\tilde{y}(t)$ to (3.2.31) is sufficiently smooth. Then the constants α_0, α_1, β_j, and γ_j, $j = 1, 2, \ldots, m$ appearing in (6.2.1) satisfy the system of equations*

$$\begin{cases} \alpha_0 + \alpha_1 = 0, \\ \dfrac{(-1)^k}{k!}\alpha_0 + \displaystyle\sum_{j=1}^{m}\left(\beta_j \frac{(c_j-1)^{k-1}}{(k-1)!} + \gamma_j \frac{c_j^{k-1}}{(k-1)!}\right) = 0, \\ \qquad k = 1, 2, \ldots, p, \\ \left(\dfrac{(-1)^{p+1}}{(p+1)!} - C_p(-1)\right)\alpha_0 + \displaystyle\sum_{j=1}^{m}\left(\beta_j \frac{(c_j-1)^p}{p!} + \gamma_j \frac{c_j^p}{p!}\right) = 1. \end{cases} \tag{6.2.2}$$

Proof:. Because of the localizing assumption, it is $\tilde{y}(t) = y_n$ (compare (4.3.10)) and since the method (3.2.1) is of order p it is locally of order $p+1$ and we have

$$y_{n-1} = \tilde{y}(t_{n-1}) - C_p(-1)h^{p+1}\tilde{y}^{(p+1)}(t_n) + O(h^{p+2}).$$

We have also

$$P(t_n + sh) = \tilde{y}(t_n + sh) + O(h^{p+1}), \quad s \in [-1, 1].$$

Substituting these relations and $y_n = \tilde{y}(t_n)$ into (6.2.1) we obtain

$$\begin{aligned} h^{p+1}\tilde{y}(t_n) &= \alpha_0\Big(\tilde{y}(t_n - h) - C_p(-1)h^{p+1}\tilde{y}^{(p+1)}(t_n)\Big) + \alpha_1\tilde{y}(t_n) \\ &+ h\sum_{j=1}^{m}\Big(\beta_j\tilde{y}'(t_n + (c_j-1)h) + \gamma_j\tilde{y}'(t_n + c_j h)\Big). \end{aligned}$$

Expanding $\tilde{y}(t_{n-1})$, $\tilde{y}(t_{n+1})$, $\tilde{y}'(t_n + (c_j - 1)h)$, and $\tilde{y}'(t_n + c_j h)$ into Taylor series around the point t_n and comparing the terms of order $O(h^k)$ for $k = 0, 1, \ldots, p+1$ leads to the system (6.2.2). □

Observe that (6.2.2) constitutes a system of $p+2$ equations with respect to $2m+2$ unknown coefficients α_0, α_1, β_j, and γ_j, $j = 1, 2, \ldots, m$. We have the following theorem.

Theorem 6.2.2 *Assume that $c_i \neq c_j$ and $c_i \neq c_j - 1$ for $i \neq j$. Then the system (6.2.2) corresponding to $p = m + r$, where $r = 1, 2, \ldots, m$, has a*

family of solutions depending on $m-r$ free parameters which may be chosen as, for example, $\beta_{r+1}, \beta_{r+2}, \ldots, \beta_m$ or $\gamma_{r+1}, \gamma_{r+2}, \ldots, \gamma_m$. In particular, if $r = m$ then the solution to the system (6.2.2) is unique. This system does not have solutions if $r = m + 1$.

Proof: The proof is similar to that of Theorem 3.2.3 and is therefore omitted. □

Other choices of free parameters than those indicated in Theorem 6.2.2 are also possible. For example, if $r = m - 1 \geq 1$ there is one free parameter which may be chosen as α_1, if $r = m - 2 \geq 1$ there are two free parameters which may be chosen as α_1 and α_0, if $r = m - 3 \geq 1$ there are three free parameters which may be chosen as α_1, α_0, and β_1 or γ_1, and if $r = m-k \geq 1$, $k > 3$, there are k free parameters which may be chosen as α_1, α_0, and β_j or γ_j, $j = 1, 2, \ldots, k-2$.

We will next illustrate the quality of this estimator in Chapter 7, where the results of variable stepsize experiments related to the solution of some stiff systems have been reported. The quality of this estimate for stiff differential systems can be remarkably improved by suitably applying the filtering technique, as discussed in the following section.

6.3 Assessment of the local error estimation for large step sizes

We have provided in the previous section the estimation (6.2.1) to the local truncation error, which is asymptotically correct for h_n tending to 0: this property can be tested by means of Taylor series expansion arguments, or may be obvious from its construction. However, in order to approach stiff systems, this property of correctness is not sufficient, since their solution also requires the usage of large stepsizes with respect to certain features of the problem. Shampine and Baca in [187] focused their attention on the assessment of the quality of the error estimate for large values of the stepsize, by using similar arguments as in the classical theory of absolute stability. We now specialize the results proposed in [187] to our class of methods (3.2.1).

Following the lines drawn in [187], we consider a restricted class of problems of the form $y' = Jy$, where J is a constant matrix that can be diagonalized by a similarity transformation $M^{-1}JM = \text{diag}(\xi_i)$. Then, it is sufficient

to consider the scalar problem

$$\begin{cases} y'(t) = \xi y, & t \geq 0, \\ y(0) = 1, \end{cases} \tag{6.3.1}$$

where $\xi \in \mathbb{C}$ is one the eigenvalues of J, which is supposed to have negative real part. The solution of the problem (6.3.1) is $y(t) = e^{\xi t}$ and, therefore,

$$Y^{[n]} = e^{\xi(t_n + \mathbf{c}h_n)} + O(h_n^{p+1})$$

and

$$Y^{[n-1]} = e^{\xi(t_n + (\mathbf{c}-\mathbf{e})h_{n-1})} + O(h_n^{p+1}).$$

As a consequence, we obtain

$$\begin{aligned} \mathrm{le}(t_n) &= e^{\xi t_n}\Big(e^{z\delta_n} - \varphi_0(1)e^{-z} - \varphi_1(1) - z\delta_n(v^T \otimes I)e^{z(\mathbf{c}-\mathbf{e})} \\ &- z\delta_n(w^T \otimes I)e^{z\mathbf{c}\delta_n}\Big) + O(z^{p+2}), \end{aligned}$$

where $z = \xi h_{n-1}$ and $\delta_n = \frac{h_n}{h_{n-1}}$. We next achieve an analogous expression also for the error estimate $\mathrm{est}(t_n)$, obtaining

$$\begin{aligned} \mathrm{est}(t_n) = C_p(1)e^{\xi t_n}\Big(\alpha_0 e^{-z} + \alpha_1 &+ z\delta_n(\beta^T \otimes I)e^{z(\mathbf{c}-\mathbf{e})} \\ &+ z\delta_n(\gamma^T \otimes I)e^{z\mathbf{c}\delta_n}\Big) + O(z^{p+2}). \end{aligned}$$

To investigate the behaviour of error estimates for large values of z, we define the functions $R_{\mathrm{le}}(z,\delta)$ and $R_{\mathrm{est}}(z,\delta)$, respectively defined by

$$\begin{aligned} R_{\mathrm{le}}(z,\delta) &= e^{z\delta} - \varphi_0(1)e^{-z} - \varphi_1(1) - z\delta(v^T \otimes I)e^{z\mathbf{c}\delta} - z\delta(w^T \otimes I)e^{z(\mathbf{c}-\mathbf{e})\delta}, \\ R_{\mathrm{est}}(z,\delta) &= \alpha_0 e^{-z} + \alpha_1 + z\delta(\beta^T \otimes I)e^{z(\mathbf{c}-\mathbf{e})\delta} + z\delta(\gamma^T \otimes I)e^{z\mathbf{c}\delta}, \end{aligned}$$

corresponding to $\mathrm{le}(t_n)$ and $\mathrm{est}(t_n)$. To assess the quality of $\mathrm{est}(t_n)$ for large step sizes, we examine the ratio

$$r(z,\delta) = \frac{R_{\mathrm{est}}(z,\delta)}{R_{\mathrm{le}}(z,\delta)}. \tag{6.3.2}$$

If $r(z,\delta) \sim \text{constant} \cdot z^{\mu}$, for $\mathrm{Re}(z) < 0$ as $|z| \to \infty$ with a positive integer μ, the error is grossly overestimated for large z. To compensate for

this, Shampine and Baca proposed in [187], in the context of RK methods, premultiplying $\mathrm{est}(t_n)$ by the so-called *filter matrix*,

$$\big(I - h_n \mathbf{J}(t_n)\big)^{-\mu}$$

where $\mathbf{J}(t_n)$ is an approximation to th Jacobian matrix of the problem (1.1.1) at the point t_n. This choice is suitable to damp the large, stiff error components.

Concerning two-step collocation methods (3.2.1), we observe that the ratio (6.3.2) behaves in the following way:

$$r(z,\delta) \sim -\frac{\alpha_1}{\varphi_0(1)}, \quad |z| \to \infty, \quad \mathrm{Re}(z) < 0,$$

and this behaviour would suggest that the original estimate $\mathrm{est}(t_n)$ can be used for all the values of the stepsize. However, it is important to observe that the denominator appearing in the above expression is equal to $\varphi_0(1)$ which, for zero-stability requirements, is always between -1 and 1: this means that, for small values of $\varphi_0(1)$ close to zero, the ratio $r(z,\delta)$ results to be very large and, therefore, the error estimate $\mathrm{est}(t_n)$ would not be reliable at all. On the contrary, the filtered estimation

$$\mathrm{est}'(t_n) = (I - h_n J)^{-1}\mathrm{est}(t_n), \tag{6.3.3}$$

corresponding to the filter matrix $(I - h_n J)^{-1}$ proposed in [187], results to be much more reliable than the original estimation $\mathrm{est}(t_n)$, as it has also been verified experimentally. As observed in [187], the improved error estimator does not alter the behaviour for small h_n but, on the opposite, it corrects the behaviour of the estimate for large values of h_n.

6.4 High order terms estimation

The estimation of the principal term of the local error (3.2.29) is a necessary tool in order to implement our methods in a variable stepsize environment. However, it is also important to estimate the quantities

$$h^{p+2} y^{(p+2)}(t_n), \quad h^{p+2} \frac{\partial f}{\partial y}\Big(y(t_n)\Big) y^{(p+1)}(t_n),$$

i.e. the terms of order $p+2$ appearing in the expansion (3.2.34) of the local discretization error, in view of a variable stepsize-variable order implementation of two-step collocation methods. Therefore, proceeding similarly as for the estimation of the principal term of the local error, we look for estimates of the type

$$h^{p+2}y^{(p+2)}(t_n) = \tilde{\alpha}_0 y_{n-1} + \tilde{\alpha}_1 y_n + \\ +h\sum_{j=1}^{m}\left(\tilde{\beta}_j f\Big(P(t_{n-1}+c_jh)\Big) + \tilde{\gamma}_j f\Big(P(t_n+c_jh)\Big)\right), \tag{6.4.1}$$

and

$$h^{p+2}\frac{\partial f}{\partial y}\Big(y(t_n)\Big)y^{(p+1)}(t_n) = \bar{\alpha}_0 y_{n-1} + \bar{\alpha}_1 y_n + \\ +h\sum_{j=1}^{m}\left(\bar{\beta}_j f\Big(P(t_{n-1}+c_jh)\Big) + \bar{\gamma}_j f\Big(P(t_n+c_jh)\Big)\right), \tag{6.4.2}$$

where the real parameters $\tilde{\alpha}_0, \tilde{\alpha}_1, \tilde{\beta}_j, \tilde{\gamma}_j,\ j = 1, 2, \ldots, m$, and $\bar{\alpha}_0, \bar{\alpha}_1, \bar{\beta}_j, \bar{\gamma}_j$, $j = 1, 2, \ldots, m$, can be computed according to the following result.

Theorem 6.4.1 *Setting*

$$\nu_i = \frac{(-1)^{p+i}}{(p+i)!} - C_p(-1),$$

and

$$\mu_\ell(t) = \frac{t^p}{p!} - C_{p+\ell}(t),$$

the parameters appearing in the estimates (6.4.1) and (6.4.2) satisfy the fol-

lowing systems of equations:

$$\begin{cases} \tilde{\alpha}_0 + \tilde{\alpha}_1 = 0, \\ \dfrac{(-1)^k}{k!}\tilde{\alpha}_0 + \displaystyle\sum_{j=1}^{m}\left(\tilde{\beta}_j \frac{(c_j-1)^{k-1}}{(k-1)!} + \tilde{\gamma}_j \frac{c_j^{k-1}}{(k-1)!}\right) = 0, \qquad k = 1, 2, \ldots, p, \\ \nu_1\tilde{\alpha}_0 + \displaystyle\sum_{j=1}^{m}\Big(\tilde{\beta}_j\mu_0(c_j-1) + \tilde{\gamma}_j\mu_0(c_j)\Big) = 0, \\ \nu_2\tilde{\alpha}_0 + \displaystyle\sum_{j=1}^{m}\Big(\tilde{\beta}_j\mu_1(c_j-1) + \tilde{\gamma}_j\mu_1(c_j)\Big) = 1, \\ G_{p+1}(-1)\tilde{\alpha}_0 + \displaystyle\sum_{j=1}^{m}\Big(\tilde{\beta}_j G_{p+1}(c_j-1) + \tilde{\gamma}_j G_{p+1}(c_j)\Big) = 0, \end{cases} \tag{6.4.3}$$

and

$$\begin{cases} \bar{\alpha}_0 + \bar{\alpha}_1 = 0, \\ \dfrac{(-1)^k}{k!}\bar{\alpha}_0 + \displaystyle\sum_{j=1}^{m}\left(\bar{\beta}_j \frac{(c_j-1)^{k-1}}{(k-1)!} + \bar{\gamma}_j \frac{c_j^{k-1}}{(k-1)!}\right) = 0, \qquad k = 1, 2, \ldots, p, \\ \nu_1\bar{\alpha}_0 + \displaystyle\sum_{j=1}^{m}\Big(\bar{\beta}_j\mu_0(c_j-1) + \bar{\gamma}_j\mu_0(c_j)\Big) = 0, \\ \nu_2\bar{\alpha}_0 + \displaystyle\sum_{j=1}^{m}\Big(\bar{\beta}_j\mu_1(c_j-1) + \bar{\gamma}_j\mu_1(c_j)\Big) = 0, \\ G_{p+1}(-1)\bar{\alpha}_0 + \displaystyle\sum_{j=1}^{m}\Big(\bar{\beta}_j G_{p+1}(c_j-1) + \bar{\gamma}_j G_{p+1}(c_j)\Big) = 1. \end{cases} \tag{6.4.4}$$

Proof: The proof follows along the same lines provided in [47]. Since the method (3.2.1) has order p, we have

$$\begin{aligned}
y_{n-1} &= y(t_n-1) - h^{p+1}C_p(-1)y^{(p+1)}(t_n) - h^{p+2}C_{p+1}(-1)y^{(p+2)}(t_n)\\
&\quad -h^{p+2}G_{p+1}(-1)\frac{\partial f}{\partial y}(y(t_n)y^{(p+1)}(t_n) + O(h^{p+3}),\\
Y_j^{[n-1]} &= y(t_n+(c_j-1)h) - h^{p+1}C_p(c_j-1)y^{(p+1)}(t_n) - h^{p+2}C_{p+1}(c_j-1)\\
&\quad y^{(p+2)}(t_n) - h^{p+2}G_{p+1}(c_j-1)\frac{\partial f}{\partial y}(y(t_n)y^{(p+1)}(t_n) + O(h^{p+3}),\\
Y_j^{[n]} &= y(t_n+c_jh) - h^{p+1}C_p(c_j)y^{(p+1)}(t_n) - h^{p+2}C_{p+1}(c_j)y^{(p+2)}(t_n)\\
&\quad -h^{p+2}G_{p+1}(c_j)\frac{\partial f}{\partial y}(y(t_n)y^{(p+1)}(t_n) + O(h^{p+3}),
\end{aligned}$$

and

$$P(t_n+sh) = \tilde{y}(t_n+sh) + O(h^{p+1}), \quad s\in[-1,1].$$

Substituting these relations and into (6.4.1) and taking into account the localizing assumption $y(t_n)=y_n$, we obtain

$$\begin{aligned}
h^{p+2}y^{(p+2)}(t_n) &= \tilde{\alpha}_0\Big(y(t_n-1) - h^{p+1}C_p(-1)y^{(p+1)}(t_n)\\
&\quad -h^{p+2}C_{p+1}(-1)y^{(p+2)}(t_n) - h^{p+2}G_{p+1}(-1)\tfrac{\partial f}{\partial y}(y(t_n))y^{(p+1)}(t_n)\Big) + \tilde{\alpha}_1 y(t_n)\\
&\quad +h\sum_{j=1}^{m}\Bigg(\beta_j\Big(y'(t_n+(c_j-1)h) - h^{p+1}C_p(c_j-1)y^{(p+1)}(t_n)\\
&\quad -h^{p+2}C_{p+1}(c_j-1)y^{(p+2)}(t_n) - h^{p+2}G_{p+1}(c_j-1)\tfrac{\partial f}{\partial y}(y(t_n))y^{(p+1)}(t_n)\Big)\\
&\quad +\gamma_j\Big(y'(t_n+c_jh) - h^{p+1}C_p(c_j)y^{(p+1)}(t_n) - h^{p+2}C_{p+1}(c_j)y^{(p+2)}(t_n)\\
&\quad -h^{p+2}G_{p+1}(c_j)\tfrac{\partial f}{\partial y}(y(t_n))y^{(p+1)}(t_n)\Big)\Bigg).
\end{aligned}$$

Expanding $y(t_{n-1})$, $y(t_{n+1})$, $y'(t_n+(c_j-1)h)$, and $y'(t_n+c_jh)$ into Taylor series around the point t_n and comparing the terms of order $O(h^k)$ for $k=0,1,\ldots,p+2$, leads to the system (6.4.3). Proceeding in analogous way for the term $h^{p+2}\frac{\partial f}{\partial y}(y(t_n))y^{(p+1)}(t_n)$ leads to the system (6.4.4) □

6.5 Stepsize control strategy

The strategy we apply in order to set an efficient control of the stepsize is strictly close to the ones developed in the classical solvers for systems of ODEs (1.1.1), and it is based on the following observations. Once we have derived an estimation to the local error, we can decide whether to increase or decrease the stepsize in the advancing from the point t_n to the point t_{n+1} according to the following control (see [16])

$$\|est(n)\| \leq Rtol \cdot \max\{\|y_{n-1}\|, \|y_n\|\} + Atol, \tag{6.5.1}$$

where $Atol$ and $Rtol$ are given absolute and relative tolerances. In our numerical experiments we have used $Atol = Rtol = tol$, where tol is the accuracy we aim to achieve. If the control (6.5.1) is not satisfied, the stepsize h_n is halved. Otherwise, the stepsize is accepted and a new stepsize for the following step is computed, according to a suitable control strategy. The standard step control strategy (see [119])

$$h_{n+1} = h_n \min\left(2, \left(\frac{fac \cdot tol}{\|est(t_n)\|}\right)^{\frac{1}{p+1}}\right), \tag{6.5.2}$$

which only depends on the estimate computed in the previous step, can often determine useless stepsize rejections, "with disruptive and wasteful increases and decreases" of the stepsize (see [42]). Gustafsson, Lundh and Söderlind [117, 191, 192] introduced a different stepsize control, the so-called PI stepsize control, based on control theory arguments. The PI control involve the estimation of the local errors related to the two most recent step points, as follows

$$h_{n+1} = h_n \cdot \min\left(2, \left(\frac{tol}{\|est(t_n)\|}\right)^{\sigma_1} \left(\frac{tol}{\|est(t_{n-1})\|}\right)^{\sigma_2}\right), \tag{6.5.3}$$

where σ_1 and σ_2 must be suitably chosen. In [122], [191], [192] the derivation of σ_1 and σ_2 is discussed, according to some control theory arguments. In our case, we have experimentally found some values for σ_1 and σ_2 in order to obtain a PI stepsize control which is competitive with the standard one in the implementation of our methods: they are $\sigma_1 \approx 0.3$ and $\sigma_2 \approx 0.04$.

When we advance from t_n to t_{n+1} with stepsize h_n, another problem occurs, i.e. the computation of the missing approximations $\tilde{y}_{n-1}$ to $y(\tilde{t}_{n-1})$, with $\tilde{t}_{n-1} = t_n - h_n$, and $\tilde{Y}_i^{[n-1]}$ to $y(\tilde{t}_{n-1} + c_i h_n)$, with $i = 1, 2, \ldots, m$. The computation of such approximations can be efficiently derived taking into account the special structure of the methods we are implementing: collocation methods are particularly suitable for the design of a numerical solver in a variable stepsize environment, since every time the stepsize changes, the missing approximations to the solution in previous points can be suitably computed by evaluating the collocation polynomial in these points. In fact, let us suppose that k is the minimum integer such that $\tilde{t}_{n-1}$ belongs to the interval $[t_k, t_{k+1}]$ of length h_k. The point $\tilde{t}_{n-1}$ is then uniquely determined by the time scaled variable

$$\tilde{s} = \frac{\tilde{t}_{n-1} - t_k}{h_k}.$$

The value of $\tilde{y}_{n-1}$ can next be computed by evaluating the collocation polynomial $P(t_k + sh_k)$ (3.2.1) in correspondence of $s = \tilde{s}$, obtaining

$$\tilde{y}_{n-1} = \varphi_0(\tilde{s}) y_{k-1} + \varphi_1(\tilde{s}) y_k + h_k \sum_{i=1}^{m} \Big(\chi_i(\tilde{s}) f(Y_i^{[k-1]}) + \psi_i(s) f(Y_i^{[k]}) \Big).$$

In an analogous way, we can derive the values of $\tilde{Y}_i^{[n-1]}$, $i = 1, 2, \ldots, m$. Let us assume that r is the minimum integer such that $\tilde{t}_{n-1} + c_i h$, for a fixed value of the index i, belongs to the interval $[t_r, t_{r+1}]$ of length h_r. The point $\tilde{t}_{n-1} + c_i h$ corresponds to the value of the time scaled variable

$$\tilde{s}_i = \frac{\tilde{t}_{n-1} + c_i h_r - t_r}{h_r}.$$

The missing value of $\tilde{Y}_i^{[n-1]}$ can then be computed by evaluating the collocation polynomial $P(t_r + sh_r)$ (3.2.1) in correspondence of $s = \tilde{s}_i$, obtaining

$$\tilde{Y}_i^{[n-1]} = \varphi_0(\tilde{s}_i) y_{r-1} + \varphi_1(\tilde{s}_i) y_r + h_r \sum_{j=1}^{m} \Big(\chi_j(\tilde{s}_i) f(Y_j^{[r-1]}) + \psi_j(s_i) f(Y_j^{[r]}) \Big).$$

6.6 Computation of the stage values

Two-step collocation methods are implicit formulae and, therefore, they

require the solution of a system of nonlinear equations of dimension $md \times md$ at each time step. We solve this system by means of Newton-type iterations, in the following way. We define

$$\Phi(Y^{[n]}) = Y^{[n]} - (u \otimes I)y_{n-1} - (\tilde{u} \otimes I)y_n - h\Big((A \otimes I)F^{[n-1]} - (B \otimes I)F^{[n]}\Big),$$

and aim to solve the system $\Phi(Y^{[n]}) = 0$. We take as initial guess the vector

$$Y^{[n],0} = [y_n, \ldots, y_n]^T \in \mathbb{R}^{md},$$

and start the following Newton-type iterative procedure

$$Y^{[n],i+1} = Y^{[n],i} - \big(\partial\Phi(Y^{[n],i})\big)^{-1}\Phi(Y^{[n],i}), \tag{6.6.1}$$

for $i = 0, 1, \ldots, \mu - 1$, where

$$\partial\Phi(Y^{[0]}) = I_{md} - h(B \otimes I_d)J \in \mathbb{R}^{md \times md},$$

and J is the jacobian matrix of $F^{[n]}$, i.e. the block diagonal matrix

$$J = \begin{bmatrix} \partial f(Y_1^{[n]}) & & \\ & \ddots & \\ & & \partial f(Y_m^{[n]}) \end{bmatrix},$$

where $\partial f(Y_j^{[n]})$ is the jacobian matrix of f evaluated in $Y_j^{[n]}$, for $j = 1, 2, \ldots, m$. The expression (6.6.1) is equivalent to the linear system

$$-\partial\Phi(Y^{[n],i})\delta Y^{[n]} = \Phi(Y^{[n],i}), \tag{6.6.2}$$

where $\delta Y^{[n]} = Y^{[n],i+1} - Y^{[n],i}$. We next solve the system (6.6.2) with respect to $\delta Y^{[n]}$, for instance by Gaussian elimination, and derive

$$Y^{[n],i+1} = Y^{[n],i} + \delta Y^{[n]}.$$

We stop the iterative scheme at the μ-th step, when $\|\delta Y^{[n]}\|_\infty < tol$ and $\|\Phi Y^{[n],\mu}\|_\infty < tol$, and take $Y^{[n]} = Y^{[n],\mu}$.

The numerical solution of the nonlinear system $\Phi(Y^{[n]}) = 0$ can be efficiently approached if the matrix B has a structured shape, e.g. lower triangular or diagonal: in these cases, instead of solving a nonlinear system of dimension md, we solve m successive or independent nonlinear systems of dimension d and, in particular, when these systems are independent, their solution can be fastly computed in a parallel environment. The construction of such numerical methods is treated in [98] and in Section 4.4.

Chapter 7

Numerical results

This chapter is devoted to the solution of a selection of stiff ODEs (1.1.1) by using the numerical methods developed in the previous chapters. In particular, we present two different groups of experiments: the first one contains some fixed stepsize numerical evidences aiming to confirm that the methods we have introduced do not suffer from order reduction and compare them with Runge-Kutta methods which, on the contrary suffer from this phenomenon; the second part, instead, regards some variable stepsize experiments obtained by implementing highly stable two-step almost collocation methods (3.2.1) and particularly aims to assert the reliability of the error estimate we have derived in Chapter 6, together with the effectiveness of the developed methods.

The selection of problems we aim to treat is the following:

1. the Prothero-Robinson problem [180]

$$\begin{cases} y'(t) = \lambda\big(y(t) - F(t)\big) + F'(t), \quad t \in [t_0, T], \\ y(t_0) = y_0, \end{cases} \tag{7.0.1}$$

where $\mathrm{Re}(\lambda) < 0$ and $F(t)$ is a slowly varying function on the interval $[t_0, T]$. In our experiments, we have considered $F(t) = \sin(t)$. As observed by Hairer and Wanner [122] in the context of Runge-Kutta methods this equation provides much new insight into the behaviour of numerical methods for stiff problems. This equation with $t_0 = 0$, $F(t) = \exp(\mu t)$, and $y_0 = 1$, was also used by Butcher [36] to investigate order reduction for Runge-Kutta-Gauss methods of order $p = 2s$;

2. the van der Pol oscillator (see VDPOL problem in [122])

$$\begin{cases} y_1' = y_2, & y_1(0) = 2, \\ y_2' = \big((1-y_1^2)y_2 - y_1\big)/\epsilon, & y_2(0) = -2/3, \end{cases} \tag{7.0.2}$$

$t \in [0, T]$, with stiffness parameter ϵ;

3. the Hires problem in [122]

$$\begin{cases} y_1' = -1.71y_1 + 0.43y_2 + 8.32y_3 + 0.0007 \\ y_2' = 1.71y_1 - 8.75y_2 \\ y_3' = -10.03y_3 + 0.43y_4 + 0.035y_5 \\ y_4' = 8.32y_2 + 1.71y_3 - 1.12y_4 \\ y_5' = -1.745y_5 + 0.43y_6 + 0.43y_7 \\ y_6' = -280y_6y_8 + 0.69y_4 + 1.71y_5 - 0.43y_6 + 0.69y_7 \\ y_7' = 280y_6y_8 - 1.81y_7 \\ y_8' = -280y_6y_8 + 1.81y_7 \\ y_1(0) = 1, \quad y_2(0) = \ldots = y_7(0) = 1 \quad y_8(0) = 0.0057, \end{cases} \tag{7.0.3}$$

with $t \in [0, 321.8122]$.

7.1 Fixed stepsize experiments

Similarly as in [122], in order to reduce the influence of round-off errors, we rewrite TSRK methods (2.2.1) in the form

$$\begin{cases} Z^{[n]} &= h(A \otimes I_m) f(Z^{[n-1]} + \theta(e \otimes I_m)y_{n-2} + (1-\theta)(e \otimes I_m)y_{n-1}) \\ &+ h(B \otimes I_m) f(Z^{[n]} + \theta(e \otimes I_m)y_{n-1} + (1-\theta)(e \otimes I_m)y_n), \\ y_n &= y_{n-1} + h(v^T \otimes I_m) f(Z^{[n-1]} + \theta(e \otimes I_m)y_{n-2} + (1-\theta)(e \otimes I_m)y_{n-1}) \\ &+ h(w^T \otimes I_m) f(Z^{[n]} + \theta(e \otimes I_m)y_{n-1} + (1-\theta)(e \otimes I_m)y_n), \end{cases} \tag{7.1.1}$$

$n = 2, 3, \ldots, N$, where

$$Z^{[n]} := Y^{[n]} - \theta(e \otimes I_m)y_{n-1} - (1-\theta)(e \otimes I_m)y_n$$

is usually smaller than $Y^{[n]}$. Define

$$\begin{aligned} G(Z^{[n]}) &:= Z^{[n]} - h(A \otimes I_m) f(Z^{[n-1]} + (e \otimes I_m) y_{n-2}) \\ &- h(B \otimes I_m) f(Z^{[n]} + (e \otimes I_m) y_{n-1}), \end{aligned}$$

and denote by $J = J(y_{n-1})$ the Jacobian of the right hand side of (7.1.1) computed at y_{n-1}. Then similarly as in [122] an approximation to $Z^{[n]}$ is computed by simplified Newton iterations

$$\begin{cases} \big(I - h(B \otimes J)\big) \Delta Z_k^{[n]} = -G(Z_k^{[n]}), \quad k = 0, 1, 2, \ldots, \\ Z_{k+1}^{[n]} = Z_k^{[n]} + \Delta Z_k^{[n]}, \end{cases}$$

with $Z_0^{[n]} = 0$ and a stopping criterion similar to that used in [122] in case of Runge-Kutta methods.

For each problem, we have implemented both methods with a fixed stepsize

$$h = (T - t_0)/2^k,$$

with several integer values of k, and listed norms of errors $\|e_h^{RK}(T)\|$ and $\|e_h^{TSRK}(T)\|$ at the endpoint of integration T and the observed order of convergence p computed from the formula

$$p = \frac{\log\big(\|e_h(T)\| / \|e_{h/2}(T)\|\big)}{\log(2)},$$

where $e_h(T)$ and $e_{h/2}(T)$ are errors corresponding to stepsizes h and $h/2$ for Runge-Kutta and TSRK methods.

7.1.1 Numerical results for continuous TSRK methods

In this section we will experimentally confirm that continuous TSRK methods described in Section 4.2 of order p and stage order $q = p$ do not suffer from order reduction in the integration of stiff differential systems, which is the case for classical Runge-Kutta formulae. This phenomenon, in fact, does not occur for continuous TSRK methods because they possess high stage order equal to their uniform order of convergence over the entire integration interval. On the other hand, Runge-Kutta methods do not possess the same feature, because their stage order is only equal to m, where m is the number of stages. To illustrate this we have applied the two-stage Runge-Kutta-Gauss

method of order $p = 4$ and stage order $q = 2$ and the continuous TSRK method of uniform order $p = 4$ given in the Example 8 of Section 4.2.9, i.e.

$$\begin{aligned}
&\varphi_0(s) = 0, \quad \varphi_1(s) = 1, \quad \chi_1(s) = -s^3\Big(\tfrac{63}{100} - \tfrac{223}{150}s + \tfrac{13}{10}s^2 - \tfrac{2}{5}s^3\Big),\\
&\chi_2(s) = \tfrac{125840873}{10156165010}s^3\Big(189 - 446s + 390s^2 - 120s^3\Big),\\
&\chi_3(s) = \tfrac{313000831}{6093699006}s^3(189 - 446s + 390s^2 - 120s^3), \quad \chi_4(s) = 0,\\
&\psi_1(s) = s\Big(1 - \tfrac{223}{126}s - \tfrac{110596774973233}{9597575934450}s^2 + \tfrac{48055456715852}{1599595989075}s^3\\
&\qquad\qquad - \tfrac{2838443145187}{106639732605}s^4 + \tfrac{873367121596}{106639732605}s^5\Big),\\
&\psi_2(s) = s^2\Big(\tfrac{75}{7} - \tfrac{13154611771291}{639838395630}s + \tfrac{671254535668}{35546577535}s^2 - \tfrac{80390326549}{7109315507}s^3 + \tfrac{24735485092}{7109315507}s^4\Big),\\
&\psi_3(s) = -s^2\Big(\tfrac{175}{9} - \tfrac{2867265551881}{54843291054}s + \tfrac{575594042414}{9140548509}s^2 - \tfrac{130770083795}{3046849503}s^3 + \tfrac{40236948860}{3046849503}s^4\Big),\\
&\psi_4(s) = s^2\Big(\tfrac{21}{2} - \tfrac{28900702732187}{914054850900}s + \tfrac{2081690316751}{50780825050}s^2 - \tfrac{290054503193}{10156165010}s^3 + \tfrac{44623769722}{5078082505}s^4\Big),
\end{aligned} \tag{7.1.2}$$

with $c = [0, \frac{7}{10}, \frac{9}{10}, 1]^T$. Let us first consider the results obtained for the Prothero-Robinson problem in the interval $[0, 50]$, which are presented in Table 7.1.1 and Table 7.1.1, for the Runge-Kutta-Gauss method and the continuous TSRK one respectively, in correspondence of several values for the stiffness parameter λ.

We can observe that in the case $\lambda = -10^3$, for which the Prothero-Robinson problem is mildly stiff, both methods are convergent with expected order $p = 4$. However, for $\lambda = -10^5$, the problem in stiff and the Runge-Kutta-Gauss method exhibits the order reduction phenomenon and its order of convergence drops to about $p = 2$ which corresponds to the stage order $q = 2$. This is not the case for TSRK method which preserves order of convergence $p = q = 4$, which leads to higher accuracy.

We next consider the Van der Pol oscillator, which is observed in the interval $[0, 3/4]$, i.e. for the slowly varying parts of the solution, where the problem is stiff for small values of the parameter ϵ (the problem is not stiff on the interval where the solution is changing rapidly). The results are presented in Table 7.1.1 and Table 7.1.1, for several values of the parameter ϵ.

Also in this case we can observe that for the values of $\epsilon = 10^{-1}$ and $\epsilon = 10^{-3}$ for which the problem (7.0.2) is not stiff and mildly stiff both methods are convergent with expected order $p = 4$. However, for small

	$\lambda = -10^3$			$\lambda = -10^5$	
k	$e_h^{RKG}(T)$	p	k	$e_h^{RKG}(T)$	p
10	$1.55 \cdot 10^{-5}$				
11	$7.80 \cdot 10^{-7}$	3.89	7	$1.11 \cdot 10^{-3}$	
12	$4.94 \cdot 10^{-8}$	3.98	8	$2.78 \cdot 10^{-4}$	2.00
13	$3.09 \cdot 10^{-9}$	3.99	9	$6.80 \cdot 10^{-5}$	2.02
14	$1.93 \cdot 10^{-10}$	4.00	10	$1.68 \cdot 10^{-5}$	2.01

Table 7.1: Numerical results for Runge-Kutta-Gauss method of order $p = 4$ and stage order $q = 2$ for the Prothero-Robinson problem

	$\lambda = -10^3$			$\lambda = -10^5$	
k	$e_h^{TSRK}(T)$	p	k	$e_h^{TSRK}(T)$	p
10	$3.29 \cdot 10^{-11}$				
11	$2.11 \cdot 10^{-12}$	3.97	7	$1.12 \cdot 10^{-9}$	
12	$1.34 \cdot 10^{-13}$	3.98	8	$7.75 \cdot 10^{-11}$	3.86
13	$8.43 \cdot 10^{-15}$	3.98	9	$4.97 \cdot 10^{-12}$	3.96
14	$5.55 \cdot 10^{-16}$	3.92	10	$3.03 \cdot 10^{-13}$	4.03

Table 7.2: Numerical results for continuous TSRK method (7.1.2) of uniform order $p = 4$ for the Prothero-Robinson problem

	$\epsilon = 10^{-1}$		$\epsilon = 10^{-3}$		$\epsilon = 10^{-6}$	
k	$\|e_h^{RKG}(T)\|$	p	$\|e_h^{RKG}(T)\|$	p	$\|e_h^{RKG}(T)\|$	p
6	$1.88 \cdot 10^{-8}$		$2.25 \cdot 10^{-4}$		$1.49 \cdot 10^{-3}$	
7	$1.18 \cdot 10^{-9}$	4.00	$1.68 \cdot 10^{-5}$	3.74	$3.71 \cdot 10^{-4}$	2.01
8	$8.21 \cdot 10^{-11}$	3.84	$1.11 \cdot 10^{-6}$	3.93	$8.84 \cdot 10^{-5}$	2.07
9	$1.43 \cdot 10^{-11}$	2.52	$7.02 \cdot 10^{-8}$	3.98	$1.87 \cdot 10^{-5}$	2.24

Table 7.3: Numerical results for Runge-Kutta-Gauss method of order $p = 4$ and stage order $q = 2$ for the Van der Pol problem

	$\epsilon = 10^{-1}$		$\epsilon = 10^{-3}$		$\epsilon = 10^{-6}$	
k	$\|e_h^{TSRK}(T)\|$	p	$\|e_h^{TSRK}(T)\|$	p	$\|e_h^{TSRK}(T)\|$	p
6	$5.82 \cdot 10^{-8}$		$1.58 \cdot 10^{-5}$		$1.54 \cdot 10^{-5}$	
7	$3.66 \cdot 10^{-9}$	3.99	$1.17 \cdot 10^{-6}$	3.75	$1.09 \cdot 10^{-6}$	3.81
8	$2.32 \cdot 10^{-10}$	3.98	$7.85 \cdot 10^{-8}$	3.90	$7.34 \cdot 10^{-8}$	3.90
9	$1.46 \cdot 10^{-11}$	3.99	$4.80 \cdot 10^{-9}$	4.03	$4.75 \cdot 10^{-9}$	3.94

Table 7.4: Numerical results for TSRK method (7.1.2) of order $p = 4$ and stage order $q = 4$ for the Van der Pol problem

k	$\|e_h^{RKG}(T)\|$	p
8	$3.05 \cdot 10^{-6}$	
9	$6.42 \cdot 10^{-7}$	2.25
10	$1.47 \cdot 10^{-7}$	2.12
11	$3.52 \cdot 10^{-8}$	2.06
12	$8.62 \cdot 10^{-9}$	2.03

Table 7.5: Numerical results for Runge-Kutta-Gauss method of order $p = 4$ and stage order $q = 2$ for the Hires problem

k	$\|e_h^{TSRK}(T)\|$	p
6	$4.85 \cdot 10^{-5}$	
7	$3.31 \cdot 10^{-6}$	3.87
8	$2.16 \cdot 10^{-7}$	3.93

Table 7.6: Numerical results for TSRK method (7.1.2) of order $p = 4$ and stage order $q = 4$ for the Hires problem

values of ϵ ($\epsilon = 10^{-6}$) for which the van der Pol oscillator is stiff the Runge-Kutta-Gauss method exhibits order reduction phenomenon and its order of convergence drops to about $p = 2$ which corresponds to the stage order $q = 2$. This is not the case for TSRK method which preserves order of convergence $p = q = 4$, which leads to higher accuracy.

We conclude our analysis presenting the results obtained for the Hires problem, included in Table (7.1.1) and Table (7.1.1). Also in this case, the order reduction phenomenon is evident for the Runge-Kutta-Gauss method, while it is not present on the continuous TSRK method considered.

7.1.2 Numerical experiments with diagonally implicit two-step collocation methods

In this section we present some numerical evidences arising from the appli-

cation of the two-step collocation methods derived in Section 4.4 equivalent to TSRK methods (2.2.1) with B lower triangular, in order to confirm the theoretical expectations on the order of accuracy. Also in this case we aim to provide an experimental confirmation that, since the derived methods have high stage order (equal to the order of convergence), they do not suffer from order reduction in the integration of stiff differential systems, which is the case for diagonally implicit Runge–Kutta formulae, whose effective order is only equal to m, where m is the number of stages. To illustrate these features we compare the following numerical methods:

- TS3: L-stable two-step almost collocation method (3.2.1), with $m = 2$ and

$$\begin{aligned}
\varphi_0(s) &= s^2\left(-\frac{21}{10}+\frac{7}{5}s\right), & \varphi_1(s) &= 1+\frac{21}{10}s^2-\frac{7}{5}s^3,\\
\chi_1(s) &= -s\left(1-\frac{23}{10}s+\frac{6}{5}s^2\right), & \chi_2(s) &= s\left(1-\frac{191}{40}s+\frac{57}{20}s^2\right),\\
\psi_1(s) &= -s^2\left(\frac{9}{8}-\frac{3}{4}s\right), & \psi_2(s) &= s\left(1+\frac{3}{2}s-s^2\right),
\end{aligned}$$

  of order 3 and stage order 3, equivalent to the type 2 TSRK method

$$\begin{array}{c|cc|cc}
14/5 & -12/5 & -57/10 & 3/2 & 0\\
-7/10 & 1/10 & -37/40 & -3/8 & 3/2\\
\hline
-7/10 & 1/10 & -37/40 & -3/8 & 3/2
\end{array}, \qquad c=[2,1]^T;$$

  the method has been obtained in correspondence of the point $(c_1, \alpha_1) = (2, -1/2)$ of the shaded region reported in Figure 4.1.

- SDIRK3: two-stage singly diagonally implicit Runge–Kutta method [36]

$$\begin{array}{c|cc}
(3+\sqrt{3})/6 & (3+\sqrt{3})/6 & 0\\
(3-\sqrt{3})/6 & -\sqrt{3}/3 & (3+\sqrt{3})/6\\
\hline
 & 1/2 & 1/2
\end{array}$$

  of order 3 and stage order 2.

We apply these methods to the van der Pol oscillator (7.0.2) in the interval $[0, 3/4]$, for different values of the stiffness parameter ϵ. We have implemented

	$\epsilon = 10^{-1}$		$\epsilon = 10^{-3}$		$\epsilon = 10^{-6}$	
k	$\|e_h^{\text{SDIRK3}}\|$	p	$\|e_h^{\text{SDIRK3}}\|$	p	$\|e_h^{\text{SDIRK3}}\|$	p
7	$2.45 \cdot 10^{-7}$		$4.94 \cdot 10^{-4}$		$8.78 \cdot 10^{-4}$	
8	$3.41 \cdot 10^{-8}$	2.84	$1.06 \cdot 10^{-4}$	2.21	$2.29 \cdot 10^{-4}$	1.93
9	$4.49 \cdot 10^{-9}$	2.92	$2.00 \cdot 10^{-5}$	2.40	$5.85 \cdot 10^{-5}$	1.97
10	$5.78 \cdot 10^{-10}$	2.95	$3.31 \cdot 10^{-6}$	2.59	$1.47 \cdot 10^{-5}$	1.98
11	$7.52 \cdot 10^{-11}$	2.94	$4.92 \cdot 10^{-7}$	2.74	$3.69 \cdot 10^{-6}$	1.99
12	$9.81 \cdot 10^{-12}$	2.93	$6.80 \cdot 10^{-8}$	2.85	$9.27 \cdot 10^{-7}$	2.00

Table 7.7: Numerical results for SDIRK3 method on the Van der Pol oscillator

	$\epsilon = 10^{-1}$		$\epsilon = 10^{-3}$		$\epsilon = 10^{-6}$	
k	$\|e_h^{\text{TS3}}\|$	p	$\|e_h^{\text{TS3}}\|$	p	$\|e_h^{\text{TS3}}\|$	p
7	$1.57 \cdot 10^{-5}$		$6.36 \cdot 10^{-4}$		$8.32 \cdot 10^{-4}$	
8	$2.12 \cdot 10^{-6}$	2.89	$8.49 \cdot 10^{-5}$	2.90	$1.11 \cdot 10^{-4}$	2.90
9	$2.78 \cdot 10^{-7}$	2.93	$1.15 \cdot 10^{-5}$	2.88	$1.47 \cdot 10^{-5}$	2.92
10	$3.56 \cdot 10^{-8}$	2.96	$1.57 \cdot 10^{-6}$	2.87	$1.90 \cdot 10^{-6}$	2.94
11	$4.51 \cdot 10^{-9}$	2.98	$2.88 \cdot 10^{-7}$	2.88	$2.43 \cdot 10^{-7}$	2.97
12	$5.70 \cdot 10^{-10}$	2.99	$2.84 \cdot 10^{-8}$	2.90	$3.06 \cdot 10^{-8}$	2.98

Table 7.8: Numerical results for TS3 on the Van der Pol problem

both methods with a fixed stepsize $h = (T - t_0)/2^k$, for several integer values of k, and listed norms of errors $\|e_h^{\mathrm{TS3}}(T)\|$ and $\|e_h^{\mathrm{SDIRK3}}(T)\|$ at the endpoint of integration T and the observed order of convergence p.

We can observe that for the values of $\epsilon = 10^{-1}$ and $\epsilon = 10^{-3}$ for which the problem (7.0.2) is not stiff and mildly stiff both methods are convergent with expected order $p = 3$. However, for small values of ϵ ($\epsilon = 10^{-6}$) for which the van der Pol oscillator is stiff, the SDIRK3 method exhibits order reduction phenomenon and its order of convergence drops to about $p = 2$ which corresponds to the stage order $q = 2$. This is not the case for TS3 method which preserves order of convergence $p = q = 3$, which leads to higher accuracy.

7.1.3 Numerical results for TSRK methods with Inherent Quadratic Stability

In this section we aim to illustrate that the TSRK methods with IQS discussed in Chapter 5 do not suffer from order reduction, which is the case for classical Runge-Kutta formulae. To illustrate this we have applied the Runge-Kutta-Gauss method of order $p = 4$ and stage order $q = 2$ and TSRK method of order $p = 4$ and stage order $q = 4$ given in Example 4 in Chapter 5, i.e.

$$A = \begin{bmatrix} -\frac{73571}{418565} & \frac{316790}{450193} & -\frac{383309}{370547} & -\frac{1102057}{1459404} \\ -\frac{324116}{495273} & \frac{3108022}{1186313} & -\frac{2008351}{521461} & -\frac{1905671}{677809} \\ -\frac{813738}{787901} & \frac{4021146}{972541} & -\frac{6409321}{1054477} & -\frac{6349415}{1430988} \\ -\frac{426460}{370257} & \frac{4154204}{900915} & -\frac{12185608}{1797671} & -\frac{6621076}{1338039} \end{bmatrix},$$

$$B = \begin{bmatrix} \frac{1082275}{789096} & -\frac{47158}{1102905} & -\frac{20658}{230377} & \frac{16548}{733283} \\ \frac{2053468}{392523} & \frac{173881}{1660851} & -\frac{337517}{836884} & \frac{86197}{880374} \\ \frac{13765224}{1684843} & \frac{119918}{620675} & -\frac{387828}{932779} & \frac{214966}{1621163} \\ \frac{8694859}{954168} & \frac{68987}{727614} & -\frac{198815}{935168} & \frac{90358}{331129} \end{bmatrix},$$

$$v = \begin{bmatrix} -\frac{426460}{370257} & \frac{4154204}{900915} & -\frac{12185608}{1797671} & -\frac{6621076}{1338039} \end{bmatrix}^T,$$

$$w = \begin{bmatrix} \frac{8694859}{954168} & \frac{68987}{727614} & -\frac{198815}{935168} & \frac{90358}{331129} \end{bmatrix}^T,$$

	$\epsilon = 10^{-1}$		$\epsilon = 10^{-3}$		$\epsilon = 10^{-6}$	
N	$\|e_h^{RKG}(T)\|$	p	$\|e_h^{RKG}(T)\|$	p	$\|e_h^{RKG}(T)\|$	p
32	$3.02 \cdot 10^{-7}$		$2.19 \cdot 10^{-3}$		$5.83 \cdot 10^{-3}$	
64	$1.88 \cdot 10^{-8}$	4.00	$2.25 \cdot 10^{-4}$	3.28	$1.49 \cdot 10^{-3}$	1.97
128	$1.18 \cdot 10^{-9}$	4.00	$1.68 \cdot 10^{-5}$	3.74	$3.71 \cdot 10^{-4}$	2.01
256	$8.21 \cdot 10^{-11}$	3.84	$1.11 \cdot 10^{-6}$	3.93	$8.84 \cdot 10^{-5}$	2.07
512	$1.43 \cdot 10^{-11}$	2.52	$7.02 \cdot 10^{-8}$	3.98	$1.87 \cdot 10^{-5}$	2.24

Table 7.9: Numerical results for Runge-Kutta-Gauss method of order $p = 4$ and stage order $q = 2$

with $c = [0, \frac{1}{3}, \frac{2}{3}, 1]^T$, to the Van der Pol problem (7.0.2). We have implemented both methods with a fixed stepsize $h = T/N$, and observed the order of convergence of numerical approximations to the slowly varying parts of the solution, where the problem is stiff for small values of the parameter ϵ.

The results of numerical experiments for fixed stepsize implementations of Runge-Kutta-Gauss method of order $p = 4$ and stage order $q = 2$ and TSRK method with IQS of order $p = 4$ and stage order $q = 4$ correspond to $T = 2/3$, $h = T/N$, and $N = 32$, 64, 128, 256 and 512. We have listed norms of errors $\|e_h^{RKG}(T)\|$ and $\|e_h^{TSRK}(T)\|$ at the endpoint of integration T and the observed order of convergence p.

We can observe that for the values of $\epsilon = 10^{-1}$ and $\epsilon = 10^{-3}$ for which the problem (7.0.2) is not stiff and mildly stiff both methods are convergent with expected order $p = 4$ (although there is an unexpected reduction to order $p = 2.52$ only for Runge-Kutta method for $N = 512$). However, for small values of ϵ ($\epsilon = 10^{-6}$) for which the van der Pol oscillator (7.0.2) is stiff the Runge-Kutta Gauss method exhibits order reduction phenomenon and its order of convergence drops to about $p = 2$ which corresponds to the stage order $q = 2$. This is not the case for TSRK method which preserves order of convergence $p = q = 4$, which leads to higher accuracy.

We also present in Fig. 7.1 the comparison of fixed stepsize implementations of SDIRK method of order $p = 3$ and stage order $q = 2$ (see [36], p. 234) and TSRK method of order $p = 3$ and stage order $q = 3$ given in

	$\epsilon = 10^{-1}$		$\epsilon = 10^{-3}$		$\epsilon = 10^{-6}$	
N	$\|e_h^{TSRK}(T)\|$	p	$\|e_h^{TSRK}(T)\|$	p	$\|e_h^{TSRK}(T)\|$	p
32	$7.83 \cdot 10^{-7}$		$1.85 \cdot 10^{-4}$		$2.44 \cdot 10^{-4}$	
64	$1.03 \cdot 10^{-7}$	2.93	$1.94 \cdot 10^{-5}$	3.25	$2.65 \cdot 10^{-5}$	3.21
128	$7.67 \cdot 10^{-9}$	3.75	$1.57 \cdot 10^{-6}$	3.62	$2.20 \cdot 10^{-6}$	3.59
256	$5.17 \cdot 10^{-10}$	3.89	$1.09 \cdot 10^{-7}$	3.85	$1.59 \cdot 10^{-7}$	3.79
512	$4.21 \cdot 10^{-11}$	3.62	$6.52 \cdot 10^{-9}$	4.06	$1.08 \cdot 10^{-8}$	3.89

Table 7.10: Numerical results for TSRK method of order $p = 4$ and stage order $q = 4$

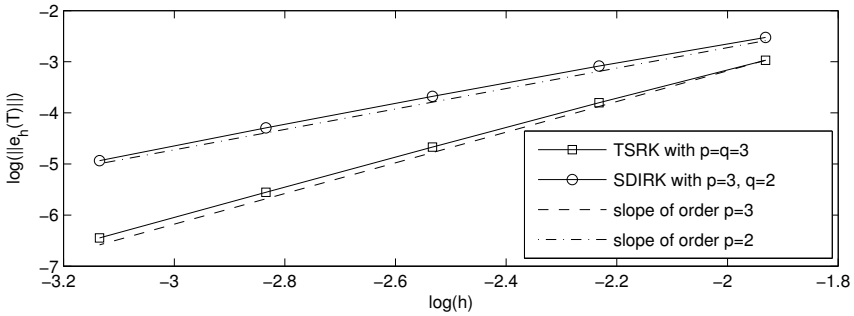

Figure 7.1: Numerical results for SDIRK formula with $p = 3$ and $q = 2$ and TSRK method with $p = q = 3$ on the problem (7.0.2) with $\epsilon = 10^{-4}$

Example 3 in Section 4.4. These results correspond to $T = 3/4$ and $\epsilon = 10^{-4}$. We can see again that in contrast to SDIRK formula TSRK method does not suffer from order reduction phenomenon.

7.2 Variable stepsize experiments based on two-step collocation methods

In this section we present some variable stepsize numerical experiments which aim to indicate the effectiveness of two-step collocation methods, especially in the implementation of stiff problems, and the reliability of the local error estimation. The implementation issues we have used in order to carry out the following experiments are the ones described in Chapter 6.

We apply the L-stable method (3.2.1), with $c = [1/2, 1]^T$ and

$$\begin{aligned} \varphi_0(s) &= -\frac{60}{19}s + \frac{45}{19}s^2, \quad \varphi_1(s) = 1 + \frac{60}{19}s - \frac{45}{19}s^2, \\ \chi_1(s) &= -\frac{8}{3}s + 2s^2, \qquad \chi_2(s) = -\frac{4}{3}s + s^2, \\ \psi_1(s) &= \frac{182}{57}s - \frac{36}{19}s^2, \quad \psi_2(s) = -\frac{77}{57}s + \frac{24}{19}s^2, \end{aligned} \tag{7.2.1}$$

described in Section 4.3.2, having narrowed contribution of the high order terms in the local discretization error. The method (7.2.1) of uniform effective order 2 has been applied to the Prothero-Robinson problem (7.0.1), and the van der Pol oscillator (7.0.2) in the interval $[0, 2]$. This equation constitutes a challenging problem for numerical methods: small oscillations are amplified, while large oscillations are damped (compare [119]).

The results of the implementation are reported in the figures below. Concerning the Prothero-Robinson problem (7.0.1), it is known that such problem is much more stiff when the stiffness parameter λ is negative and large in modulus. The experimental results reported in Figures 7.2, 7.3, 7.4 and 7.5 are referred to the cases $\lambda = -1e6$ and $\lambda = -1e10$: in correspondence of these values, the problem (7.0.1) is very stiff. In particular, Figures 7.3 and 7.5 show the reliability of the error estimate, also when the problem is very stiff. Moreover, as suggested by Figures 7.2 and 7.4, the stepsize pattern is very smooth, especially because of the high stability properties of the implemented method and in force of the used stepsize control strategy: this control also avoid useless stepsize refusions.

Figures 7.6 and 7.7 report the results concerning the numerical solution of the Van der Pol problem for $\varepsilon = 1e-6$ and $tol = 1e-4$. In correspondence of this value of the stiffness parameter ε, the problem is stiff. We observe

that also in this case the error estimate is absolutely reliable and the stepsize pattern is very smooth. Also the number of refused stepsize is very low: its percentage with respect to the total number of steps is lower then 1%: most of the refusions occur at the very first step point, because of the presence of an initial transient. We also observe that no hump phenomena (see [122]) occur: this is due to the L-stability of the method we have implemented.

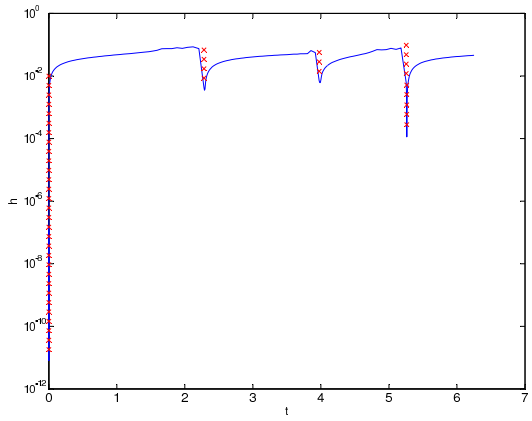

Figure 7.2: Stepsize pattern related to the solution of the Prothero-Robinson problem (7.0.1) with $tol = 1e-6$ and $\lambda = -1e6$, using the method (7.2.1). The crosses represent the refused stepsizes

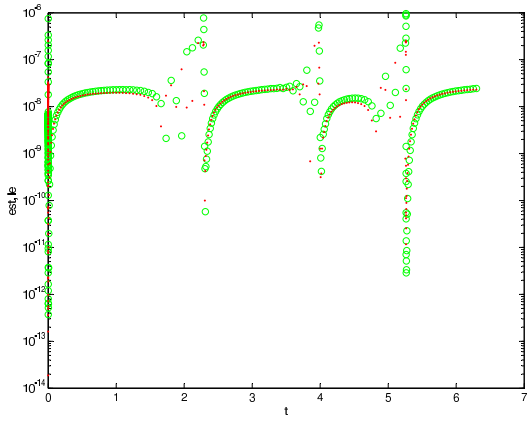

Figure 7.3: Comparison between the local error and its estimate for the solution of the Prothero-Robinson problem (7.0.1) with $tol = 1e-6$ and $\lambda = -1e6$, using the method (7.2.1). The circles represent the true local error in each step point, while the dots represent the corresponding estimation

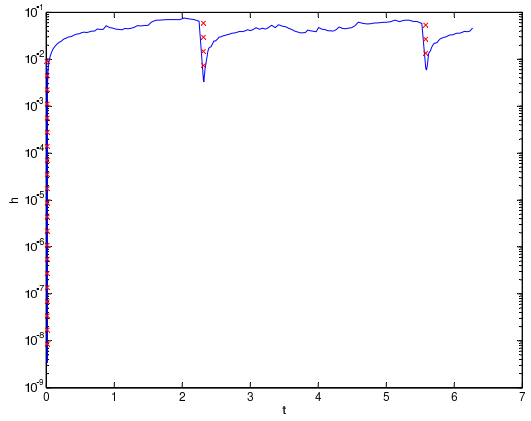

Figure 7.4: Stepsize pattern related to the solution of the Prothero-Robinson problem (7.0.1) with $tol = 1e-6$ and $\lambda = -1e10$, using the method (7.2.1)

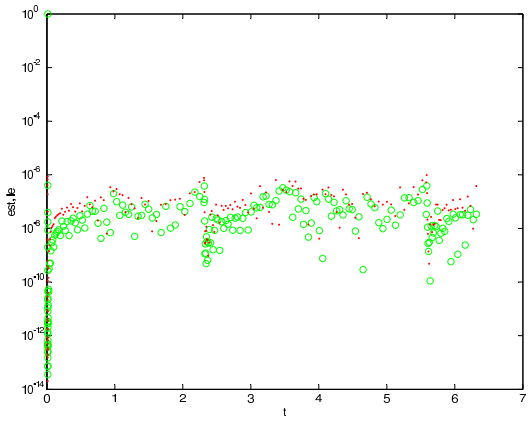

Figure 7.5: Comparison between the local error and its estimate for the solution of the Prothero-Robinson problem (7.0.1) with $tol = 1e-6$ and $\lambda = -1e10$, using the method (7.2.1)

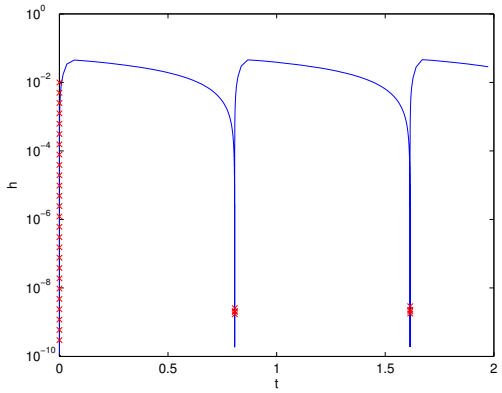

Figure 7.6: Stepsize pattern related to the solution of the Van der Pol oscillator (7.0.2) with $tol = 1e-4$ and $\varepsilon = 1e-6$, using the method (7.2.1)

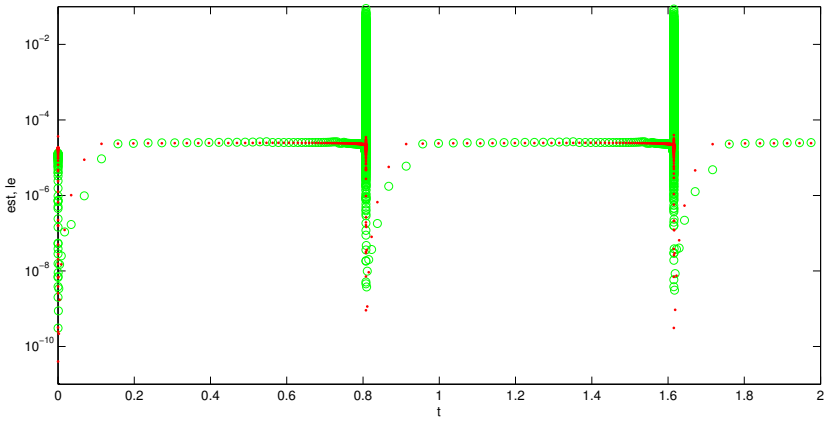

Figure 7.7: Comparison between the local error and its estimate for the solution of the Van der Pol oscillator (7.0.2) with $tol = 1e-4$ and $\varepsilon = 1e-6$, using the method (7.2.1)

Part II

Numerical solution of differential systems of the second order

Chapter 8

Collocation methods for second order ODEs of special type: state of the art

In this chapter we concentrate our attention on the historical evolution of the collocation technique for the numerical solution of initial value problems based on second order ordinary differential equations with periodic and oscillating solution

$$\begin{cases} y'(t) = f(t, y(t)), \quad t \in [t_0, T], \\ y'(t_0) = y'_0 \in \mathbb{R}^d, \\ y(t_0) = y_0, \end{cases} \tag{8.0.1}$$

where $f : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ is assumed to be a is sufficiently smooth function, in order to ensure the existence and the uniqueness of the solution. Although the problem (8.0.1) 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 seems more natural and efficient.

8.1 Direct and indirect collocation methods

In the context of collocation methods for second order equations, two

possibilities have been taken into account in the literature, i.e. methods based on *indirect* or *direct collocation* [204]. Indirect collocation methods are generated by applying a collocation based Runge-Kutta method to the first order representation of (8.0.1), which has doubled dimension. If

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

is the Butcher array of a collocation Runge-Kutta method, the tableau of the corresponding indirect collocation method is

$$\begin{array}{c|c} c & A^2 \\ \hline & A^T b \\ & b^T \end{array}$$

which results in a Runge-Kutta-Nyström method [119]. The theory of indirect collocation methods completely parallels the well–known theory of collocation methods for first order equations (see [204]) and, therefore, the properties of a collocation method are totally inherited by the corresponding indirect collocation method. Thus, the maximum attainable order is $2m$, where m is the number of stages, and it is achieved by Gauss-type methods, which are also A-stable, while L-stability is achieved by Radau IIA-type methods, of order $2m-1$.

In the case of direct collocation methods, the collocation polynomial is derived directly for the second order problem. Van der Houwen et al. in [204] studied the order, stage order of direct collocation methods and also provided their stability analysis, extending the results of Kramarz [146]. Concerning order and stage order, the following result holds (see [204]):

Theorem 8.1.1 *Direct and indirect collocation methods with the same collocation nodes have the same order. The stage order of direct collocation methods is one higher whenever*

$$\int_0^1 \prod_{i=1}^{m} (s - c_i) ds = 0. \qquad \square$$

Therefore, while indirect and direct collocation methods have the same order, their stage order is different and, in particular, direct methods have higher stage order. However, they are not competitive in terms of stability. Van der Houwen et al. in [204] clearly state that "From a practical point

of view, direct collocation methods based on Gauss, Radau and Lobatto collocation points are of limited value, because the rather small stability or periodicity boundaries make them unsuitable for stiff problems. The A-stable indirect analogues are clearly more suitable for integrating stiff problems".

Moreover, Coleman [61] proved that no P-stable one step symmetric collocation methods exist. P-stability (see Lambert-Watson paper [148]) is a very relevant property for the numerical treatment of a second order system whose theoretical solution is periodic with a moderate frequency and a high frequency oscillation of small amplitude superimposed. This phenomenon is known in literature as *periodic stiffness* [178], which can be reasonably faced using P-stable methods, exactly as A-stable methods are suitable for stiff problems. In other terms, P-stability ensures that the choice of the stepsize is independent from the values of the frequencies, but it only depends on the desired accuracy [65, 173].

In [148], the authors proved that P-stable linear multistep methods

$$\sum_{j=0}^{k} \alpha_j y_{n+j} = h^2 \sum_{j=0}^{k} \beta_j f(t_{n+j}, y_{n+j}).$$

can achieve maximum order 2. In the context of Runge–Kutta–Nyström methods

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

many A-stable and P-stable methods exist, but the ones falling in the subclass of collocation methods, whose coefficients (see [119]) are of the form

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

have only bounded stability intervals and are not P-stable [173].

8.2 Two-step Runge-Kutta-Nyström methods

We have observed in the previous paragraph that P-stability is a desirable property that only few methods in the context of linear multistep methods and Runge-Kutta-Nyström methods possess. In order to create a good balance between high order and strong stability properties, further steps in the literature have been devoted to the development of multistep Runge-Kutta-Nystrom methods for second order problems. Much of this work has been done by Paternoster (see [173, 174, 175, 176, 159, 177]). In particular, the author proved that no P-stable methods can be found in the class of indirect collocation TSRK methods, while it was possible to find P-stable methods in the context of *two-step Runge-Kutta-Nyström methods*

$$
\begin{aligned}
Y_j^{[n-1]} &= y_{n-1} + c_j h y'_{n-1} + h^2 \sum_{k=1}^{m} a_{jk} f(t_{n-1} + c_k h, Y_k^{[n-1]}), \quad j = 1, 2, ..., m, \\
Y_j^{[n]} &= y_n + c_j h y'_n + h^2 \sum_{k=1}^{m} a_{jk} f(t_n + c_k h, Y_k^{[n]}), \quad j = 1, 2, ..., m, \\
y_{n+1} &= (1-\theta) y_n + \theta y_{n-1} + h \sum_{j=1}^{m} (v_j y'_{n-1} + w_j y'_n) \\
&+ h^2 \sum_{j=1}^{m} \bar{v}_j f(t_{n-1} + c_j h, Y_j^{[n-1]}) + \bar{w}_j f(t_n + c_j h, Y_j^{[n]}), \\
y'_{n+1} &= (1-\theta) y'_n + \theta y'_{n-1} + h \sum_{j=1}^{m} (v_j f(t_{n-1} + c_j h, Y_j^{[n-1]}) \\
&+ w_j f(t_n + c_j h, Y_j^{[n]})),
\end{aligned}
$$

which represent the extension to second order problems of the two-step Runge-Kutta methods introduced in [139] for first order problems.

8.3 Mixed collocation methods

The development of classical collocation methods (i.e. methods based on algebraic polynomials), even if it is not the most suitable choice for second order problems that do not possess solutions with polynomial behaviour, is the first necessary step in order to construct collocation methods whose collocation function is expressed as linear combination of different functions, e.g. trigonometric polynomials, mixed or exponential basis (see, for instance, [64, 135]), which can better follow the qualitative behaviour of the solution. It is indeed more realistic to choose basis functions which are not polynomials.

Many authors have considered in literature different functional basis, instead of the polynomial one, e.g. [25, 62, 65, 91, 100, 110, 135, 145, 170, 172, 174, 176, 189]. In particular we mention here the work by Coleman and Duxbury [64], where the authors introduced mixed collocation methods applied to the Runge-Kutta-Nyström scheme, where the collocation function is expressed as linear combination of trigonometric functions and powers, in order to provide better approximations for oscillatory solutions. The methods are derived in order to exactly integrate the harmonic oscillator

$$y'' = -k^2 y,$$

where k is a constant, a feature which is not achievable by algebraic polynomial collocation. The term *mixed interpolation* appeared for the first time in [100] to describe interpolation by a linear combination of a sine and cosine of a given frequency, and powers of the relevant variable, and later used by Brunner et al. in [25] in the context of Volterra integral equations. The solution on the generic integration interval $[t_n, t_{n+1}]$ is approximated by the collocating function

$$u(t_n + sh) = a\cos\theta s + b\sin\theta s + \sum_{i=0}^{k} \Gamma_i s^i, \tag{8.3.2}$$

which satisfies the following collocation and interpolation conditions

$$\begin{aligned} u(t_n) &= y_n, \qquad u'(t_n) = y'_n, \\ u''(t_n + c_j h) &= f(t_n + c_j h, u(t_n + c_j h)), \quad j = 1, \ldots, m. \end{aligned}$$

Integrating (8.3.2) twice, we obtain the Runge-Kutta-Nystrom formulation of the methods, i.e.

$$\begin{aligned} u'(t_n + sh) &= y'_n + h\sum_{i=1}^{m} \alpha_i(s) f(t_n + c_i h, u(t_n + c_i h)), \\ u(t_n + sh) &= y_n + shy'_n + h^2 \sum_{i=1}^{m} \beta_i(s) f(t_n + c_i h, u(t_n + c_i h)), \end{aligned}$$

where

$$\alpha_i(s) = \int_0^s L_i(\tau) d\tau, \qquad \beta_i(s) = \int_0^s (s - \tau) L_i(\tau) d\tau.$$

Outside collocation, many authors derived methods having frequency dependent parameters (see, for instance, [135, 145, 171, 189] and references therein contained). The linear stability analysis of these methods is carried out in [65]. In [91] also a method with parameters depending on two frequency is presented, and the modification in the stability analysis is performed, leading to a three dimensional region.

Chapter 9

Collocation based two-step hybrid methods

9.1 Collocation based two-step hybrid methods

In the numerical integration of second order ODEs (8.0.1) through collocation, many possibilities can be taken into account: for example, Runge–Kutta–Nyström methods provide an approximation to the solution and its first derivative at each step point. However, as Henrici observed in [123], "If one is not particularly interested in the values of the first derivatives, it seems unnatural to introduce them artificially". For this reason, other types of methods have been taken into account in the literature, i.e. methods which provide an approximation to the solution without computing any approximation to the first derivative: these formulae are denoted in literature as *hybrid methods*. Coleman introduced in [63] the following class of two-step hybrid methods for second order ODEs:

$$Y_i^{[n]} = u_i y_{n-1} + (1-u_i) y_n + h^2 \sum_{j=1}^{m} a_{ij} f(t_n + c_j h, Y_j^{[n]}), \qquad (9.1.1)$$

$$i = 1, \ldots, m,$$

$$y_{n+1} = \theta y_{n-1} + (1-\theta) y_n + h^2 \sum_{j=1}^{m} w_j f(t_n + c_j h, Y_j^{[n]}), \qquad (9.1.2)$$

which can also be represented through the Butcher array

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

with $c = (c_1, c_2, ..., c_m)^T$, $A = (a_{ij})_{i,j=1}^m$, $b = (b_1, b_2, ..., b_m)^T$, where m is the number of stages. The interest in this class of methods, as also pointed out by Coleman in [63], 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 (8.0.1), towards a class of General Linear Methods [42, 97, 138] for this problem. This class of methods has been further investigated in [59, 91, 108, 205, 206]. We report in this chapter the results contained in the paper [90], concerning the construction and the analysis of collocation based two-step hybrid methods of the type (9.1.1)-(9.1.2). In the next sections we will define the collocation polynomial $P(t_n + sh)$ associated to the methods (9.1.1)-(9.1.2), discuss its construction and handle the study of the order and the stability properties of the resulting methods. It is important to observe that algebraic collocation, even if not particularly suited to reproduce the qualitative behaviour of problems whose solutions is particularly oscillating, is the first necessary step in order to construct collocation methods whose collocation function is expressed as linear combination of different functions, e.g. trigonometric polynomials, mixed or exponential basis (see, for instance, [64, 135]).

9.2 Derivation of methods

We now discuss the modification of the technique introduced by Hairer and Wanner in [122] for first order ODEs, in order to derive two–step collocation methods of the form (9.1.1), (9.1.2). We require that the corresponding collocation polynomial assumes the form

$$P(t_n + sh) = \varphi_1(s)y_{n-1} + \varphi_2(s)y_n + h^2 \sum_{j=1}^{m} \chi_j(s)P''(t_n + c_j h), \qquad (9.2.1)$$

where $s \in [0, 1]$, which is a linear combination of the basis polynomials $\{\varphi_1(s), \varphi_2(s), \chi_j(s), j = 1, 2, \ldots, m\}$ of degree at most equal to $m + 1$. The

collocation polynomial has to satisfy the following set of $m+2$ interpolation and collocation conditions

$$\begin{aligned} P(t_{n-1}) &= y_{n-1}, \qquad P(t_n) = y_n, && (9.2.2)\\ P''(t_n + c_j h) &= f(t_n + c_j h, P(t_n + c_j h)), \qquad j = 1, \ldots, m. && (9.2.3) \end{aligned}$$

which allow us to derive a polynomial of degree at most $m+1$. The counterpart of these conditions for the basis functions $\{\varphi_1(s), \varphi_2(s), \chi_j(s), j = 1, 2, \ldots, m\}$ can then be expressed as follows

$$\begin{array}{lll} \varphi_1(-1) = 1, & \varphi_2(-1) = 0, & \chi_j(-1) = 0, \\ \varphi_1(0) = 0, & \varphi_2(0) = 1, & \chi_j(0) = 0, \\ \varphi_1''(c_i) = 0, & \varphi_2''(c_i) = 0, & \chi_j''(c_i) = \delta_{ij}, \end{array} \qquad (9.2.4)$$

for $i, j = 1, \ldots, m$. Such conditions are useful to determine the coefficients of the unknown basis functions, which are given by solving $m+2$ linear systems having the following coefficient matrix

$$H = \begin{pmatrix} 1 & -1 & 1 & \ldots & (-1)^i & \ldots & (-1)^{m+1} \\ 1 & 0 & 0 & \ldots & 0 & \ldots & 0 \\ 0 & 0 & 2 & \ldots & i(i-1)c_1^{i-2} & \ldots & (m+1)mc_1^{m-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 2 & \ldots & i(i-1)c_m^{i-2} & \ldots & (m+1)mc_m^{m-1} \end{pmatrix},$$

which is a nonsingular matrix (apart for some exceptional values of the collocation abscissa) because of Vandermonde type (see [158]). After computing the basis functions, the class of methods takes the following form

$$\begin{aligned} Y_i^{[n]} &= \varphi_1(c_i) y_{n-1} + \varphi_2(c_i) y_n + h^2 \sum_{j=1}^{m} \chi_j(c_i) P''(t_n + c_j h), && (9.2.5)\\ y_{n+1} &= \varphi_1(1) y_{n-1} + \varphi_2(1) y_n + h^2 \sum_{j=1}^{m} \chi_j(1) P''(t_n + c_j h). && (9.2.6) \end{aligned}$$

9.3 Order conditions

We now derive the set of continuous order conditions, by considering $P(t_n + sh)$ as an uniform approximation of $y(t_n + sh)$ on the whole integration interval.

Theorem 9.3.1 *Assume that the function f in (8.0.1) is sufficiently smooth. Then the method (9.2.5), (9.2.6) has uniform order p if the following conditions are satisfied:*

$$1 - \varphi_1(s) - \varphi_2(s) = 0, \tag{9.3.1}$$

$$s + \varphi_1(s) = 0, \tag{9.3.2}$$

$$\frac{s^k}{k!} - \varphi_1(s)\frac{(-1)^k}{k!} - \sum_{j=1}^{m} \chi_j(s)\frac{c_j^{k-2}}{(k-2)!} = 0, \tag{9.3.3}$$

with $k = 2, \ldots, p$ and $t \in [0, 1]$.

Proof: We consider the local discretization error

$$\xi(t_n+sh) = y(t_n+sh)-\varphi_1(s)y(t_n-h)-\varphi_2(s)y(t_n)-h^2\sum_{j=1}^{m}\chi_j(s)y''(t_n+c_jh), \tag{9.3.4}$$

and expand $y(t_n+sh)$, $y(t_n-h)$ and $y''(t_n+c_jh)$ in Taylor series around the point t_n, obtaining

$$\begin{aligned}
\xi(t_n+sh) &= y(t_n) + shy'(t_n) + \cdots + \frac{(sh)^p}{p!}y^{(p)}(t_n) + \\
&- \varphi_1(s)\left(y(t_n) - hy'(t_n) + \cdots + \frac{(-1)^p h^p}{p!}y^{(p)}(t_n)\right) - \varphi_2(s)y(t_n) \\
&- h^2\sum_{j=1}^{m}\chi_j(s)\left(y''(t_n) + c_jhy'''(t_n) + \cdots + \frac{(c_jh)^{p-2}}{(p-2)!}y^{(p)}(t_n)\right) + O(h^{p+1}).
\end{aligned}$$

We then compare the coefficients of the same power of h, achieving the thesis. □

Theorem 9.3.1 allows us to prove that every two–step collocation method of the type (9.2.5), (9.2.6) has order $p = m$ on the whole integration interval, and this is result is in keeping with [63]. In the context of GLMs, the condition (9.3.1) is the *preconsistency condition*, while the (9.3.2) is the *consistency condition*. In order to be the method preconsistent and consistent, it must be $\varphi_1(s) = -s$ and $\varphi_2(s) = 1+s$, i.e. the methods (9.1.1), (9.1.2) exactly fall in the class of Coleman hybrid methods, since $\theta = -1$ and $u_i = c_i$, $i = 1, ..., m$.

9.4 Linear Stability Analysis

We now examine the linear stability properties of the resulting methods by using the procedure indicated in [178, 203, 204]. We apply the class of methods (9.1.1), (9.1.2), to the test equation

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

obtaining

$$Y_i^{[n]} = u_i y_{n-1} + (1-u_i) y_n - z^2 \sum_{j=1}^{m} a_{ij} Y_j^{[n]}, \tag{9.4.1}$$

$$y_{n+1} = \theta y_{n-1} + (1-\theta) y_n - z^2 \sum_{j=1}^{m} w_j Y_j^{[n]}, \tag{9.4.2}$$

where $z^2 = \omega^2 h^2$. In matrix notation,

$$Y^{[n]} = u y_{n-1} + \tilde{u} y_n - z^2 A Y^{[n]}, \tag{9.4.3}$$

$$y_{n+1} = \theta y_{n-1} + (1-\theta) y_n - z^2 w^T Y^{[n]}, \tag{9.4.4}$$

where $Y^{[n]} = (Y_i^{[n]})_{i=1}^m$, $u = (u_i)_{i=1}^m$, $\tilde{u} = (1-u_i)_{i=1}^m$, $w = (w_i)_{i=1}^m$, $A = (a_{ij})_{i,j=1}^m$. The following expression for the stage values holds:

$$Y^{[n]} = Q[u y_{n-1} + \tilde{u} y_n], \tag{9.4.5}$$

where $Q = [I + z^2 A]^{-1}$ and I is the identity matrix of dimension m. If we substitute this expression in (9.4.4), the following recurrence relation arises:

$$y_{n+1} = [\theta - z^2 w^T Q u] y_{n-1} + [1 - \theta - z^2 w^T Q \tilde{u}] y_n \tag{9.4.6}$$

or, equivalently,

$$\begin{bmatrix} y_{n+1} \\ y_n \end{bmatrix} = \begin{bmatrix} M_{11}(z^2) & M_{12}(z^2) \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y_n \\ y_{n-1} \end{bmatrix}, \tag{9.4.7}$$

where

$$M_{11}(z^2) = 1 - \theta - z^2 w^T Q \tilde{u}, \qquad M_{12}(z^2) = \theta - z^2 w^T Q u.$$

The characteristic polynomial, associated to the recursion (9.4.7), is the *stability polynomial*, while the *stability* or *amplification matrix* [203, 204] associated to the methods takes the form

$$M(z^2) = \begin{bmatrix} M_{11}(z^2) & M_{12}(z^2) \\ 1 & 0 \end{bmatrix}. \tag{9.4.8}$$

Let us denote the spectral radius of $(M(z^2))$ with $\rho(M(z^2))$. From [203, 204], we consider the following definitions.

Definition 9.4.1 *$(0, \beta^2)$ is a stability interval for the method (9.1.1), (9.1.2) if, for $z^2 \in (0, \beta^2)$, it is*

$$\rho(M(z^2)) < 1. \tag{9.4.9}$$

The above condition is equivalent to the fact that the roots of the stability polynomial are in modulus less than 1, $\forall z^2 \in (0, \beta^2)$. Setting $S(z^2) = \mathrm{tr}(M^2(z^2))$ and $P(z^2) = \det(M^2(z^2))$, the condition (9.4.9) is equivalent to

$$P(\nu^2) < 1, \quad |S(\nu^2)| < P(\nu^2) + 1, \quad \nu \in (0, \beta^2) \tag{9.4.10}$$

Definition 9.4.2 *The method (9.1.1), (9.1.2) is A-stable if $(0, \beta^2) = (0, +\infty)$.*

In order to reach A-stability, it must be $\rho(M(\nu^2)) < 1$, for any value of ν^2, where $\rho(M(\nu^2))$ is the spectral radius of the stability matrix, i.e. both the eigenvalues λ_1, λ_2 of $M(\nu^2)$ must satisfy the condition $|\lambda_1| < 1$, $|\lambda_2| < 1$. For $m = 1$, through an analytical study of the stability matrix (9.4.8), it is possible to prove the following result which characterizes A-stable methods.

Theorem 9.4.1 (One-stage A-stable methods) *For $m = 1$, the method (9.2.5), (9.2.6) is A-stable if and only if $c \in (\frac{1}{\sqrt{2}}, 1]$*

If the eigenvalues of the stability matrix (9.4.8) (or in equivalent way, 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 9.4.3 *$(0, H_0^2)$ is a periodicity interval if, for $z^2 \in (0, H_0^2)$ the roots $r_1(z^2)$, $r_2(z^2)$ of the stability polynomial $\pi(\lambda) = \det[M(z^2) - \lambda I]$ are complex conjugate and $|r_1(z^2)| = |r_2(z^2)| = 1$.*

Definition 9.4.4 *The method (9.1.1), (9.1.2) is P-stable if its periodicity interval is $(0, +\infty)$.*

Through a numerical search, it is possible to find nonempty periodicity intervals. For instance, in the case $m = 1$, for any $c \in [0, \frac{1}{50})$, the periodicity interval of the resulting methods is $[0, 4]$.

9.5 Numerical Experiments

We consider the following initial value problem

$$\begin{cases} y''(t) = \begin{pmatrix} \mu - 2 & 2\mu - 2 \\ 1 - \mu & 1 - 2\mu \end{pmatrix} y(t), \\ y(0) = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \\ y'(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \end{cases} \tag{9.5.1}$$

with $\mu \in \mathbb{R}$ and $t \in [0, 20\pi]$, which provides an useful illustration of the need to know the stability properties of the methods when we aim to integrate stiff systems. The exact solution is $y_1(t) = 2\cos t,\ y_2(t) = -\cos t$, i.e. it is independent on μ. In particular, when $\mu = 2500$, the problem (10.2.15) is the Kramarz system [146], which is often used in numerical experiments on stiffness in second order ODEs.

The eigenvalues of the coefficient matrix of the system (10.2.15) are -1 and $-\mu$: as a consequence, the analytical solution of the system exhibits the two frequencies 1 and $\sqrt{\mu}$, but the initial conditions eliminate the high frequency component, which corresponds to $\sqrt{\mu}$ when $\mu >> 1$. Notwithstanding this, its presence in the general solution of the system dictates strong restrictions on the choice of the stepsize, so that the system exhibits the phenomenon of *periodic stiffness* [178].

In the usage of a numerical method with constant parameters and having a limited interval of stability $(0, H)$, it is known that the method is stable when $0 < h < \sqrt{H}/\sqrt{\mu}$. As μ increases, the value of h has to be chosen smaller and smaller, to make the computation stable; in this sense the parameter μ is a measure of the stiffness of the system. It is obvious that, for an A-stable method, the choice of the stepsize is governed only by accuracy demands.

Hybrid method, m=1, c=3/4, p=1			
h	fe	cd	ge
0.01	9215	0.6046	0.2485
0.005	18008	0.8930	0.1279
0.0025	33923	1.1877	0.0648
0.00125	61937	1.4856	0.0326
0.000625	109219	1.7850	0.0164
0.0003125	201060	2.0853	0.0082

Radau IIA method, m=1, p=1			
h	fe	cd	ge
0.01	12361	0.4881	0.3250
0.005	24294	0.7724	0.1688
0.0025	46940	1.0650	0.0861
0.00125	88872	1.3617	0.0435
0.000625	165310	1.6606	0.0218
0.0003125	302663	1.9606	0.0109

In the above tables, fe is the counter of function evaluations, cd is the number of correct digits and ge is the norm of the global error in the last step point. The numerical results reveal that two–step hybrid collocation methods (9.2.5), (9.2.6) essentially show the same behaviour of the indirect collocation Radau IIA method, but with a lower computational cost in the fixed stepsize implementation.

Chapter 10

Efficient solution of oscillatory problems: special purpose two-step hybrid methods

Classical numerical methods for ODEs (8.0.1) may not be well-suited to follow a prominent periodic or oscillatory behaviour of the solutions 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 (within the round-off error) problems of type (8.0.1) whose solution can be expressed as linear combination of functions other than polynomials: this is the spirit of the *exponential fitting* technique (compare the monography [135] and references therein contained), where the adapted numerical method is developed in order to be exact on problems whose solution is linear combination of the following basis functions

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

with K and P integer numbers, and of *trigonometrical fitting*, where the reference function basis is

$$\{1, t, \ldots, t^M, t^j \cos(\pm\omega t), \ldots, t^\ell \sin(\pm\omega t),\ j = 0, 1, \ldots, Q,\ \ell = 0, 1, \ldots, R\},$$

where M, Q and R are integer numbers. The coefficients of the adapted methods depend on the values of the parameters appearing in the solution, e.g. the frequency of the oscillations, and are supposed to be estimated in advance.

In the context of linear multistep methods for second order ODEs, Gautschi [110], Stiefel-Bettis [195] considered trigonometric functions depending on one or more frequencies, while Lyche [157] derived methods exactly integrating initial value problems of order r whose solution can be expressed as linear combination of powers and exponentials; Raptis-Allison [181] and Ixaru-Rizea [133] derived special purpose linear multistep methods for the numerical treatment of the radial Schrödinger equation $y'' = (V(t) - E)y$, 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 [53], Franco [107], Simos [145, 190] and Vanden Berghe [201], while their trigonometrically-fitted version has been developed by Paternoster in [170]; mixed-collocation based Runge–Kutta–Nyström methods have been introduced by Coleman and Duxbury in [64]. Recent adaptations of the Numerov method have been provided in [106, 128, 202]. For a more extensive bibliography see [135] and references within.

The purpose of this chapter is the derivation and the analysis of the adapted version of two-step hybrid methods (9.1.1)-(9.1.2), suitable to integrate problems (8.0.1) whose solution shows an oscillatory or exponential behaviour, in such a way that it exactly integrates linear combinations of power, exponentials and/or trigonometric functions depending on the values of one and two parameters, which we suppose can be estimated in advance. Frequency-dependent methods within the class (9.1.1)-(9.1.2) have already been considered in [205], where the methods are modified to produced phase-fitted and amplification-fitted methods.

The results contained in this chapter deal with the papers [91, 87], where we have respectively introduced the trigonometrically and exponentially fitted versions of the methods (9.1.1)-(9.1.2), depending on one and two parameters related to the behaviour of the solution.

10.1 The role of linear operators: two-step hybrid methods as $\mathcal{A}$-methods

An important role in the derivation of adapted formulae is played by some linear operators associated to the linearized version of the methods we are analyzing. In order to linearize the methods (9.1.1)-(9.1.2), we follow the approach introduced by Albrecht (cfr. [5, 6, 7, 8, 147]), mainly consisting in rewriting the class of methods in the form of the so-called $\mathcal{A}$-methods. We now describe how to apply this approach in the case of two-step hybrid methods (9.1.1)-(9.1.2) to approximate the solution of the problem (8.0.1) which is assumed, for simplicity, to be scalar.

We first define the following vectors in $\mathbb{R}^{s+2}$

$$\begin{aligned} Y_{n+1} &= [Y_1^{[n]}, ..., Y_s^{[n]}, y_n, y_{n+1}]^T, \\ F(t_n, Y_{n+1}; h) &= [f(t_n + c_1 h, Y_1), ..., f(t_n + c_s h, Y_s), f(t_n, y_n), f(t_n, y_{n+1})]^T. \end{aligned}$$

In this way, a two-step hybrid method (9.1.1)-(9.1.2) can be expressed in the following form of $\mathcal{A}$-*method*

$$Y_{n+1} = \mathcal{A} Y_n + h^2 \mathcal{B} F(t_n, Y_{n+1}; h), \tag{10.1.1}$$

with

$$\mathcal{A} = \begin{bmatrix} 0 & -c & \mathbf{e}+c \\ 0 & -1 & 2 \\ 0 & 0 & 1 \end{bmatrix}, \ \mathcal{B} = \begin{bmatrix} A & 0 & 0 \\ b^T & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{(s+2)\times(s+2)}, \tag{10.1.2}$$

where $\mathbf{e}$ is the unit vector of $\mathbb{R}^s$. This representation is very useful, because it constitutes a $s+2$ linear stages representation, in the sense that each of the s internal stages and the external stages are linear, thus we can look at them as a generalized linear multistep formula on a nonequidistant grid. For this reason, we can consider the following $s+1$ linear operators

$$\begin{aligned} \mathcal{L}_i[z(t); h] &= z(t + c_i h) - (1 + c_i) z(t) + c_i z(t - h) \\ &- h^2 \sum_{j=1}^{s} a_{ij} z''(t + c_j h), \quad i = 1, ..., s \end{aligned} \tag{10.1.3}$$

$$\hat{L}[z(t); h] = z(t + h) - 2z(t) + z(t - h) - h^2 \sum_{i=1}^{s} b_i z''(t + c_i h), \tag{10.1.4}$$

where $z(t)$ is a smooth enough function. Expanding in power series of h around t we obtain

$$\mathcal{L}_i[z(t);h] = C_{i2}h^2 z^{(2)}(t) + C_{i3}h^3 z^{(3)}(t) + ..., \quad i = 1, ..., s,$$
$$\hat{L}[z(t);h] = \hat{C}_2 h^2 z^{(2)}(t) + \hat{C}_3 h^3 z^{(3)}(t) + ...,$$

where

$$C_{iq} = \frac{c_i^q}{q!} + \frac{(-1)^q}{q!} - \frac{1}{(q-2)!}\sum_{j=1}^{s} a_{ij}c_j^{q-2}, \quad i = 1, ..., s,\ q = 2, 3, ...,$$

$$\hat{C}_q = \frac{1}{q!} + \frac{(-1)^q}{q!} - \frac{1}{(q-2)!}\sum_{j=1}^{s} b_j c_j^{q-2},\ q = 2, 3, ...$$

As a consequence, the following definition arises.

Definition 10.1.1 *The i^{th} internal stage (9.1.1) of a two-step hybrid method (9.1.1)-(9.1.2) has order p_i if*

$$C_{i2} = 0,\ C_{i3} = 0,\ C_{ip_i+1} = 0, \quad C_{ip_i+2} \neq 0, \tag{10.1.5}$$

while the external stage (9.1.2) has order p if

$$\hat{C}_2 = 0,\ \hat{C}_3 = 0,\ \hat{C}_{p+1} = 0, \quad \hat{C}_{p+2} \neq 0. \tag{10.1.6}$$

We know that necessary condition for a two-step hybrid method (9.1.1)-(9.1.2) to have order p is that the external stage must have order p, i.e.

$$b^T c^{q-2} = \frac{1 + (-1)^q}{q(q-1)}, \quad q = 2, 3, ..., p+1, \tag{10.1.7}$$

where the vector power is componentwise. In order to look for conditions that are also sufficient, in line to Albrecht's approach, we need to look at the global error. We omit the details achieving order conditions, because they are outside the original aim of our analysis, and which can be found in [63]. Table 10.1 shows the set of order conditions up to 4.

In order to derive s-stage methods of type (9.1.1)-(9.1.2) with constant coefficients and having order q, we annihilate the linear operators (10.1.3)-(10.1.4) on the functional basis

$$\{1, t, t^2, ..., t^q\} \tag{10.1.8}$$

Order	Order conditions
1	$\sum_i b_i = 1$
2	$\sum_i b_i c_i = 0$
3	$\sum_i b_i c_i^2 = \frac{1}{6}$ $\sum_i \sum_j b_i a_{ij} = \frac{1}{12}$
4	$\sum_i b_i c_i^3 = 0$ $\sum_i \sum_j b_i c_i a_{ij} = \frac{1}{12}$ $\sum_i \sum_j b_i a_{ij} c_j = 0$

Table 10.1: Order conditions for two-step hybrid methods (9.1.1)-(9.1.2) of order up to 4.

with $q = s + 1$. It trivially occurs that

$$\begin{aligned} \mathcal{L}_i[1;h] &= \hat{L}[1;h] = 0, \quad i = 1,\ 2, ...,\ s, \\ \mathcal{L}_i[t;h] &= \hat{L}[t;h] = 0, \quad i = 1,\ 2, ...,\ s, \end{aligned}$$

while, for $2 \leq k \leq q$, it is

$$\begin{aligned} \mathcal{L}_j[t^k;h] &= 0,\ i = 1,\ 2, ...,\ s \quad \Leftrightarrow \quad \frac{c_i^k + (-1)^k c_i}{k(k-1)} = \sum_{j=1}^{s} a_{ij} c_j^{k-2} \\ \hat{L}[t^k;h] &= 0 \quad \Leftrightarrow \quad \frac{1 + (-1)^k}{k(k-1)} = \sum_{i=1}^{s} b_i c_i^{k-2}. \end{aligned}$$

Therefore, we obtain the set of conditions

$$\begin{cases} \dfrac{c_i^k + (-1)^k c_i}{k(k-1)} = \displaystyle\sum_{j=1}^{s} a_{ij} c_j^{k-2} \\ \dfrac{1 + (-1)^k}{k(k-1)} = \displaystyle\sum_{j=1}^{s} b_j c_j^{k-2}, \end{cases} \tag{10.1.9}$$

for $i = 1,\ 2, ...,\ s$ and $2 \leq k \leq q$, which is a system of $s(s+1)$ equations in the unknowns a_{ij}, b_i, for $i,\ j = 1,\ 2, ...,\ s$.

It is easy to verify that the methods obtained by solving the order conditions (10.1.9) are equal to the methods described in [90], which are based on collocation through algebraic polynomials, and have been derived by extending the multistep collocation technique described in [122]. Therefore conditions (10.1.9) are the order conditions for collocation methods within class (9.1.1)-(9.1.2).

10.2 Trigonometrically-fitted two-step hybrid methods

This section is devoted to the derivation of trigonometrically-fitted two-step hybrid methods (9.1.1)-(9.1.2), aiming for the efficient integration of second order ODEs (8.0.1) whose solution exhibits a prominent oscillatory behaviour, under the assumption that it depends on the values of one or two frequencies, which we suppose can be estimated in advance.

In particular, we require that both the internal and external stages of the resulting methods exactly integrate linear combinations of the following basis functions:

$$\{1, t, \ldots, t^q, \cos(\omega_i t), \sin(\omega_i t), \ldots\} \tag{10.2.1}$$

depending on the frequencies ω_i, for $i = 1$ or $i = 2$. The idea to use a mixed basis containing both powers and trigonometric functions follows the idea of using mixed interpolation of type

$$a\cos(\omega t) + b\sin(\omega t) + \sum_{i=0}^{s-1} c_i t^i, \tag{10.2.2}$$

introduced in [100]. The abscissa vector $\mathbf{c}$ is considered to be free: each value of c_i, $i = 1, 2, ..., s$, can be chosen in order to improve the stability properties of the methods or, imposing a special set of constraints, in order to achieve superconvergence.

10.2.1 Construction of methods with coefficients depending on one frequency

In order to derive numerical methods for second order ODEs whose solution depends on the frequency ω, a priori known, we consider the function basis

$$\{1, t, \ldots, t^q, \cos(\omega t), \sin(\omega t),\ q = 1, 2, \ldots\}.$$

We first consider the derivation of two-stage methods, corresponding to the basis

$$\{1, \cos(\omega t), \sin(\omega t)\},$$

and we impose that the numerical method exactly integrates second order ODEs whose solution is a linear combination of the basis functions. In this way we obtain the class of trigonometrically fitted two-step hybrid methods. As it automatically happens that

$$\mathcal{L}_j[1;h] = \hat{L}[1;h] = 0,\ j = 1, 2,$$

we need to impose the following set of conditions

$$\begin{array}{llll} \mathcal{L}_j[\cos\omega t;h] = 0, & j = 1,2, & \hat{L}[\cos\omega t;h] = 0, \\ \mathcal{L}_j[\sin\omega t;h] = 0, & j = 1,2, & \hat{L}[\sin\omega t;h] = 0, \end{array} \tag{10.2.3}$$

which constitutes a 6×6 linear system in the unknowns $a_{11},\ a_{12},\ a_{21},\ a_{22},\ b_1,\ b_2$.

In order to construct methods with 3 or more stages, we impose that the numerical method exactly integrates second order ODEs whose solution is a linear combination of the basis function

$$\{1, t, t^2, ...t^{s-1}, \cos(\omega t), \sin(\omega t)\}$$

depending on the frequency ω. In this case we obtain a class of *mixed-trigonometrically fitted two-step hybrid methods*. As also $\mathcal{L}_j[t;h] = \hat{L}[t;h] = 0,\ j = 1, ..., s$, it is sufficient to impose that

$$\begin{array}{llll} \mathcal{L}_j[t^q;h] = 0, & \hat{L}[t^q;h] = 0, & j = 1,2,...,s, & q = 2,3...,s-1, \\ \mathcal{L}_j[\cos\omega t;h] = 0, & \hat{L}[\cos\omega t;h] = 0, & j = 1,2,...,s, & \\ \mathcal{L}_j[\sin\omega t;h] = 0, & \hat{L}[\sin\omega t;h] = 0, & j = 1,2,...,s. & \end{array}$$

It arises a system of $s(s+1)$ conditions in the unknowns a_{ij}, b_i, $i, j = 1, 2, ..., s$.

This system is equivalent to the following set of conditions:

$$\begin{aligned}
\sum_{i=1}^{s} b_i c_i^{q-2} &= \frac{1+(-1)^q}{q(q-1)}, \quad q = 2, \ldots s, \\
\sum_{j=1}^{s} a_{ij} c_j^{q-2} &= c_i^q \frac{1+(-1)^q c_i}{q(q-1)}, \quad i = 1, 2, \ldots s, \quad q = 2, \ldots s, \\
\sum_{i=1}^{s} b_i \cos(c_i\theta) &= 2\frac{(1-\cos\theta)}{\theta^2}, \\
\sum_{j=1}^{s} a_{ij} \cos(c_j\theta) &= -\frac{\cos(c_i\theta)+1+c_i+c_i\cos\theta}{\theta^2} \quad i = 1, 2, \ldots s, \\
\sum_{i=1}^{s} b_i \sin(c_i\theta) &= 0, \\
\sum_{j=1}^{s} a_{ij} \sin(c_j\theta) &= \frac{c_i\sin(\theta)-\sin(c_i\theta)}{\theta^2} \quad i = 1, 2, \ldots s,
\end{aligned}$$

where $\theta = \omega h$. In both cases, the coefficients of the resulting methods are subjected to heavy numerical cancellation, so it is necessary to represent them through their expansion in power series of θ, as it is shown in the examples of methods provided in Section 10.2.4. It is also possible to prove that, for $\theta \to 0$, the coefficients of the resulting trigonometrically–fitted method tend to the coefficient of the corresponding collocation two–step hybrid methods [90], also reported in Section 10.2.2.

10.2.2 Construction of methods with coefficients depending on two frequencies

We now deal with the case of second order ODEs whose solution depends on two frequencies ω_1 and ω_2, both estimated in advance. We require that the methods must exactly solve the problem when its solution is linear combination of the basis functions

$$\{1, t, \ldots, t^{s-3}, \cos(\omega_1 t), \sin(\omega_1 t), \cos(\omega_2 t), \sin(\omega_2 t)\},$$

with $s \geq 4$. In order to derive such methods, we impose the following set of conditions

$$\begin{aligned}
\mathcal{L}_j[\cos\omega_1 t; h] &= 0, \quad \hat{L}[\cos\omega_1 t; h] = 0, \quad j = 1, \ldots, s, \\
\mathcal{L}_j[\sin\omega_1 t; h] &= 0, \quad \hat{L}[\sin\omega_1 t; h] = 0, \quad j = 1, \ldots, s, \\
\mathcal{L}_j[\cos\omega_2 t; h] &= 0, \quad \hat{L}[\cos\omega_2 t; h] = 0, \quad j = 1, \ldots, s, \\
\mathcal{L}_j[\sin\omega_2 t; h] &= 0, \quad \hat{L}[\sin\omega_2 t; h] = 0, \quad j = 1, \ldots, s.
\end{aligned} \tag{10.2.4}$$

Then if we are interested in methods with 4 stages, we only have to solve the system (10.2.4) in the unknowns $a_{ij}, b_i, \quad i, j = 1, ..., 4$. It has now become clear that, if we annihilate also

$$\mathcal{L}_j[t^q; h] = 0, \quad \hat{L}[t^q; h] = 0 \quad j = 1, 2, ..., s, \quad q = 2, 3..., s - 3, \tag{10.2.5}$$

more stages are necessary. The methods derived by solving (10.2.4), whose coefficients are listed in Appendix, have parameters depending on $\theta_1 = \omega_1 h$ and $\theta_2 = \omega_2 h$.

Also in this case, for $\theta_1 \to 0$ and $\theta_2 \to 0$, the coefficients of the derived methods tend to the coefficients of the two–step collocation hybrid methods of Section 10.2.2 and Chapter 9.

10.2.3 Linear stability analysis

We next analyze the linear stability properties [65, 203, 204] of the resulting methods, taking into account the dependency on the parameters. The following definitions are the natural adaptation of the ones formulated in Section 9.4 for methods depending on constant coefficients.

We first analyze the stability properties of mixed trigonometrically fitted methods depending on one frequency. In [66] the authors discussed the modifications introduced in the linear stability analysis, when the parameters depend on one fitted frequency ω. As a consequence of the presence of the fitted frequency ω, the interval of stability now becomes a two-dimensional region for the one parameter family of methods.
In this analysis, we denote the stability matrix as $M(\nu^2, \theta)$ and $R(\nu^2, \theta) = \frac{1}{2}\mathrm{tr}(M(\nu^2, \theta))$, $P(\nu^2, \theta) = \det(M(\nu^2, \theta))$, because it depends not only on

$\nu^2 = \lambda^2 h^2$ but also on $\theta = \omega h$.
The eigenvalues of the stability matrix $M(\nu^2, \theta)$ satisfy the following equation

$$\xi^2 - 2R(\nu^2, \theta)\xi + P(\nu^2, \theta) = 0. \tag{10.2.6}$$

It is known in literature (see [64, 66, 146]) that methods such that

$$|P(\nu^2, \theta)| \equiv 1, \tag{10.2.7}$$

i.e. the roots of (10.2.6) lie on the unit circle, are of particular interest. For example, Runge–Kutta Nyström methods based on polynomial approximations with symmetric abscissas c_i in $[0, 1]$, have an interval of periodicity, but they are not P–stable, if collocation based. If (10.2.7) holds, the study of periodicity can be developed just looking at the so–called *stability function* $R(\nu^2, \theta)$, in agreement with the following definition [66].

Definition 10.2.1 *For a trigonometrically fitted method of the type (9.1.1)-(9.1.2) satisfying $|P(\nu^2, \theta)| \equiv 1$, we define the primary interval of periodicity as the largest interval $(0, h_0)$ such that $|R(\nu^2, \theta)| < 1$, for all steplengths $h \in (0, h_0)$. If, when h_0 is finite, $|R(\nu^2, \theta)| < 1$ also for $\gamma < h\delta$, where $\gamma > h_0$, then the interval (γ, δ) is a secondary interval of periodicity.*

Exponentially fitted linear multistep methods in [66] verify (10.2.7), but only for few methods in the literature condition (10.2.7) holds. In our analysis, we found that for two-stage methods (9.1.1)-(9.1.2) depending on one frequency, the values of the abscissas such that $|P(\nu^2, \theta)| \equiv 1$ are only $c_1 = 0, c_2 = 1$. We next relax the definition 2 of region of stability in [66], in order to consider also methods for which $P(\nu^2, \theta) < 1$, in the following way [106]:

Definition 10.2.2 *A region of stability Ω is a region of the (ν, θ) plane, such that $\forall(\nu^2, \theta) \in \Omega$*

$$P(\nu^2, \theta) < 1, \quad |R(\nu^2, \theta)| < \frac{1}{2}(P(\nu^2, \theta) + 1). \tag{10.2.8}$$

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

Figure 10.1 and fig. 10.2 show some examples of stability regions for one–frequency depending methods, in the cases $s = 2$ and $s = 3$.

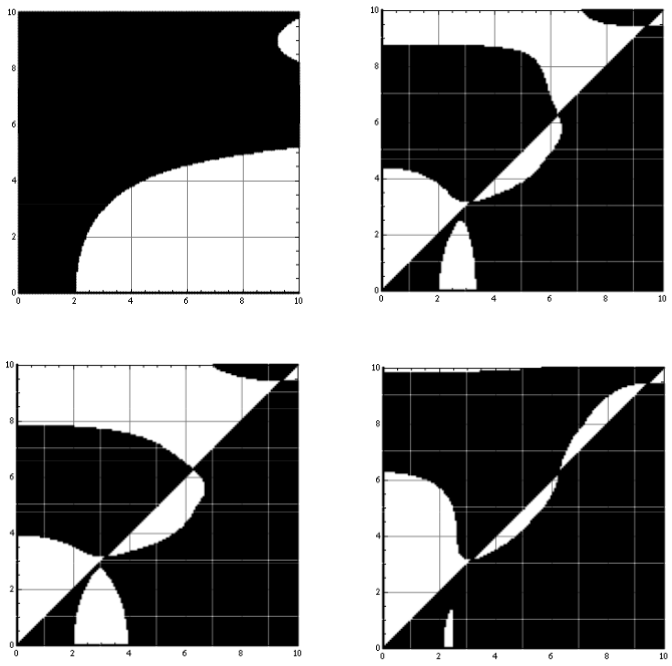

Figure 10.1: Regions of stability in the (ν, θ)–plane for the two-step methods for $s = 2$ with nodes $(0, 1)$, $(\frac{1}{7}, \frac{6}{7})$, $(\frac{1}{10}, \frac{9}{10})$, $(\frac{3}{4}, 1)$ respectively.

We next consider the linear stability analysis of methods depending on two frequencies. As stated before, for methods with constant coefficient, the stability region is an interval on the real axis, while methods depending on one frequency have bidimensional stability region. In the case of methods depending on the values of two frequencies, ω_1, ω_2, opportunely adapting the approach that Coleman and Ixaru in [66] introduced for one frequency depending methods, the stability region becomes tridimensional. We now denote the stability matrix of the methods as $M(\nu^2, \theta_1, \theta_2)$, with $\nu^2 = \lambda^2 h^2, \theta_1 = \omega_1 h, \theta_2 = \omega_2 h$. Its eigenvalues satisfy the following equation

$$\xi^2 - 2R(\nu^2, \theta_1, \theta_2)\xi + P(\nu^2, \theta_1, \theta_2) = 0, \tag{10.2.9}$$

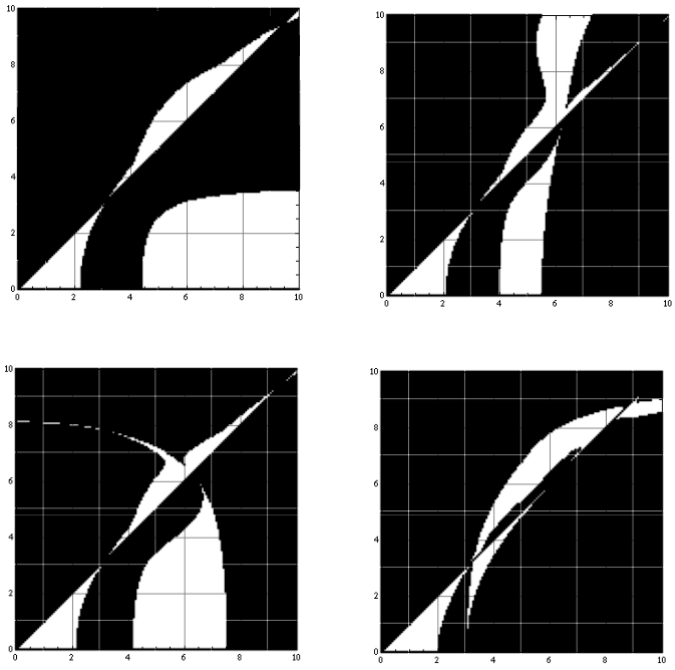

Figure 10.2: Regions of stability in the (ν, θ)–plane for the two-step methods for $s = 3$ with nodes $(0, \frac{1}{2}, 1)$, $(\frac{1}{4}, \frac{1}{2}, \frac{3}{4})$, $(\frac{1}{9}, \frac{1}{2}, \frac{8}{9})$, $(\frac{1}{2}, \frac{3}{4}, 1)$ respectively.

where $R(\nu^2, \theta_1, \theta_2) = \frac{1}{2}\text{tr}(M(\nu^2, \theta_1, \theta_2))$ and $P(\nu^2, \theta_1, \theta_2) = \det(M(\nu^2, \theta_1, \theta_2))$ are rational functions of ν^2. The definition of stability region for two frequencies depending methods can now be adapted as follows [106]:

Definition 10.2.3 *A three dimensional region Ω of the $(\nu^2, \theta_1, \theta_2)$ space is said to be the region of stability of the corresponding two-frequencies depending method if $\forall(\nu^2, \theta_1, \theta_2) \in \Omega$*

$$P(\nu^2, \theta_1, \theta_2) < 1, \quad |R(\nu^2, \theta_1, \theta_2)| < \frac{1}{2}(P(\nu^2, \theta_1, \theta_2) + 1). \tag{10.2.10}$$

Any closed curve defined by

$$P(\nu^2, \theta_1, \theta_2) \equiv 1, \quad |R(\nu^2, \theta_1, \theta_2)| = \frac{1}{2}(P(\nu^2, \theta_1, \theta_2) + 1). \tag{10.2.11}$$

is a stability boundary for the method.

Figure 10.3 shows an example of three dimensional stability region, while Figure 10.4 shows the projection of three dimensional regions on a particular plane.

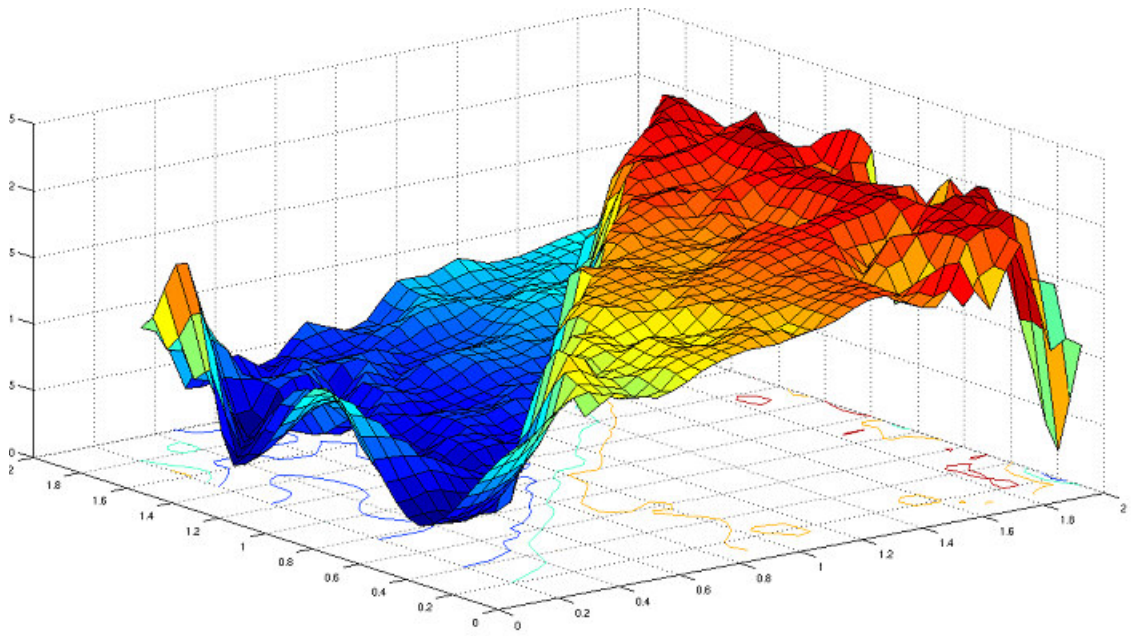

Figure 10.3: Region of stability in the $(\nu^2, \theta_1, \theta_2)$–plane for the two-step methods for $m = 4$ with nodes $(0, \frac{1}{3}, \frac{2}{3}, 1)$.

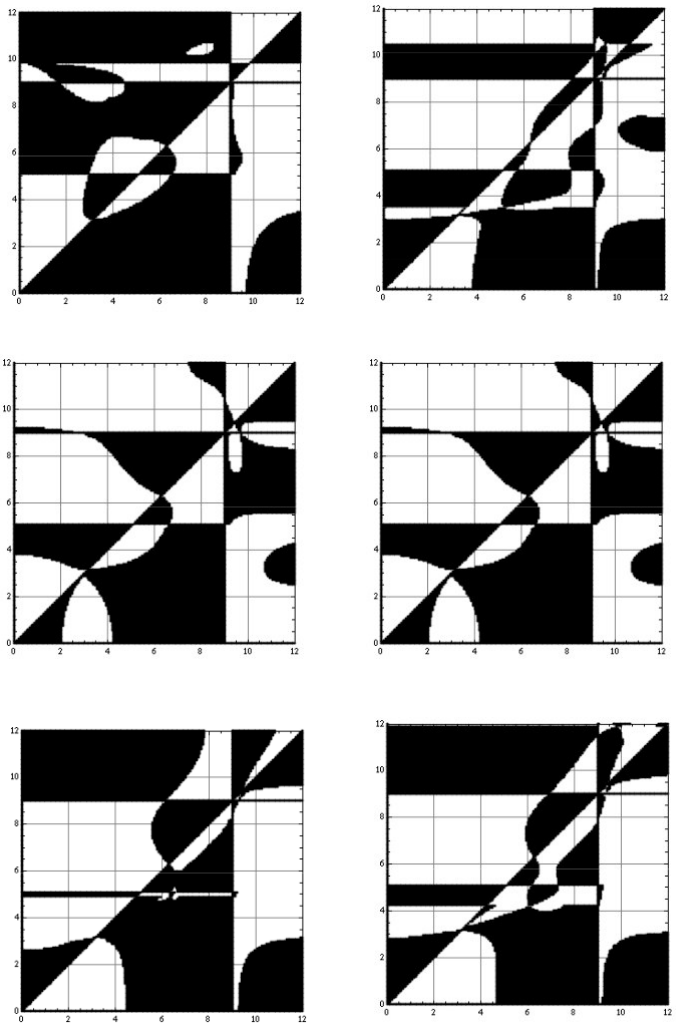

Figure 10.4: Regions of stability in the (ν, θ_1)–plane for the two-step methods for $m = 4$ with nodes $(0, \frac{1}{3}, \frac{2}{3}, 1)$, $(0, \frac{1}{10}, \frac{9}{10}, 1)$, $(0, \frac{2}{5}, \frac{3}{5}, 1)$, $(0, \frac{9}{20}, \frac{11}{20}, 1)$, $(\frac{1}{4}, \frac{3}{4})$, $(\frac{1}{5}, \frac{4}{5})$ respectively.

10.2.4 Examples of methods

We now provide some examples of trigonometrically-fitted methods (9.1.1)-

(9.1.2), using the analysis provided in the previous sections.

1. **Two-stage methods depending on one frequency.**
The solution of the system (10.2.3) for $s = 2$ is

$$\begin{aligned}
a_{11} &= -\frac{\csc((c_1-c_2)\theta)(\sin((c_1-c_2)\theta)+(1+c_1)\sin(c_2\theta)-c_1\sin((1+c_2)\theta))}{\theta^2},\\
a_{12} &= \frac{\csc((c_1-c_2)\theta)((1+c_1)\sin(c_1\theta)-c_1\sin((1+c_1)\theta)))}{\theta^2}\\
a_{21} &= \frac{(\csc((c_1-c_2)\theta)(-((1+c_2)\sin(c_2\theta))+c_2\sin((1+c_2)\theta)))}{\theta^2}\\
a_{22} &= \frac{(\csc((c_1-c_2)\theta)((1+c_2)\sin(c_1\theta)-c_2\sin((1+c_1)\theta)-\sin((c_1-c_2)\theta)))}{\theta^2}\\
b_1 &= \frac{2(-1+\cos(\theta))\csc((c_1-c_2)\theta)\sin(c_2\theta)}{\theta^2}\\
b_2 &= \frac{-2(-1+\cos(\theta))\csc((c_1-c_2)\theta)\sin(c_1\theta)}{\theta^2}
\end{aligned}$$

where $\theta = \omega h$. The coefficients expressed in this form are not of practical utility, because they are subject to heavy numerical cancellation: this is the reason why we handle their Taylor series expansion. For brevity, we give only the Taylor series expansion of the coefficients of the two stage method having $c = [3/4, 1]$

$$\begin{aligned}
a_{11} &= \tfrac{91}{32} - \tfrac{4375}{6144}\theta^2 + \tfrac{198451}{2949120}\theta^4 - \tfrac{263429}{75497472}\theta^6 + \tfrac{62606173}{543581798400}\theta^8\\
&- \tfrac{4225415771}{1607262661509120}\theta^{10} + O(\theta^{12})\\
a_{12} &= -\tfrac{35}{16} + \tfrac{287}{768}\theta^2 - \tfrac{475}{18432}\theta^4 + \tfrac{2921}{2949120}\theta^6 - \tfrac{5149}{212336640}\theta^8 + \tfrac{1606607}{3923981107200}\theta^{10}\\
&+ O(\theta^{12})\\
a_{21} &= 4 - \tfrac{23}{24}\theta^2 + \tfrac{2071}{23040}\theta^4 - \tfrac{24019}{5160960}\theta^6 + \tfrac{4565341}{29727129600}\theta^8 - \tfrac{110038253}{31391848857600}\theta^{10}\\
&+ O(\theta^{12})\\
a_{22} &= -3 + \tfrac{1}{2}\theta^2 - \tfrac{11}{320}\theta^4 + \tfrac{71}{53760}\theta^6 - \tfrac{2503}{77414400}\theta^8 + \tfrac{3719}{6812467200}\theta^{10} + O(\theta^{12})\\
b_1 &= 4 - \tfrac{23}{24}\theta^2 + \tfrac{2071}{23040}\theta^4 - \tfrac{24019}{5160960}\theta^6 + \tfrac{4565341}{29727129600}\theta^8 - \tfrac{110038253}{31391848857600}\theta^{10}\\
&+ O(\theta^{12})\\
b_2 &= -3 + \tfrac{1}{2}\theta^2 - \tfrac{11}{320}\theta^4 + \tfrac{71}{53760}\theta^6 - \tfrac{2503}{77414400}\theta^8 + \tfrac{3719}{6812467200}\theta^{10} + O(\theta^{12}).
\end{aligned}$$

This representation of the coefficients is very expressive, because it allows us to easily consider what follows. First of all we can notice that,

for $\theta \to 0$, these coefficients tend to the ones of the corresponding polynomial collocation method. Moreover, we observe that the following relations hold

$$b_1 + b_2 = 1 + O(\theta^2)$$
$$b_1 c_1 + b_2 c_2 = O(\theta^2)$$
$$b_1 c_1^2 + b_2 c_2^2 = -c_1 c_2 + O(\theta^2)$$

and, therefore, the method has algebraic order 2. The stability region in the (ν, θ)-plane of the methods corresponding to some values of c_1 and c_2 are drawn in Figure 10.1.

2. **Three stage mixed trigonometrically fitted methods depending on one frequency.**

Solving the system of equations for $m = 3$, we derive the coefficients of three stage methods depending on one frequency. We omit their expression because it is huge and it has no practical utility because of the heavy numerical cancellation it is subject to; anyway the expression of the coefficients can be required to the authors. In our numerical experiment we have used the Taylor expansion of the coefficients which, in correspondence to the abscissa vector $(c_1, c_2, c_3) = (\frac{1}{2}, \frac{3}{4}, 1)$, take the following form:

$$a_{11} = \frac{7}{2} - \frac{1081\theta^2}{2560} + \frac{251761\theta^4}{10321920} - \frac{1384021\theta^6}{1651507200} + O(\theta^8)$$
$$a_{12} = -\frac{21}{4} + \frac{2293\theta^2}{3840} - \frac{163169\theta^4}{5160960} + \frac{279607\theta^6}{275251200} + O(\theta^8)$$
$$a_{13} = \frac{17}{8} - \frac{1343\theta^2}{7680} + \frac{24859\theta^4}{3440640} - \frac{293621\theta^6}{1651507200} + O(\theta^8)$$
$$a_{21} = \frac{693}{128} - \frac{78169\theta^2}{122880} + \frac{431851\theta^4}{11796480} - \frac{3163721\theta^6}{2516582400} + O(\theta^8)$$
$$a_{22} = -\frac{511}{64} + \frac{18403\theta^2}{20480} - \frac{279851\theta^4}{5898240} + \frac{5752283\theta^6}{3774873600} + O(\theta^8)$$
$$a_{23} = \frac{413}{128} - \frac{32249\theta^2}{122880} + \frac{42617\theta^4}{3932160} - \frac{2013403\theta^6}{7549747200} + O(\theta^8)$$

$$a_{31} = \frac{22}{3} - \frac{17\theta^2}{20} + \frac{2953\theta^4}{60480} - \frac{129767\theta^6}{77414400} + O(\theta^8)$$
$$a_{32} = -\frac{32}{3} + \frac{6\theta^2}{5} - \frac{3827\theta^4}{60480} + \frac{78647\theta^6}{38707200} + O(\theta^8)$$
$$a_{33} = \frac{13}{3} - \frac{7\theta^2}{20} + \frac{437\theta^4}{30240} - \frac{27527\theta^6}{77414400} + O(\theta^8)$$
$$b_1 = \frac{22}{3} - \frac{17\theta^2}{20} + \frac{2953\theta^4}{60480} - \frac{129767\theta^6}{77414400} + O(\theta^8)$$
$$b_2 = -\frac{32}{3} + \frac{6\theta^2}{5} - \frac{3827\theta^4}{60480} + \frac{78647\theta^6}{38707200} + O(\theta^8)$$
$$b_3 = \frac{13}{3} - \frac{7\theta^2}{20} + \frac{437\theta^4}{30240} - \frac{27527\theta^6}{77414400} + O(\theta^8).$$

We next compute the order of this method, applying the set of order conditions

$$\begin{aligned}
\sum_{i=1}^{3} b_i &= 1 + O(\theta^2) \\
\sum_{i=1}^{3} b_i c_i &= O(\theta^2) \\
\sum_{i=1}^{3} b_i c_i^2 &= \frac{1}{6} + O(\theta^2) \\
\sum_{i=1}^{3}\sum_{j=1}^{3} b_i a_{ij} &= \frac{1}{12} + O(\theta^2) \\
\sum_{i=1}^{3} b_i c_i^3 &\neq O(\theta^2).
\end{aligned}$$

obtaining that the method has algebraic order 3. The stability region is drawn in Figure 10.2.

3. **Four stage methods depending on two frequencies.**
We now consider four stage methods of order 4 depending on two frequencies. The following method comes out setting $(c_1, c_2, c_3, c_4) = (0, \frac{1}{3}, \frac{2}{3}, 1)$. We report some terms of the Taylor series expansion of

its coefficients.

$$
\begin{aligned}
a_{11} &= 0, \quad a_{12} = 0, \quad a_{13} = 0, \quad a_{14} = 0, \\
a_{21} &= \frac{13676040 - 658854\theta_2^2}{16533720} - \frac{\theta_1^2(355781160 - 6480270\theta_2^2)}{8928208800} + O(\theta_1^4) + O(\theta_2^4) \\
a_{22} &= -\frac{71}{54} + \frac{833\theta_2^2}{9720} + \frac{\theta_1^2(510095880 - 35167230\theta_2^2)}{5952139200} + O(\theta_1^4) + O(\theta_2^4) \\
a_{23} &= \frac{26}{27} - \frac{7\theta_2^2}{135} - \frac{\theta_1^2(7 - 533\theta_2^2)}{367416} + O(\theta_1^4) + O(\theta_2^4) \\
a_{24} &= \frac{-8368920 + 198450\theta_2^2}{33067440} + \frac{\theta_1^2(21432600 - 103626\theta_2^2)}{3571283520} + O(\theta_1^4) + O(\theta_2^4) \\
a_{31} &= \frac{539}{324} - \frac{929\theta_2^2}{11664} - \frac{\theta_1^2(929 - 27443\theta_2^2)}{18895680} + O(\theta_1^4) + O(\theta_2^4) \\
a_{32} &= -\frac{137}{54} + \frac{37\theta_2^2}{216} + \frac{\theta_1^2(203915880 - 14061762\theta_2^2)}{1190427840} + O(\theta_1^4) + O(\theta_2^4) \\
a_{33} &= 42661080 - 2285010\theta_2^2 - \frac{\theta_1^2(246781080 - 6909354\theta_2^2)}{2380855680} + O(\theta_1^4) + O(\theta_2^4) \\
a_{34} &= -8368920 + 198450\theta_2^2 + \frac{\theta_1^2(21432600 - 103626\theta_2^2)}{1785641760} + O(\theta_1^4) + O(\theta_2^4) \\
a_{41} &= \frac{5}{2} - \frac{43\theta_2^2}{360} + \frac{\theta_1^2(-8777160 + 160110\theta_2^2)}{73483200} + O(\theta_1^4) + O(\theta_2^4) \\
a_{42} &= -\frac{15}{4} + \frac{37\theta_2^2}{144} + \frac{\theta_1^2(7552440 - 520722\theta_2^2)}{29393280} + O(\theta_1^4) + O(\theta_2^4) \\
a_{43} &= 3 - \frac{7\theta_2^2}{45} + \frac{\theta_1^2(-5715360 + 160110\theta_2^2)}{36741600} + O(\theta_1^4) + O(\theta_2^4) \\
a_{44} &= -\frac{3}{4} + \frac{13\theta_2^2}{720} + \frac{\theta_1^2(2653560 - 12690\theta_2^2)}{146966400} + O(\theta_1^4) + O(\theta_2^4) \\
b_1 &= \frac{5}{2} - \frac{43\theta_2^2}{360} + \frac{\theta_1^2(-8777160 + 160110\theta_2^2)}{73483200} + O(\theta_1^4) + O(\theta_2^4) \\
b_2 &= -\frac{15}{4} + \frac{37\theta_2^2}{144} + \frac{\theta_1^2(7552440 - 520722\theta_2^2)}{29393280} + O(\theta_1^4) + O(\theta_2^4) \\
b_3 &= \frac{3 - (7\theta_2^2)}{45} + \frac{\theta_1^2(-5715360 + 160110\theta_2^2)}{36741600} + O(\theta_1^4) + O(\theta_2^4) \\
b_4 &= -\frac{3}{4} + \frac{13\theta_2^2}{720} + \frac{\theta_1^2(2653560 - 12690\theta_2^2)}{146966400} + O(\theta_1^4) + O(\theta_2^4)
\end{aligned}
$$

Figures 10.3 and 10.4 show the stability region of this method and other methods, obtained in correspondence of different values of the abscissa.

10.2.5 Numerical experiments

We now show some numerical results we have obtained applying our families of solvers to some linear and nonlinear problems depending on one or two frequencies, in order to test the accuracy of the derived methods and also to compare them with ones already considered in literature for second order ODEs. We will consider the following solvers:

- COLEM: two–step hybrid method, [63],

$$\begin{array}{c|cc} \frac{1}{\sqrt{6}} & \frac{1+\sqrt{6}}{12} & 0 \\ -\frac{1}{\sqrt{6}} & -\frac{\sqrt{6}}{12} & \frac{1}{12} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array} \tag{10.2.12}$$

- TRIGFIT1: trigonometrically fitted two–step hybrid method, with 2 stages and order 2, $c = [0, 1]$;
- TRIGFIT2: trigonometrically fitted two–step hybrid method, with 2 stages and order 2, $c = [0, 3/4]$;
- TRIGFIT3: trigonometrically fitted two–step hybrid method, with 2 stages and order 2, $c = [3/4, 1]$;
- POL: two–step hybrid method [90], with 2 stages and order 2;
- MTRIGFIT: mixed–trigonometrically fitted two–step hybrid method, with 3 stages and order 3,$c = [1/3, 1/2, 1]$;
- TRIGFIT4S: trigonometrically fitted two–step hybrid method, with 4 stages and order 4, $c = [0, 1/3, 2/3, 1]$.

Test 1. We consider the following test equation

$$\begin{cases} y''(t) = -25y(t), \quad t \in [0, 2\pi], \\ y'(0) = 0, \\ y(0) = 1 \end{cases} \tag{10.2.13}$$

whose exact solution is $y(t) = \cos(5t)$, so it depends on the frequency $\omega = 5$. Table 10.2.5 compares the new methods with classical ones having constant

Method	$h = \pi/64$	cd	$h = \pi/128$	cd
COLEM	7.31e-02	1.13	4.20e-03	2.37
POL	2.87e-03	2.54	6.94e-04	3.16
TRIGFIT1	4.22e-15	14.37	4.44e-16	15.35
TRIGFIT2	4.44e-15	14.35	4.44e-16	15.35
TRIGFIT3	2.89e-15	14.54	8.88e-16	15.05
MTRIGFIT	2.22e-15	14.65	1.11e-13	13.95

Table 10.1: Numerical results for the problem (10.2.13).

coefficients, reporting the global error in the final point of the integration interval and the number *cd* of achieved correct digits. Trigonometrically fitted methods would solve this kind of problem exactly, of course, in exact arithmetic. The errors are the effect of the accumulation of round off errors in finite precision calculation.

Test 2. The Prothero-Robinson problem

$$y'' + v^2[y - \cos(10t)]^3 = -100y, \quad t \in [0, 20\pi], \tag{10.2.14}$$

with $v >> 0$, $y(0) = 1$, $y'(0) = 0$, whose exact solution is $y(t) = \cos(10t)$, is an example of nonlinear equation, depending on the frequency $\omega = 10$.

Method	$h = \pi/8$	cd	$h = \pi/16$	cd	$h = \pi/32$	cd
TRIGFIT1	4.44e-16	15.35	8.88e-15	14.05	6.22e-15	14.21
TRIGFIT2	1.77e-15	14.75	4.22e-15	14.37	6.88e-15	14.16
TRIGFIT3	2.00e-15	14.70	4.22e-15	14.37	6.88e-15	14.16
MTRIGFIT	9.40e-13	12.07	1.49e-13	12.83	2.19e-14	13.66

Table 10.2: Numerical results for the problem (10.2.14).

Numerical results show that trigonometrically fitted methods and mixed-trigonometrically fitted ones are both *exact* also for this nonlinear problem. Small differences in numerical errors are due to stability properties.

Test 3. We test our methods also on a well known example of stiff system, from [146]

$$y''(t) = \begin{pmatrix} \mu - 2 & 2\mu - 2 \\ 1 - \mu & 1 - 2\mu \end{pmatrix} y(t), \quad y(0) = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \; y'(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{10.2.15}$$

we have already discussed in Section 9.5. The behaviour of our class of solvers is still similar to the one shown in the previous cases. The choice of the stepsize is such that the methods result stable, and it is possible to integrate this problem with a large stepsize. On the contrary, methods with constant coefficients are stable only for small values of the stepsize. Anyway methods in Table 10.2.5 show an accumulation of the error, due to finite precision calculation, that is bigger than in the other considered problems.

Method	$h = \pi/2$	cd	$h = \pi/4$	cd
TRIGFIT3	2.34e-10	9.63	1.95e-09	8.71
TRIGFIT4S	3.80e-10	9.42	9.65e-08	7.01
MTRIGFIT	1.50e-08	7.84	5.01e-08	7.30

Table 10.3: Numerical results for the problem 10.2.15.

10.3 Exponentially-fitted two-step hybrid methods

This section aims to develop the exponentially-fitted version of two-step hybrid methods (9.1.1)-(9.1.2), aiming for the efficient integration of second order ODEs (8.0.1) whose solution exhibits a prominent exponential behaviour, also in this case under the assumption that it depends on the values of one or two parameters, which can be estimated in advance.

In particular, we require that both the internal and external stages of the resulting methods exactly integrate linear combinations of the following basis functions:

$$\{1, t, \ldots, t^K, \exp(\pm\mu_i t), t\exp(\pm\mu_i t), \ldots, \ldots, t^P \exp(\pm\mu_i t)\}, \tag{10.3.1}$$

depending on the values of the parameters μ_i, for $i = 1$ or $i = 2$. Also in this case the abscissa vector $\mathbf{c}$ is free from any assumption and, therefore, can be suitably chosen according to the properties we aim to achieve.

10.3.1 Construction of the methods

We now present the constructive technique we used to derive EF methods within the class (9.1.1)-(9.1.2), based on the so-called six-step procedure, introduced by Ixaru and Vanden Berghe in [135] 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 sin, cos 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 $\eta_k(s)$ functions introduced by Ixaru (see [132, 135] and references therein contained) and, as a consequence, the effects of numerical cancellation are notably reduced. We next report the first five steps of the procedure, while the remaining one, i.e. the local error analysis, is reported in the following section.

- step (i) *Computation of the classical moments.* The reduced classical moments (see [135], p. 42) are defined, in our case, as

$$L^*_{im}(\mathbf{a}) = h^{-(m+1)}\mathcal{L}_i[h;\mathbf{a}]t^m,\ i=1,\ldots,s,\ m=0,1,... \quad (10.3.2)$$
$$L^*_{m}(\mathbf{b}) = h^{-(m+1)}\mathcal{L}[h;\mathbf{b}]t^m,\ m=0,1,2,\ldots \quad (10.3.3)$$

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

$$L^*_{im}(\mathbf{a}) = 0,\ i=1,\ldots,s,\ m=0,1,\ldots,M-1 \quad (10.3.4)$$
$$L^*_{m}(\mathbf{b}) = 0,\ m=0,1,\ldots,M'-1. \quad (10.3.5)$$

to find out the maximal values of M and M' for which the above systems

are compatible. From the following relations

$$
\begin{array}{lll}
L_0^* = 0, & L_1^* = 0, & L_2^* = 2(1 - b_1 - b_2), \\
L_3^* = 6(-b_1c_1 - b_2c_2), & & L_4^* = 12(\frac{1}{6} - b_1c_1^2 - b_2c_2^2), \\
\\
L_{10}^* = 0, & L_{11}^* = 0, & L_{12}^* = c_1 + c_1^2 - 2(a_{11} + a_{12}), \\
L_{13}^* = -c_1(1 + 6a_{11} - c_1^2) - 6a_{12}c_2, & & L_{14}^* = c_1 + c_1^4 - 12(a_{11}c_1^2 + a_{12}c_2^2), \\
\\
L_{20}^* = 0, & L_{21}^* = 0, & L_{22}^* = c_2 + c_2^2 - 2(a_{21} + a_{22}), \\
L_{23}^* = -c_2(1 + 6a_{22} - c_2^2) - 6a_{21}c_1, & & L_{24}^* = c_2 + c_2^4 - 12(a_{21}c_1^2 + a_{22}c_2^2),
\end{array}
$$

we obtain $M = M' = 4$.

- 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 [135], p. 42), i.e.

$$
\begin{aligned}
E_{0i}^*(\pm z, \mathbf{a}) &= \exp \pm \mu t \mathcal{L}_i[h, \mathbf{a}] \exp(\pm \mu t), i = 1, \ldots, s, \quad (10.3.6) \\
E_0^*(\pm z, \mathbf{b}) &= \exp \pm \mu t \mathcal{L}[h, \mathbf{b}] \exp(\pm \mu t). \quad (10.3.7)
\end{aligned}
$$

Once computed the reduced exponential moments, we can derive the G functions, defined in the following way:

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

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

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

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 [135])

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

and obtaining

$$\begin{aligned}
G^{+(p)}(Z,\mathbf{b}) &= \frac{1}{2^{p-1}}\eta_{p-1}(Z) - \sum_{j=1}^{2} b_j \frac{d^p}{dZ^p}\left(Z\eta_{-1}(c_j^2 Z)\right),\\
G^{-(p)}(Z,\mathbf{b}) &= -\sum_{j=1}^{2} b_j c_j \frac{d^p}{dZ^p}\left(Z\eta_{-1}(c_j^2 Z)\right),\\
G_i^{+(p)}(Z,\mathbf{b}) &= \frac{c_i^{2p}}{2^p}\eta_{p-1}(c_i^2 Z) + \frac{c_i}{2^p}\eta_{p-1}(Z) - \sum_{j=1}^{2} a_{ij}\frac{d^p}{dZ^p}\left(Z\eta_{-1}(c_j^2 Z)\right),\\
G_i^{-(p)}(Z,\mathbf{b}) &= \frac{c_i^{2p+1}}{2^p}\eta_p(c_i^2 Z) - \frac{c_i}{2^p}\eta_p(Z) - \sum_{j=1}^{s} a_{ij}c_j\frac{d^p}{dZ^p}\left(Z\eta_0(c_j^2 Z)\right).
\end{aligned}$$

- 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 integrated (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.

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

where K and P are integer numbers satisfying the relation

$$K + 2P = M - 3 = 1. \tag{10.3.8}$$

Let us next consider the set of M' functions

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

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:

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

We focus our attention on the complete analysis of two-stage EF methods with $K = 1$ and $P = 0$ within the class (9.1.1)-(9.1.2), whose coefficients have been reported in Section 10.3.5. Therefore, in the case, we deal with the functional basis

$$\{1, t, \exp(\pm\mu t)\}. \tag{10.3.9}$$

Moreover, it is possible to extend the above procedure in order to derive EF methods belonging to the class (9.1.1)-(9.1.2), in the case of more than one frequency. In particular, the appendix reports the coefficients of two-parameters EF methods with 4 stages, with respect to the basis of functions

$$\{1, t, \exp(\pm\mu_1 t), \exp(\pm\mu_2 t)\}. \tag{10.3.10}$$

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

10.3.2 Error analysis and estimation of the parameters

According to the used procedure [135], the general expression of the local truncation error for an EF method with respect to the basis of functions (10.3.9) takes the form

$$lte^{EF}(t) = (-1)^{P+1} h^M \frac{L^*_{K+1}(\mathbf{b}(Z))}{(K+1)Z^{P+1}} D^2(D^2 - \mu^2)y(t), \tag{10.3.11}$$

with K, P and M satisfying the condition (10.3.8). Taking into account that, in our case, $K = 1$, $P = 0$ and $M = 4$, we obtain

$$lte^{EF}(t) = -h^2 \frac{L^*_2(\mathbf{b}(Z))}{2\mu^2} D^2(D^2 - \mu^2)y(t). \tag{10.3.12}$$

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

$$lte^{EF}(t_n) = -\frac{1}{6}\Big((1 + 6c_1c_2)(\mu^2 y^{(2)}(t_n) - y^{(4)}(t_n))\Big)h^4 + \mathcal{O}(h^5). \tag{10.3.13}$$

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 [134, 135] and references therein), generally based on the minimization of the leading term of the local discretization error. For this reason we annihilate the term $\mu^2 y^{(2)}(t_n) - y^{(4)}(t_n)$ and estimate the parameter in the following way:

$$\mu = \sqrt{\frac{y^{(4)}(t_n)}{y^{(2)}(t_n)}}, \quad \text{if} \quad y^{(2)}(t_n) \neq 0. \tag{10.3.14}$$

The expressions for the occurring derivatives can be obtained analytically from the given ODEs and, therefore, equation (10.3.14) can be re-written as

$$\mu = \sqrt{\frac{f_{yy}(t_n, y(t_n))y'(t_n) + f_y(t_n, y(t_n))f(t_n, y(t_n))}{f(t_n, y(t_n))}}, \quad \text{if} \quad f(t_n, y(t_n)) \neq 0. \tag{10.3.15}$$

The value of the first derivative $y'(t_n)$ can be approximated by means of sufficiently accurate finite differences.

Similar considerations can be applied in the case of four-stage EF methods depending on the values of two parameters μ_1 and μ_2, with respect to the functional basis (10.3.10). We consider the local truncation error, which depends on the reference differential equation

$$D^2(D^2-\mu_1^2)(D^2-\mu_2^2)y(t).$$

We next expand the in Taylor series around t, consider the leading term and evaluate it in t_n, obtaining the expression

$$3f_{yy}f^2+f(\beta-\alpha f_y+f_y^2+6f_{yyy}y\prime^2)+y\prime^2\bigg((5f_y-\alpha)f_{yy}+f_{yyyy}y\prime^2\bigg), \quad (10.3.16)$$

where $\alpha=\mu_1^2+\mu_2^2$, $\beta=\mu_1^2\mu_2^2$. For the sake of brevity, we have omitted in (10.3.16) the dependency of the function f and its derivatives on $(t_n,y(t_n))$ and the dependency of the solution y on t_n. An estimate to the unknown parameters μ_1 and μ_2 arises by annihilating the expression (10.3.16).

10.3.3 Examples of methods

We report the coefficients of EF methods (9.1.1)-(9.1.2) with $m=2$ and $m=4$ with respect to the basis (10.3.9) and (10.3.10) respectively. Two-stage EF methods within this class and exact on the functional basis (10.3.9)have the following coefficients:

$$b_1=-\tfrac{2c_2}{\delta}\bigg(-1+\eta_{-1}(Z)\bigg)\eta_0(c_2^2Z),$$
$$b_2=\tfrac{2c_1}{\delta}\bigg(-1+\eta_{-1}(Z)\bigg)\eta_0(c_1^2Z),$$
$$a_{11}=\tfrac{1}{\delta}\bigg(c_1\eta_{-1}(c_2^2,Z)(-\eta_0(Z)+\eta_0(c_1^2,Z))+c_2(1+c_1-c_1\eta_{-1}(Z)-\eta_{-1}(c_1^2Z))\eta_0(c_2^2Z)\bigg),$$
$$a_{12}=\tfrac{c_1}{\delta}\bigg(\eta_{-1}({c_1}^2Z)\eta_0(Z)+(-1-c_1+c_1\eta_{-1}(Z))\eta_0({c_1}^2Z)\bigg),$$
$$a_{21}=\tfrac{c_2}{\delta}\bigg(-(\eta_{-1}(c_2^2Z)\eta_0(Z))+(1+c_2-c_2\eta_{-1}(Z))\eta_0(c_2^2Z)\bigg),$$
$$a_{22}=\tfrac{c_1}{\delta}\bigg((-1-c_2+c_2\eta_{-1}(Z)+\eta_{-1}(c_2^2Z))\eta_0(c_1^2Z)+c_2\eta_{-1}(c_1^2Z)(\eta_0(Z)-\eta_0(c_2^2Z)\bigg),$$

where

$$\delta = Z(c_1\eta_{-1}(c_2^2 Z)\eta_0(c_1^2 Z) - c_2\eta_{-1}(c_1^2 Z)\eta_0(c_2^2 Z)).$$

It is easy to prove that, for Z tending to 0, such coefficients tend to the ones of based on algebraic collocation (see [90]): therefore, applying the order conditions derived in [63] for Z tending to 0, we discover that these methods have order 2. Fig. 10.1 shows an example of stability region for two-stage one-parameter depending method with $c_1 = \frac{2}{3}$, $c_1 = \frac{4}{5}$.

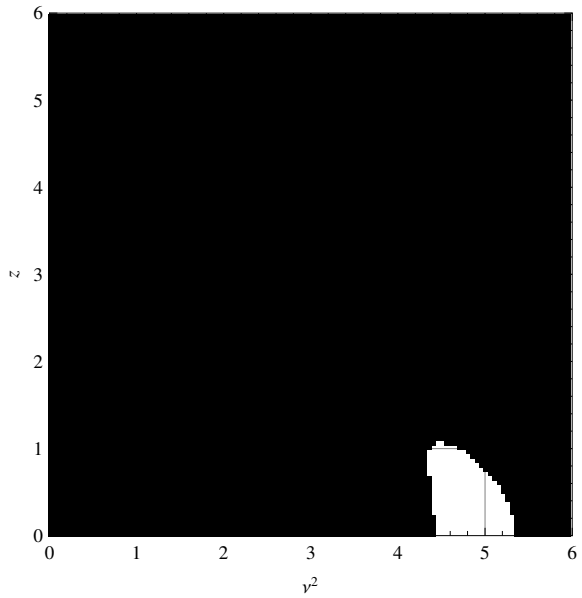

Figure 10.1: Region of stability in the $(\nu^2, Z)-$plane for $m = 2$, with $c_1 = \frac{2}{3}$, $c_1 = \frac{4}{5}$.

The coefficients of four-stage EF methods within the class (9.1.1)-(9.1.2) and exact on the functional basis (10.3.10) are too long to be reported here

and, for this reason, we present their truncated power series expansion.

$$
\begin{aligned}
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{581Z_2^2}{14580} + \left(\frac{581}{14580} + \frac{24001Z_2^2}{33067440}\right)Z_1^2 + O(Z_1^4) + O(Z_2^4),\\
a_{22} &= -\frac{71}{54} - \frac{833Z_2^2}{9720} + \left(-\frac{833}{9720} - \frac{18607Z_2^2}{3149280}\right)Z_1^2 + O(Z_1^4) + O(Z_2^4),\\
a_{23} &= \frac{26}{27} + \frac{7Z_2^2}{135} + \left(\frac{7}{135} + \frac{533Z_2^2}{367416}\right)Z_1^2 + O(Z_1^4) + O(Z_2^4),\\
a_{24} &= -\frac{41}{162} - \frac{35Z_2^2}{5832} + \left(-\frac{35}{5832} - \frac{1919Z_2^2}{66134880}\right)Z_1^2 + O(Z_1^4) + O(Z_2^4),\\
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 the ones of based on algebraic collocation and the corresponding methods have algebraic order 4. Figure 10.2 shows an example of tridimensional stability region for four-stage two-parameter depending methods with $c = [0, \frac{1}{3}, \frac{2}{3}, 1]^T$.

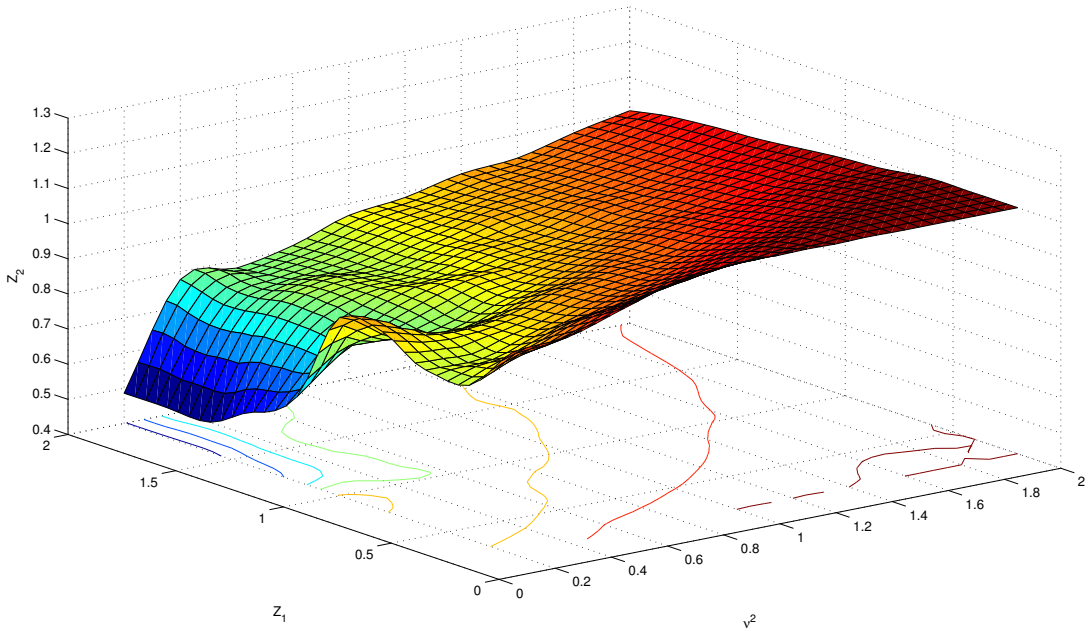

Figure 10.2: Region of stability in the (ν^2, Z_1, Z_2)-space for $m = 4$, with $c = [0, \frac{1}{3}, \frac{2}{3}, 1]^T$.

10.3.4 Numerical results

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

- COLEM2, two-step hybrid method (9.1.1)-(9.1.2) having constant coefficients (see [63])

$$\begin{array}{c|cc} \frac{1}{2} & -1 & \\ 1 & 2 & -1 \\ \hline & 2 & -1 \end{array} \tag{10.3.17}$$

  with $m = 2$ and $p = 2$;

- EXPCOLEM2, one-parameter depending exponentially-fitted method (9.1.1)-(9.1.2), with $m = 2$ and $p = 2$, whose coefficients are reported in the appendix.

We implement such methods in a fixed stepsize environment, with step $h = \frac{1}{2^k}$, with k positive integer number. The problems we aim to consider in this section are both linear and nonlinear, having solutions depending on one and two parameters. The numerical evidence confirm that EF-based methods within the class (9.1.1)-(9.1.2) are able to exactly integrate all those 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 (9.1.1)-(9.1.2) are quite often unstable and this also confirms that, 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

$$\begin{cases} y''(t) = \lambda^2 y(t), \\ y(0) = 1, \\ y'(0) = -\lambda, \end{cases} \tag{10.3.18}$$

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 10.1 shows the results we have obtained by using the above numerical methods.

λ	k	COLEM2	EXPCOLEM2
2	4	8.32e-1	4.20e-14
	5	2.29e-1	7.38e-15
	6	5.96e-2	1.20e-13
3	4	Unstable	1.92e-13
	5	Unstable	3.17e-13
	6	Unstable	2.50e-14
4	4	Unstable	1.91e-12
	5	Unstable	2.46e-12
	6	Unstable	3.83e-13

Table 10.1: Relative errors corresponding to the solution of the problem (10.3.18), for different values of λ and k.

Problem 2. We examine the following linear equation

$$\begin{cases} y''(t) - y(t) = t - 1, \\ y(0) = 2, \\ y'(0) = -2, \end{cases} \tag{10.3.19}$$

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 (10.3.9). The obtained results are reported in Table 10.2.

k	COLEM2	EXPCOLEM2
4	Unstable	3.57e-13
5	8.52e-1	1.67e-15
6	2.71e-1	5.07e-14

Table 10.2: Relative errors corresponding to the solution of the problem (10.3.19).

k	COLEM2	EXPCOLEM2
3	Unstable	1.21e-15
4	Unstable	4.68e-15
5	Unstable	1.30e-14

Table 10.3: Relative errors corresponding to the solution of the problem (10.3.20), with $\nu = 100$.

Problem 3. We next focus on the Prothero-Robinson problem [180]

$$\begin{cases} y''(t) + \nu^2[y(t) - \exp(-\lambda t)]^3 = \lambda^2 y, \\ y(0) = 1, \\ y'(0) = -\lambda, \end{cases} \tag{10.3.20}$$

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 10.3.

Chapter 11

Runge-Kutta-Nystrom stability for a class of General Linear Methods for $y'' = f(x, y)$

We have already discussed in Chapter 2 the great potential of General Linear Methods for first order Ordinary Differential Equations (ODEs) and used here and there in the dissertation some of the results developed up to now for GLMs. The interest in this theory (compare [42, 138]) is involving also different functional equations, for instance Volterra Integral Equations [137]. For special second order ODEs $y" = f(t, y)$ no systematic investigation on GLMs has begun till now, even if many linear and nonlinear methods appeared in the literature (see, for instance, [135]). We would like to begin our investigation on GLMs for second order ODEs, starting from the linear stability analysis of a special class of methods which, up to now, constitute one of the most general family of methods for (8.0.1), i.e. the class of *two-step*

Runge-Kutta-Nyström (TSRKN) methods

$$\begin{aligned}
Y_j^{[i-1]} &= y_{i-1} + hc_j y'_{i-1} + h^2 \sum_{s=1}^{m} \bar{a}_{js} f(t_{i-1} + c_s h, Y_s^{[i-1]}), \qquad j = 1, \ldots, m \\
Y_j^{[i]} &= y_i + hc_j y'_i + h^2 \sum_{s=1}^{m} \bar{a}_{js} f(t_i + c_s h, Y_s^{[i]}), \qquad j = 1, \ldots, m, \\
y_{i+1} &= (1-\theta) y_i + \theta y_{i-1} + h \sum_{j=1}^{m} v_j y'_{i-1} + h \sum_{j=1}^{m} w_j y'_i \\
&+ h^2 \sum_{j=1}^{m} \big(\bar{v}_j f(t_{i-1} + c_j h, Y_j^{[i-1]}) + \bar{w}_j f(t_i + c_j h, Y_j^{[i]}) \big), \\
y'_{i+1} &= (1-\theta) y'_i + \theta y'_{i-1} + h \sum_{j=1}^{m} \big(v_j f(t_{i-1} + c_j h, Y_j^{[i-1]}) \\
&+ w_j f(t_i + c_j h, Y_j^{[i]}) \big),
\end{aligned} \tag{11.0.1}$$

represented by the Butcher array

$$\begin{array}{c|c} \mathbf{c} & \bar{\mathbf{A}} \\ \hline & \bar{\mathbf{v}}^T \\ \theta & \bar{\mathbf{w}}^T \\ & \mathbf{v} \\ & \mathbf{w} \end{array} . \tag{11.0.2}$$

Such methods have been introduced and analyzed in [173, 174, 175, 176, 177]. In particular, they were derived in [173] as a family of indirect methods [204] through a transformation of two-step Runge-Kutta (TSRK) methods [139]. The analysis of convergence and consistency for TSRKN methods has already been provided in [173]. In this paper we derive new TSRKN methods, possessing the same stability properties of the best Runge-Kutta-Nyström (RKN) methods, that are the indirect Gauss-Legendre collocation-based methods [204], introducing the concept of RKN-stability, following the lines already drawn in the derivation of GLMs for first order ODEs [42, 138].

11.1 Stability analysis

In this section we present the framework on which the analysis of the linear stability properties of TSRKN methods is based. The stability matrix of TSRKN methods is (see [173])

$$\mathbf{M}(z^2) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \theta + \alpha(\bar{\mathbf{v}}, \mathbf{e}) & 1 - \theta + \alpha(\bar{\mathbf{w}}, \mathbf{e}) & \mathbf{v}^T\mathbf{e} + \alpha(\bar{\mathbf{v}}, \mathbf{c}) & \mathbf{w}^T\mathbf{e} + \alpha(\bar{\mathbf{w}}, \mathbf{c}) \\ 0 & 0 & 0 & 1 \\ \alpha(\mathbf{v}, \mathbf{e}) & \alpha(\mathbf{w}, \mathbf{e}) & \mathbf{v}^T\mathbf{e} + \theta + \alpha(\mathbf{v}, \mathbf{c}) & 1 - \theta + \alpha(\mathbf{w}, \mathbf{c}) \end{bmatrix}, \tag{11.1.1}$$

where $\alpha(\mathbf{x}, \mathbf{y}) = -z^2\mathbf{x}^T\mathbf{N}^{-1}\mathbf{y}$, $\mathbf{N} = \mathbf{I} + z^2\mathbf{A}$, $\mathbf{I}$ is the identity matrix of order m and $e = (1, \ldots, 1) \in \mathbb{R}^m$. We next consider the following definitions (see [203, 204]).

Definition 11.1.1 *$(0, \beta^2)$ is a stability interval for the method (11.0.1) if, $\forall z^2 \in (0, \beta^2)$, the spectral radius $\rho(\mathbf{M}(z^2))$ of the matrix $\mathbf{M}(z^2)$ is such that*

$$\rho(\mathbf{M}(z^2)) < 1. \tag{11.1.2}$$

The condition (11.1.2) 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 $S(z^2) = \mathrm{tr}(\mathbf{M}^2(z^2))$ and $P(z^2) = \det(\mathbf{M}^2(z^2))$, for a one-step RKN method the condition (11.1.2) is equivalent to

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

Definition 11.1.2 *The method (11.0.1) is A-stable if $(0, \beta^2) = (0, +\infty)$.*

If the eigenvalues $r_1(z^2), r_2(z^2), r_3(z^2)$, and $r_4(z^2)$ of the stability matrix (13.2.4) (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 11.1.3 *$(0, H_0^2)$ is a periodicity interval for (11.0.1) if, $\forall z^2 \in (0, H_0^2)$, $r_1(z^2)$ and $r_2(z^2)$ are complex conjugate and $|r_{1,2}(z^2)| = 1$, while $|r_{3,4}(z^2)| \leq 1$.*

For a one-step RKN method, the interval of periodicity is then defined by

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

Definition 11.1.4 *The method (11.0.1) is P-stable if its periodicity interval is $(0, +\infty)$.*

11.2 Runge-Kutta-Nyström stability

In the context of the numerical integration of first order ODEs, Runge-Kutta methods possess strong stability properties which are, in particular, superior to the ones of linear multistep methods. For this reason, in recent times, the attention of many authors has been devoted to the construction of general linear methods for first order ODEs having the same stability properties of Runge-Kutta methods (see [42, 138] and references therein). If $\mathbf{M}(z) \in \mathbb{R}^{s\times s}$ is the stability matrix of a certain GLM, this method is said to be *Runge-Kutta stable* if its stability polynomial $p(\omega, z)$ takes the form

$$p(\omega, z) = \omega^{s-1}(\omega - R(z)),$$

where $R(z)$ is a rational function. Butcher and Wright (2003; see also [42] and references therein) characterized Runge-Kutta stability in terms of algebraic conditions on the coefficient matrices of the method, introducing the concept of *inherent Runge-Kutta stability*.

In the case of second order ODEs, indirect RKN methods generated by highly stable Runge-Kutta methods inherit those stability properties: for example, Gaussian Runge-Kutta methods are A-stable and generate A-stable indirect RKN methods; Runge-Kutta methods based on Radau IIA collocation points are L-stable and generate L-stable indirect RKN methods (see [**?**]). Therefore, following the lines drawn in the literature in the context of GLMs for first order ODEs [42, 138], we aim for TSRKN methods having the same stability properties of highly stable RKN methods, according to the following definition.

Definition 11.2.1 *A TSRKN method (11.0.1) is said to be* ***Runge-Kutta-Nyström stable*** *if its stability polynomial exhibits the form*

$$p(\omega, z^2) = \omega^2\big(q_2(z)\omega^2 + q_1(z)\omega + q_0(z)\big), \tag{11.2.1}$$

where $q_2(z)\omega^2 + q_1(z)\omega + q_0(z)$ is the stability polynomial of a certain Runge-Kutta-Nyström method.

In other words, the stability properties of TSRKN methods having Runge-Kutta-Nyström stability (*RKN-stability*) are determined by the polynomial

$$q_2(z)\omega^2 + q_1(z)\omega + q_0(z),$$

which is exactly the stability polynomial of a RKN method. Therefore TSRKN methods with *RKN-stability* on Gaussian points are A-stable and, in particular, P-stable, while TSRKN method with *RKN-stability* on Radau IIA points are L-stable. In order to derive TSRKN methods with *RKN-stability*, we use the following procedure:

- the stability polynomial of the TSRKN method has, in general, degree 4 and its coefficients are rational functions of z, depending on the parameters of the method. It is transformed into a polynomial $\tilde{p}(\omega, z)$ of the type (11.2.1), i.e.

$$\tilde{p}(\omega, z^2) = \omega^2\big(\alpha_2(z)\omega^2 + \alpha_1(z)\omega + \alpha_0(z)\big);$$

- if we want to reproduce the stability properties of a certain RKN method whose stability function is $q_2(z)\omega^2 + q_1(z)\omega + q_0(z)$, we have to solve the system of equations

$$\alpha_2(z) = q_2(z), \quad \alpha_1(z) = q_1(z), \quad \alpha_0(z) = q_0(z). \tag{11.2.2}$$

Even if the system (11.2.2) is not solvable, the first step of this procedure (i.e. the reduction of the degree of the stability polynomial) would produce a polynomial having only two nonzero roots and this property is very similar to that of RKN methods. This leads to the following definition.

Definition 11.2.2 *A TSRKN method (11.0.1) is said to be* ***almost Runge-Kutta-Nyström stable*** *if its stability polynomial exhibits the form*

$$p(\omega, z^2) = \omega^2\big(p_2(z)\omega^2 + p_1(z)\omega + p_0(z)\big), \tag{11.2.3}$$

where $p_0(z)$, $p_1(z)$, and $p_2(z)$ are rational functions in z.

Almost Runge-Kutta-Nyström stability (*almost RKN-stability*) constitute a desirable tool for the practical derivation of highly stable TSRKN methods, because it requires the investigation of a quadratic polynomial instead of a polynomial of degree 4. This idea will be explored in a paper in preparation.

11.2.1 RKN stability on Gaussian points

We now consider the derivation of TSRKN methods having the same stability properties of indirect RKN methods based on Gauss-Legendre points, which are P-stable. Paternoster in [173] already provided a complete characterization of P-stable indirect methods of order 1 and 2 in the class (11.0.1): we now analyze if some of those methods possess the *RKN-stability* property inherited from the indirect RKN method based on the Gauss-Legendre point, i.e.

$$\begin{array}{c|c} 1/2 & 1/4 \\ \hline & 1/2 \\ & 1 \end{array} \tag{11.2.4}$$

The stability polynomial of one-stage TSRKN of order 1 with *almost RKN-stability* is

$$p(\omega, z^2) = \omega^2 \left(\omega^2 + \frac{-2 + (-2a + c + \bar{w})z^2}{1 + az^2}\omega + \frac{1 + (1 + a - c - \bar{w})z^2}{1 + az^2} \right),$$

and it is obtained imposing $w = 1$ and $v = \bar{v} = \theta = 0$. This means that TSRKN with *almost RKN-stability* or *RKN-stability* fall inside the class of RKN methods. The stability polynomial of the indirect one-stage Gauss-Legendre RKN method is

$$q(\omega, z^2) = \omega^2 + \frac{-8 + 2z^2}{4 + z^2}\omega + 1,$$

and, therefore, *RKN-stability* is forced by solving the linear system

$$\begin{cases} -4 + 4c + 4\bar{w} = 0 \\ -4a + c + \bar{w} = 0 \\ 1 - c - \bar{w} = 0 \end{cases}$$

As a consequence, we obtain the following family of P-stable methods

$$\begin{array}{c|c} c & 1/4 \\ \hline & 0 \\ 0 & 1 - c \\ & 0 \\ & 1 \end{array} \tag{11.2.5}$$

In particular, when $c = \frac{1}{2}$, order 2 is achieved and we obtain again the RKN method on the Gaussian point. These considerations lead to the following result.

Theorem 11.2.1 *One-stage TSRKN methods with order 1 and having RKN-stability are RKN methods themselves. The only one-stage TSRKN method of order 2 and having RKN-stability is the Gauss-Legendre RKN method itself.*

In the case of two-stage TSRKN methods, we have been able to derive a family of P-stable methods of order 2 and with *RKN-stability* inherited from Gauss-Legendre RKN methods. The derived family of methods is reported in the following result.

Theorem 11.2.2 *Two-stage TSRKN methods with order 2 and having RKN-stability are*

$$\begin{array}{c|cc} 1/2 & (1-4a_{12})/4 & a_{12} \\ 1/2 & (1-4a_{12})/4 & a_{22} \\ \hline & 0 & 0 \\ 0 & (1-2\bar{w}_2)/2 & \bar{w}_2 \\ & -v_2 & v_2 \\ & 1-w_2 & w_2 \end{array} \tag{11.2.6}$$

Part III

Numerical solution of Volterra Integral Equations

Chapter 12

Collocation methods for Volterra Integral Equations: a review

Piecewise polynomial collocation methods for Volterra Integral Equations (VIEs) introduce a number of aspects not present when solving ODEs. In this chapter we will present the main results in the context of collocation and almost collocation methods for VIEs of the form

$$y(t) = g(t) + \int_0^t k(t, \tau, y(\tau))d\tau \qquad t \in I := [0, T], \tag{12.0.1}$$

where $k \in C(D \times \mathbb{R}^d)$, with $D := \{(t, \tau) : 0 \leq \tau \leq t \leq T\} \subseteq \mathbb{R}^2$, and $g \in C(I)$, also underlining connections and differences with the case of ODEs. Let $I_h = \{t_n := nh, n = 0, ..., N,\ h \geq 0, Nh = T\}$ be a uniform mesh, which constitutes the discrete counterpart of the interval $[0, T]$. Equation (12.0.1) can be rewritten, by relating it to this mesh, in the form

$$y(t) = F^{[n]}(t, y(\cdot)) + \Phi^{[n+1]}(t, y(\cdot)), \quad t \in [t_n, t_{n+1}],$$

where

$$F^{[n]}(t, y(\cdot)) := g(t) + \int_0^{t_n} k(t, \tau, y(\tau))d\tau, \quad \Phi^{[n+1]}(t, y(\cdot)) := \int_{t_n}^t k(t, \tau, y(\tau))d\tau$$

are the *lag term* and the *increment term* respectively.

12.1 Classical one–step collocation methods

Let us fix m collocation parameters $0 \leq c_1 < ... < c_m \leq 1$ which identify m internal points $t_{nj} = t_n + c_j h$. The collocation polynomial, restricted to the interval $[t_n, t_{n+1}]$, is of the form:

$$P_n(t_n + sh) = \sum_{j=1}^{m} L_j(s) Y_j^{[n+1]}, \quad s \in [0,1], \quad n = 0, ..., N-1, \tag{12.1.1}$$

where $L_j(s)$ is the j^{th} Lagrange fundamental polynomial with respect to the collocation parameters and $Y_j^{[n+1]} := P_n(t_{nj})$. *Exact* collocation methods are obtained by imposing that the collocation polynomial (12.1.1) exactly satisfies the VIE (12.0.1) in the collocation points t_{ni} and by computing $y_{n+1} = P_n(t_{n+1})$:

$$\begin{cases} Y_i^{[n+1]} = F_i^{[n]} + \Phi_i^{[n+1]} \\ y_{n+1} = \sum_{j=1}^{m} L_j(1) Y_j^{[n+1]} \end{cases}, \tag{12.1.2}$$

where

$$F_i^{[n]} = g(t_{ni}) + h \sum_{\nu=0}^{n-1} \int_0^1 k(t_{ni}, t_\nu + sh, P_\nu(t_\nu + sh)) ds \quad i = 1, ..., m \tag{12.1.3}$$

$$\Phi_i^{[n+1]} = h \int_0^{c_i} k(t_{ni}, t_n + sh, P_n(t_n + sh)) ds \quad i = 1, ..., m. \tag{12.1.4}$$

Note that the first equation in (12.1.2) represents a system of m nonlinear equations in the m unknowns $Y_i^{[n+1]}$. We obtain an approximation $P(t)$ of the solution $y(t)$ of the integral equation (12.0.1) in $[0, T]$, by considering

$$P(t)|_{(t_n, t_{n+1}]} = P_n(t) \tag{12.1.5}$$

where $P_n(t)$ given by (12.1.1).

We recall that, in contrast with what happens in the case of ODEs, generally $P(t)$ is not continuous in the mesh points, as

$$P(t) \in S_{m-1}^{(-1)}(I_h), \tag{12.1.6}$$

where

$$S_\mu^{(d)}(I_h) = \left\{ v \in C^d(I) : v|_{(t_n, t_{n+1}]} \in \Pi_\mu \ (0 \leq n \leq N-1) \right\}.$$

Here, Π_μ denotes the space of (real) polynomials of degree not exceeding μ. A complete analysis of collocation methods for linear and nonlinear Volterra integral and integro–differential equations, with smooth and weakly singular kernels is given in [24]. In particular, as shown in [24, 27], the classical one–step collocation methods for a second-kind VIE do no longer exhibit $O(h^{2m})$ superconvergence at the mesh points if collocation is at the Gauss points, in fact they have uniform order m for any choice of the collocation parameters and local superconvergence order in the mesh points of $2m-2$ (m Lobatto points or $m-1$ Gauss points with $c_m = 1$) or $2m-1$ (m Radau II points). The optimal order is recovered only in the iterated collocation solution.

We observe that, differently from the case of ODEs, the collocation equations are in general not yet in a form amenable to numerical computation, due to the presence of the memory term given by the Volterra integral operator. Thus, another discretization step, based on quadrature formulas $\bar{F}_i^{[n]} \simeq F_i^{[n]}$ and $\bar{\Phi}_i^{[n+1]} \simeq \Phi_i^{[n+1]}$ for approximating the lag term (12.1.3) and the increment function (12.1.4), is necessary to obtain the fully discretized collocation scheme, thus leading to *discretized* collocation methods. Such methods preserve, under suitable hypothesis on the quadrature formulas, the same order of the exact collocation methods [27].

The connection between collocation and implicit Runge–Kutta methods for VIEs (the so called VRK methods) is not immediate: a collocation method for VIEs is equivalent to a VRK method if and only if $c_m = 1$ (see Theorem 5.2.2 in [27]). Some other continuous extensions of Runge–Kutta methods for VIEs, which do not necessarily lead to collocation methods, have been introduced in [21].

Many efforts have been made in the literature with the aim of obtaining fast collocation and more general Runge–Kutta methods for the numerical solution of VIEs. It is known that the numerical treatment of VIEs is very expensive from computational point of view because of presence of the "lag-term", which contains the entire history of the phenomenon. To this cost, it has also to be added the one due to the "increment term" which leads, for implicit methods (generally possessing the best stability properties), to the resolution of a system of nonlinear equations at each step of integration. In order to reduce the computational effort in the lag–term computation, fast collocation and Runge–Kutta methods have been constructed for convolution VIEs of Hammerstein type, see [54, 74, 155, 156].

The stability analysis of collocation and Runge–Kutta methods for VIEs

can be found in [20, 27, 55, 79] and the related bibliography. In particular a collocation method for VIEs is A–stable if the corresponding method for ODEs is A–stable.

12.2 Multistep collocation

Multistep collocation and Runge–Kutta methods for VIEs, have been introduced in order to bring down the computational cost related to the resolution of non-linear systems for the computation of the increment term. As a matter of fact such methods, showing a dependence on stages and steps in more consecutive grid points, permit to raise the order of convergence of the classical methods, without inflating the computational cost or, equivalently, having the same order at a lower computational cost.

A first analysis of multistep collocation methods for VIEs appeared in [76, 77], where the methods are obtained by introducing in the collocation polynomial the dependence from μ previous time steps; namely we seek for a collocation polynomial, whose restriction to the interval $[t_n, t_{n+1}]$ takes the form

$$P_n(t_n + sh) = \sum_{k=0}^{\mu-1} \varphi_k(s) y_{n-k} + \sum_{j=1}^{m} \psi_j(s) Y_j^{[n+1]}, \quad s \in [0,1], \quad n = 0, ..., N-1, \tag{12.2.1}$$

where

$$Y_j^{[n+1]} := P_n(t_{nj}) \tag{12.2.2}$$

and $\varphi_k(s)$, $\psi_j(s)$ are polynomials of degree $m + \mu - 1$ to be determined by imposing the interpolation conditions at the points t_{n-k}, that is $P_n(t_{n-k}) = y_{n-k}$, and by satisfying (12.2.2). It is proved in [71, 75] that, assuming $c_i \neq c_j$ and $c_1 \neq 0$, the polynomials $\varphi_k(s)$, $\psi_j(s)$ have the form:

$$\begin{aligned} \varphi_k(s) &= \prod_{i=1}^{m} \frac{s - c_i}{-k - c_i} \cdot \prod_{\substack{i=0 \\ i \neq k}}^{\mu-1} \frac{s+i}{-k+i}, \\ \psi_j(s) &= \prod_{i=0}^{\mu-1} \frac{s+i}{c_j + i} \cdot \prod_{\substack{i=1 \\ i \neq j}}^{m} \frac{s - c_i}{c_j - c_i}. \end{aligned} \tag{12.2.3}$$

The discretized multistep collocation method is then obtained by imposing the collocation conditions, i.e. that the collocation polynomial (12.2.1)

exactly satisfies the VIE (12.0.1) at the collocation points t_{ni}, and by computing $y_{n+1} = P_n(t_{n+1})$:

$$\begin{cases} Y_i^{[n+1]} = F_i^{[n]} + \Phi_i^{[n+1]}, \\ y_{n+1} = \sum_{k=0}^{\mu-1} \varphi_k(1) y_{n-k} + \sum_{j=1}^{m} \psi_j(1) Y_j^{[n+1]}. \end{cases} \tag{12.2.4}$$

The lag–term and increment–term approximations $F_i^{[n]} \approx F^{[n]}(t_{ni}, P(\cdot))$ and $\Phi_i^{[n+1]} \approx \Phi^{[n+1]}(t_{ni}, P(\cdot))$ assume the form

$$F_i^{[n]} = g(t_{ni}) + h \sum_{\nu=0}^{n-1} \sum_{l=0}^{\mu_1} b_l k(t_{ni}, t_\nu + \xi_l h, P_\nu(t_\nu + \xi_l h)), \quad i = 1, ..., m \tag{12.2.5}$$

$$\Phi_i^{[n+1]} = h \sum_{l=0}^{\mu_0} w_{il} k(t_{ni}, t_n + d_{il} h, P_n(t_n + d_{il} h)), \quad i = 1, ..., m, \tag{12.2.6}$$

and are obtained by using quadrature formulas of the form

$$(\xi_l, b_l)_{l=1}^{\mu_1}, \quad (d_{il}, w_{il})_{l=1}^{\mu_0}, \ i = 1, ..., m, \tag{12.2.7}$$

where the quadrature nodes ξ_l and d_{il} satisfy $0 \le \xi_1 < ... < \xi_{\mu_1} \le 1$ and $0 \le d_{i1} < ... < d_{i\mu_0} \le 1$, μ_0 and μ_1 are positive integers and w_{il}, b_l are suitable weights.

The discretized multistep collocation method (12.2.1)-(12.2.4) provides a continuous approximation $P(t)$ of the solution $y(t)$ of the integral equation (12.0.1) in $[0, T]$, by considering

$$P(t)|_{(t_n, t_{n+1}]} = P_n(t) \tag{12.2.8}$$

where $P_n(t)$ is given by (12.2.1). We note that usually the polynomial constructed in the collocation methods for VIEs doesn't interpolate the numerical solution in the previous step points, resulting a discontinuous approximation of the solution (12.1.6). In this multistep extension, the collocation polynomial is instead a continuous approximation to the solution, i.e. $u(t) \in S_{m-1}^{(0)}(I_h)$. The discretized multistep collocation method (12.2.1)-(12.2.4) can be regarded as a multistep Runge–Kutta method for VIEs:

$$\begin{cases} Y_i^{[n+1]} = \bar{F}^{[n]}(t_{ni}) + h \sum_{l=1}^{\mu_0} w_{il} k \left(t_n + e_{il} h, \ t_n + d_{il} h, \ \sum_{k=0}^{\mu-1} \gamma_{ilk} y_{n-k} + \sum_{j=1}^{m} \beta_{ilj} Y_j^{[n+1]} \right), \\ y_{n+1} = \sum_{k=0}^{\mu-1} \theta_k y_{n-k} + \sum_{j=1}^{m} \lambda_j Y_i^{[n+1]}, \end{cases} \tag{12.2.9}$$

where

$$\bar{F}^{[n]}(t) = g(t) + h\sum_{\nu=0}^{n-1}\sum_{l=1}^{\mu_1} b_l k\left(t,\ t_\nu + \xi_l h,\ \sum_{k=0}^{\mu-1}\delta_{lk}y_{\nu-k} + \sum_{j=1}^{m}\eta_{lj}Y_j^{[\nu]}\right) \tag{12.2.10}$$

and

$$\begin{aligned} &e_{il} = c_i,\ \gamma_{ilk} = \varphi_k(d_{il}),\ \beta_{ilj} = \psi_j(d_{il}),\\ &\theta_k = \varphi_k(1),\ \lambda_j = \psi_j(1),\\ &\delta_{lk} = \varphi_k(\xi_l),\ \eta_{lj} = \psi_j(\xi_l). \end{aligned}$$

The reason of interest of the multistep collocation methods lies in the fact that they increase the order of convergence of collocation methods without increasing the computational cost, except for the cost due to the starting procedure. As a matter of fact, in advancing from t_n to t_{n+1}, we make use of the approximations y_{n-k}, $k = 0, \ldots, \mu-1$, which have already been evaluated at the previous steps. This permits to increase the order, by maintaining in (12.2.4) the same dimension m of the nonlinear system (12.1.2).

The μ–step m–stage collocation methods have uniform order $m+\mu$, and order of local superconvergence $2m+\mu-1$. The knowledge of the collocation polynomial, which provides a continuous approximation of uniform order of the solution, will allow a cheap variable stepsize implementation. Indeed, when the stepsize changes, the new approximation values can be computed by simply evaluating the collocation polynomial, without running into problems of order reduction, as a consequence of the uniform order.

12.3 Two-step collocation and almost collocation methods

Two-step almost collocation methods for VIEs [75] provide a continuous approximation $P_n(t_n + sh)$, $s \in [0,1]$, to the solution $y(t_n + sh)$ in the interval $[t_n, t_{n+1}]$, which employs the information about the equation on two consecutive steps and suitable sufficiently high order quadrature formulae $F_j^{[n]}$ and $\Phi_j^{[n+1]}$ for the discretization of $F^{[n]}(t_{nj}, P(\cdot))$ and $\Phi^{[n+1]}(t_{nj}, P(\cdot))$ respectively. The approximation $P(t)$ of the solution $y(t)$ of (12.0.1) on $[0,T]$ is then obtained by considering

$$P(t)|_{(t_n,t_{n+1}]} = P_n(t).$$

The method assumes the form

$$\begin{cases} P_n(t_n+sh) = \varphi_0(s)y_{n-1} + \varphi_1(s)y_n + \sum_{j=1}^{m} \chi_j(s)Y_j^{[n+1]} + \sum_{j=1}^{m} \psi_j(s)(F_j^{[n]} + \Phi_j^{[n+1]}) \\ y_{n+1} = P(t_{n+1}), \end{cases} \tag{12.3.1}$$

where the algebraic polynomial $P_n(t_n+sh)$ is expressed as linear combination of the basis functions $\varphi_0(s)$, $\varphi_1(s)$, $\chi_j(s)$ and $\psi_j(s)$, $j = 1, 2, \ldots, m$, which are determined from the continuous order conditions provided in [75]. These conditions arise from the analysis of the local truncation error

$$\begin{aligned} \eta(t_n+sh) &= y(t_n+sh) - \varphi_0(s)y(t_n-h) - \varphi_1(s)y(t_n) \\ &- \sum_{j=1}^{m} \Big(\chi_j(s)y(t_n+(c_j-1)h) + \psi_j(s)y(t_n+c_jh)\Big). \end{aligned} \tag{12.3.2}$$

and are reported in the following result.

Theorem 12.3.1 *Assume that the kernel $k(t,\eta,y)$ and the function $g(t)$ in (12.0.1) are sufficiently smooth. Then the method (12.3.1) has uniform order p, i.e.,*

$$\eta(t_n+sh) = O(h^{p+1}), \quad h \to 0,$$

for $s \in [0,1]$, if the polynomials $\varphi_0(s)$, $\varphi_1(s)$, $\chi_j(s)$ and $\psi_j(s)$, $j = 1, 2, ..., m$ satisfy the system of equations

$$\begin{cases} 1 - \varphi_0(s) - \varphi_1(s) - \sum_{j=1}^{m} \chi_j(s) - \sum_{j=1}^{m} \psi_j(s) = 0, \\ s^k - (-1)^k\varphi_0(s) - \sum_{j=1}^{m} (c_j-1)^k\chi_j(s) - \sum_{j=1}^{m} c_j^k\psi_j(s) = 0, \end{cases} \tag{12.3.3}$$

$s \in [0,1]$, $k = 1, 2, ..., p$.

Two-step collocation methods are obtained by solving the system of order conditions up to the maximum uniform attainable order $p = 2m+1$, and, in this way, all the basis functions are determined as the unique solution of such system: the resulting methods completely reproduce the class of multistep collocation methods presented in the previous section with $\mu = 2$. However, as observed in [75], it is not convenient to impose all the order conditions because it is not possible to achieve high stability properties (e.g. $A-$stability)

without getting rid of some of them. Therefore, *almost* collocation methods are introduced by relaxing a specified number r of order conditions, i.e. by a priori opportunely fixing r basis functions, and determining the remaining ones as the unique solution of the system of order conditions up to $p = 2m+1-r$. Within the class of TSAC methods many $A-$stable methods have been constructed [75].

The quadrature formulae in (12.3.1) are of the form

$$F_j^{[n]} = g(t_{nj}) + h\sum_{\nu=1}^{n}\left(b_0 k(t_{nj}, t_{\nu-1}, y_{\nu-1}) + \sum_{l=1}^{m} b_l k(t_{nj}, t_{\nu-1,l}, Y_l^{[\nu]}) + b_{m+1} k(t_{nj}, t_\nu, y_\nu)\right), \tag{12.3.4}$$

and

$$\Phi_j^{[n+1]} = h\left(w_{j0} k(t_{nj}, t_n, y_n) + \sum_{l=1}^{m} w_{jl} k(t_{nj}, t_{nl}, Y_l^{[n+1]}) + w_{j,m+1} k(t_{nj}, t_{n+1}, y_{n+1})\right), \tag{12.3.5}$$

where $Y_i^{[n]} = P_{n-1}(t_{n-1,i})$ are the stage values and b_0, b_l, b_{m+1}, w_{j0}, w_{jl}, $w_{j,m+1}$ are given weights.

The polynomial $P(t_n + sh)$ is explicitly defined after solving, at each step, the following system of $(m+1)d$ nonlinear equations in the stage values $Y_i^{[n+1]}$ and y_{n+1}, obtained by computing (12.3.1) for $s = c_i$, $i = 1, 2, \ldots, m$, and $s = 1$:

$$\begin{cases} Y_i^{[n+1]} &= \varphi_0(c_i) y_{n-1} + \varphi_1(c_i) y_n + \displaystyle\sum_{j=1}^{m} \chi_j(c_i) Y_j^{[n+1]} + \sum_{j=1}^{m} \psi_j(c_i)\left(F_j^{[n]} + \Phi_j^{[n+1]}\right), \\ y_{n+1} &= \varphi_0(1) y_{n-1} + \varphi_1(1) y_n + \displaystyle\sum_{j=1}^{m} \chi_j(1) Y_j^{[n+1]} + \sum_{j=1}^{m} \psi_j(1)\left(F_j^{[n]} + \Phi_j^{[n+1]}\right), \end{cases} \tag{12.3.6}$$

$n = 1, 2, \ldots, N-1$.

Remark 12.3.1 *Let $e_h(t) := y(t) - P(t)$ be the global error of the TSAC method (12.3.1), and suppose the hypothesis of Theorem 12.3.1 are satisfied up to order p. Then, under suitable hypothesis of sufficient regularity on the kernel k (see [75]),*

$$\|e_h\|_{\infty,[t_0,T]} = O(h^{p^*}), \quad h \to 0.$$

i.e. the method has uniform order of convergence $p^* = \min\{s+1, q, p+1\}$, where s and q are such that:

i. *the starting error is* $\|e_h\|_{\infty,[t_0,t_1]} = O(h^s)$*;*

ii. *the lag-term and increment-term quadrature formulas (12.3.4)-(12.3.5) are of order* $O(h^q)$*.*

12.4 Mixed collocation

In the case of VIEs with periodic highly oscillatory solutions, traditional methods may be inefficient, as they may require the use of a small stepsize in order to follow accurately the oscillations of high frequency. As in the case of ODEs "ad hoc" numerical methods have been constructed, incorporating the a priori knowledge of the behaviour of the solution, in order to use wider stepsizes with respect to classical methods and simultaneously to simulate with high accuracy the oscillations of the solution.

A first work on the numerical solution of VIEs with periodic solution is [23], where numerical methods were constructed by means of mixed interpolation. Recently, mixed collocation methods have been introduced in [25, 26] for VIEs and VIDEs. In particular in [25], mixed collocation methods have been introduced for linear convolution VIEs of the form

$$y(t) = g(t) + \int_{-\infty}^{t} k(t-\tau)y(\tau)d\tau, \quad t \in [0,T], \tag{12.4.1}$$

with

$$y(t) = \psi(t), \quad t \in [-\infty, 0],$$

where $k \in L^1(0,\infty)$, g is a continuous periodic function and ψ is a given bounded and continuous function. The collocation polynomial is taken in the form

$$P_n(t_n + sh) = \sum_{k=0}^{m} B_k(s) Y_k^{[n+1]}$$

where the $B_k(s)$ are combinations of trigonometric functions and algebraic polynomials given in [25]. The numerical method is of the form

$$\begin{cases} Y_i^{[n+1]} = F_i^{[n]} + \Phi_i^{[n+1]}, \\ y_{n+1} = \sum_{k=0}^{m} B_k(1) Y_k^{[n+1]}, \end{cases} \tag{12.4.2}$$

where the lag-term and increment term approximations are given by

$$F_i^{[n]} = g(t_{ni}) + \int_{-\infty}^{0} k(t_{ni} - \tau)\psi(\tau)d\tau + h\sum_{\nu=0}^{n-1}\sum_{l=0}^{m} w_l(1)k(t_{nj} - t_{\nu l})Y_l^{[\nu+1]},$$

$$\Phi_i^{[n+1]} = hc_i \sum_{l=0}^{m} w_l(1)k(t_{ni} - t_n - hc_i c_l)\left(\sum_{k=0}^{m} B_k(c_i c_l)Y_k^{[n+1]}\right),$$

with

$$w_l(s) = \int_0^s B_l(\tau)d\tau.$$

With some suitable choices for collocation parameters such methods accurately integrates systems for which the period of oscillation of the solution is known. In the paper [56] the authors introduce a family linear methods, namely Direct Quadrature (DQ) methods, specially tuned on the specific feature of the problem, based on the exponential fitting [132, 135], which is extremely flexible when periodic functions are treated. Such formulae are based on a three-term quadrature formula, that is of the same form as the usual Simpson rule, but specially tuned on integrands of the form $k(s)y(s)$ where k and y are of type

$$k(t) = e^{\alpha t}, \qquad y(t) = a + b\cos(\omega t) + c\sin(\omega t), \tag{12.4.3}$$

where $\alpha, \omega, a, b, c \in I\!R$. The coefficients of the new quadrature rule depend on the parameters of the integrand, i.e. α and ω. It has been shown as the use of exponentially fitted based three-point quadrature rules produces a definite improvement in the accuracy when compared with the results from the classical Simpson rule, and that the magnitude of the gain depends on how good is the knowledge of the true frequencies. The results also indicate that, as a rule, if the input accuracy is up to 10 percent, then the accuracy gain in the output is substantial.

Chapter 13

Diagonally implicit two-step almost collocation methods for Volterra Integral Equations

This chapter concerns with the construction of both efficient and highly stable numerical methods for Volterra Integral Equations (VIEs) of the form (12.0.1). It is well known that the best stability properties are reached by implicit numerical methods, with the disadvantage that they lead to nonlinear systems of equations to be solved at each time step. To this cost we have to add the cost arising from the computation of the lag term (containing the history of the phenomenon), which can be computed by means of fast methods developed in the literature for convolution kernels [54, 74, 183]. As regards the task of reducing the computational cost associated to the solution of the above nonlinear systems, a widespread strategy in the context of Ordinary Differential Equations (ODEs), consists in making the coefficient matrix have a structured shape. This strategy, in the field of Runge–Kutta methods for ODEs, leads to the raise of the famous classes of Diagonally Implicit and Singly Diagonally Implicit Runge-Kutta methods (DIRK and SDIRK), see [42, 122] and bibliography therein contained. Moreover, in the field of collocation-based methods for ODEs, an analogous strategy has been applied, obtaining a subclass of two-step Runge–Kutta methods (see [140]) having structured coefficient matrix [98].

In this chapter we will adopt this strategy in the context of the numerical solution of VIEs (12.0.1) and, in particular, we will derive numerical meth-

ods involving nonlinear systems with lower triangular or diagonal coefficient matrices. The derived formulae belong to the class of Two-Step Almost Collocation (TSAC) methods, which have been introduced in [75] and described in Section 12.3, in order to obtain high order and highly stable continuous methods for the problem (12.0.1). In fact, to integrate a system of d integral equations of the form (12.0.1), a collocation method generally requires the solution of md simultaneous nonlinear equations of the type (12.3.6) at each time step. A lower triangular matrix allows to solve the equations in m successive stages, with only a d-dimensional system to be solved at each stage. Moreover, if all the elements on the diagonal are equal, in solving the nonlinear systems by means of Newton-type iterations, one may hope to use repeatedly the stored LU factorization of the Jacobian. If the structure is diagonal, the problem reduces to the solution of m independent systems of dimension d, and can therefore be solved in a parallel environment.

13.1 Two-step diagonally implicit almost collocation methods

This section is devoted to the construction of high order methods belonging to the class (12.3.1) such that the coefficient matrix of the nonlinear system (12.3.6) has a structured shape, leading to the solution of nonlinear systems of lower dimension d. First of all we consider $w_{j,m+1} = 0$, $j = 1, \ldots, m$, in such a way that (12.3.6) becomes a nonlinear system of dimension md only depending on the stage values $Y_i^{[n+1]}$, $i = 1, \ldots, m$, and assumes the following form

$$\begin{cases} Y_i^{[n+1]} - h \sum_{j=1}^{m} \sum_{l=1}^{m} \psi_j(c_i) w_{jl} k(t_{nj}, t_{nl}, Y_l^{[n+1]}) = B_i^{[n]}, \\ y_{n+1} = P_n(t_{n+1}), \end{cases} \tag{13.1.1}$$

where

$$B_i^{[n]} = \varphi_0(c_i) y_{n-1} + \varphi_1(c_i) y_n + \sum_{j=1}^{m} \chi_j(c_i) Y_j^{[n]} + \sum_{j=1}^{m} \psi_j(c_i) F_j^{[n]} + h \sum_{j=1}^{m} \psi_j(c_i) w_{j0} k(t_{nj}, t_n, y_n). \tag{13.1.2}$$

By defining

$$Y^{[n+1]} = \left[Y_1^{[n+1]}, Y_2^{[n+1]}, \ldots, Y_m^{[n+1]}\right]^T, \quad B^{[n]} = \left[B_1^{[n]}, B_2^{[n]}, \ldots, B_m^{[n]}\right]^T,$$

$$\Psi = \big(\psi_j(c_i)\big)_{i,j=1}^m, \quad W = \big(w_{jl}\big)_{j,l=1}^m, \quad K(t_{nc}, t_{nc}, Y^{[n+1]}) = \left(K(t_{ni}, t_{nj}, Y_j^{[n+1]})\right)_{i,j=1}^m,$$

the nonlinear system in (13.1.1) takes the form

$$Y^{[n+1]} - h\Psi\big(W \cdot K(t_{nc}, t_{nc}, Y^{[n+1]})\big) = B^{[n]}, \tag{13.1.3}$$

where $\cdot$ denotes the usual Hadamard product. The tensor form (13.1.3) clearly shows as the matrices which determine the structure of the nonlinear system (13.1.1) are Ψ and W. Therefore, in the following subsection we will describe how to choose the basis functions $\psi_j(s)$ and how to modify the quadrature formula (12.3.5) in order to obtain lower triangular or diagonal structures.

13.1.1 Determination of the basis functions $\psi_j(s)$

In order to achieve a lower triangular or diagonal structure for the matrix Ψ, the basis functions $\psi_j(s)$ must satisfy

$$\psi_j(c_i) = 0, \text{for } j > i \text{ or } j \neq i \text{ respectively}, \tag{13.1.4}$$

i.e. $\psi_j(s)$ assumes the form

$$\psi_j(s) = \prod_{k=1}^{j-1}(s - c_k)\bar{\omega}_j(s), \ j = 2, \ldots, m \tag{13.1.5}$$

or

$$\psi_j(s) = \prod_{\substack{k=1 \\ k\neq j}}^{m}(s - c_k)\tilde{\omega}_j(s), \ j = 1, \ldots, m \tag{13.1.6}$$

respectively, where $\bar{\omega}_j(s)$ is a polynomial of degree $p - j + 1$ and $\tilde{\omega}_j(s)$ is a polynomial of degree $p - m + 1$.

Imposing (13.1.5), the remaining $m + 3$ basis functions can be computed by solving the system of order conditions (12.3.3) and, as a consequence, the

maximum attainable uniform order which can be achieved by the corresponding TSAC methods is $m+2$. On the other hand, by imposing (13.1.6), the corresponding TSAC methods would have uniform order at most equal to $m+1$. However, we can follow a different strategy in order to obtain higher order methods. The idea is to impose the conditions (13.1.4) on less than m basis functions, e.g. one or two of them (generally $\varphi_0(s)$ and $\varphi_1(s)$), in such a way that the maximum attainable order is $p = 2m + r - 1$, with $r = 1, 2$.

Let us define the following sets of interpolation and collocation conditions (see [75]):

- *interpolation conditions in 0*

$$\varphi_0(0) = 0,\ \varphi_1(0) = 1,\ \chi_j(0) = 0,\ \psi_j(0) = 0,\ \forall j, \tag{13.1.7}$$

- *interpolation conditions in -1*

$$\varphi_0(-1) = 1,\ \varphi_1(-1) = 0,\ \chi_j(-1) = 0,\ \psi_j(-1) = 0,\ \forall j, \tag{13.1.8}$$

- *interpolation conditions in $c_i - 1$, $i \in \{1, \dots, m\}$*

$$\varphi_0(c_i - 1) = 0,\ \varphi_1(c_i - 1) = 0,\ \chi_j(c_i - 1) = \delta_{ij},\ \psi_j(c_i - 1) = 0,\ \forall j. \tag{13.1.9}$$

- *collocation conditions in c_i, $i \in \{1, \dots, m\}$*

$$\varphi_0(c_i) = 0,\ \varphi_1(c_i) = 0,\ \chi_j(c_i) = 0,\ \psi_j(c_i) = \delta_{ij},\ \forall j, \tag{13.1.10}$$

Remark 13.1.1 *The name of conditions (13.1.7)-(13.1.8)-(13.1.9)-(13.1.10) arises from the fact that they respectively ensure $P_n(t_n) = y_n$, $P_n(t_{n-1}) = y_{n-1}$, $P_n(t_{n-1,i}) = Y_i^{[n]} = P_{n-1}(t_{n-1,i})$, $P_n(t_{ni}) = F_i^{[n]} + \Phi_i^{[n+1]}$. In particular, the last one means that the collocation polynomial exactly satisfies the VIE (12.0.1) in the collocation point t_{ni}, except from the error associated to the quadrature formulas (12.3.4) and (12.3.5).*

Whatever condition from the sets (13.1.7)-(13.1.8)-(13.1.9)-(13.1.10) is inherited via order conditions, i.e. if we impose that some basis functions satisfy certain interpolation/collocation conditions from the sets (13.1.7)-(13.1.8)-(13.1.9)-(13.1.10) and derive all the other basis functions by solving the system of order conditions, the same interpolation/collocation conditions

are also satisfied by the computed basis functions (and then the corresponding relation in Remark 13.1.1 is satisfied by the collocation polynomial), as proved in the following theorem.

Theorem 13.1.1 *Let us define* $\xi_1 = -1$, $\xi_2 = 0$, $\xi_{2+j} = c_j - 1$, $\xi_{m+2+j} = c_j$, $j = 1, \ldots, m$ *and* $\Gamma_1(s) = \varphi_0(s)$, $\Gamma_1(s) = \varphi_1(s)$, $\Gamma_{2+j}(s) = \chi_j(s)$, $\Gamma_{m+2+j}(s) = \psi_j(s)$, $j = 1, \ldots, m$ *and let* $i \in \{1, 2, \ldots, 2m+2\}$ *be a fixed integer. Then, supposing* $\xi_i \neq \xi_j$, $i \neq j$,

i. If $\Gamma_i(\xi_i) = 1$, *then* $\Gamma_j(\xi_i) = 0$ *for all* $j \neq i$;

ii. If $\Gamma_i(\xi_\ell) = 0$ *with* $\ell \neq i$, *then* $\Gamma_j(\xi_\ell) = \delta_{j\ell}$ *for* $j \neq i$.

Proof: The system of order conditions (12.3.3) can be rewritten in terms of ξ_j and $\Gamma_j(s)$ as

$$s^k - \sum_{j=1}^{2m+2} \xi_j^k \Gamma_j(s) = 0, \quad k = 0, 1, \ldots, p, \tag{13.1.11}$$

where we assume $\xi_2^0 = 1$. We first prove the part *i.* of the thesis: for this purpose, we evaluate (13.1.11) in $s = \xi_i$ and, as a consequence, using the assumption *i.* leads to the following linear system

$$\sum_{\substack{j=1 \\ j \neq i}}^{2m+2} \xi_j^k \Gamma_j(\xi_i) = 0, \quad k = 0, 1, \ldots, p,$$

which is a Vandermonde type linear system whose unique solution is $\Gamma_j(\xi_i) = 0$ for all $j \neq i$. In analogous way, by evaluating (13.1.11) in $s = \xi_\ell$ and taking into account the assumption *ii.*, we obtain the Vandermonde type linear system

$$\xi_\ell^k - \sum_{\substack{j=1 \\ j \neq i}}^{2m+2} \xi_j^k \Gamma_j(\xi_\ell) = 0, \quad k = 0, 1, \ldots, p,$$

whose unique solution is $\Gamma_j(\xi_\ell) = \delta_{j\ell}$ for $j \neq i$. □

Remark 13.1.2 *In some examples presented in Section 5, we will fix one or both of the polynomials $\varphi_0(s)$ and $\varphi_1(s)$ in the following way*

$$\begin{aligned}\varphi_0(s) &= \prod_{k=1}^{m}(s-c_k)(\alpha_0+\alpha_1 s+\ldots+\alpha_{p-m}s^{p-m}),\\ \varphi_1(s) &= \prod_{k=1}^{m}(s-c_k)(\beta_0+\beta_1 s+\ldots+\beta_{p-m}s^{p-m}),\end{aligned} \tag{13.1.12}$$

where α_j and β_j, $j=0,1,\ldots,p-m$, are free parameters. Then, as a consequence of Theorem 13.1.1, all the conditions (13.1.10) are satisfied and Ψ reduces to the identity matrix of dimension m.

13.1.2 Approximation of the increment term: modified quadrature formula

We observe that the quadrature formulae (12.3.4) and (12.3.5) can be obtained by applying the quadrature formulae

$$\int_0^1 f(s)ds \approx b_0 f(0)+\sum_{l=1}^{m} b_l f(c_l)+b_{m+1}f(1), \tag{13.1.13}$$

$$\int_0^{c_j} f(s)ds \approx w_{j0} f(0)+\sum_{l=1}^{m} w_{jl} f(c_l)+w_{j,m+1}f(1), \tag{13.1.14}$$

for the approximation of the integrals appearing in the right hand side of

$$\begin{aligned}F^{[n]}(t_{nj},P(\cdot)) &= g(t_{nj})+h\sum_{\nu=1}^{n}\int_0^1 k(t_{nj},t_{\nu-1}+sh,P_{\nu-1}(t_{\nu-1}+sh))ds,\\ \Phi^{[n+1]}(t_{nj},P(\cdot)) &= h\int_0^{c_j} k(t_{nj},t_n+sh,P_n(t_n+sh))ds.\end{aligned} \tag{13.1.15}$$

We aim to derive a suitable modification of the quadrature formula (13.1.14) in such a way that the matrix W is lower triangular or diagonal and, with the purpose to preserve the order, we make use of some additional quadrature

nodes, i.e. we consider quadrature formulae of the form

$$\int_0^{c_j} f(s)ds \approx w_{j0}f(0) + \sum_{l=1}^{m} \tilde{w}_{jl}f(c_l - 1) + \sum_{l=1}^{j} w_{jl}f(c_l), \tag{13.1.16}$$

where, in case of triangular structure, $\tilde{w}_{jl} = 0$, $l = 1, \ldots, j$ while, in case of diagonal structure, $\tilde{w}_{j1} = 0$ and $w_{jl} = 0$, $l = 1, \ldots, j-1$.

With the purpose of achieving the desired order (see Remark 12.3.1), quadrature formulae of the form (13.1.13) and (13.1.16) can be constructed by taking into account that the order of the corresponding lag term and increment term quadrature formulae is at least $O(h^q)$, if they are interpolatory quadrature formulae on q and $q-1$ nodes respectively [27].

Remark 13.1.3 *The quadrature formulae (13.1.13) and (13.1.16) can be further generalized if we need higher order by considering*

$$\begin{aligned} \int_0^1 f(s)ds &\approx \sum_{l=0}^{\mu_1} b_l f(d_l), \\ \int_0^{c_j} f(s)ds &\approx \sum_{l=0}^{\mu_0} \tilde{w}_{jl}f(d_{jl} - 1) + \sum_{l=1}^{j} w_{jl}f(c_l), \end{aligned}$$

where μ_0 and μ_1 depend on the desired order and (13.1.13) and (13.1.16) are special cases, obtained by setting $\mu_1 = m+1$, $d_0 = 0$, $d_l = c_l, l = 1, 2, \ldots, m$, $d_{m+1} = 1$, $\mu_0 = m$, $d_{j0} = 0$, $\tilde{w}_{j0} = 0$, $d_{jl} = c_l, j, l = 1, 2, \ldots, m$.

13.1.3 Form of the diagonally implicit TSAC methods

As a consequence of the choices reported in the previous subsections, we obtain what follows. If Ψ and W are lower triangular,

$$\Psi = \begin{bmatrix} \psi_{11} & & & \\ \psi_{21} & \psi_{22} & & \\ \vdots & & \ddots & \\ \psi_{m1} & \psi_{m2} & \ldots & \psi_{mm} \end{bmatrix}, \quad W = \begin{bmatrix} w_{11} & & & \\ w_{21} & w_{22} & & \\ \vdots & & \ddots & \\ w_{m1} & w_{m2} & \ldots & w_{mm} \end{bmatrix}, \tag{13.1.17}$$

the resulting method assumes the form

$$\begin{cases} Y_i^{[n+1]} - h\psi_i(c_i)w_{ii}k(t_{ni}, t_{ni}, Y_i^{[n+1]}) = B_i^{[n]} + \tilde{B}_i^{[n]} + h\sum_{l=1}^{i-1}\sum_{j=l}^{i}\psi_j(c_i)w_{jl}k(t_{nj}, t_{nl}, Y_l^{[n+1]}), \\ y_{n+1} = \varphi_0(1)y_{n-1} + \varphi_1(1)y_n + \sum_{j=1}^{m}\chi_j(1)Y_j^{[n]} + \sum_{j=1}^{m}\psi_j(1)\Big(F_j^{[n]} + \Phi_j^{[n+1]}\Big), \end{cases} \tag{13.1.18}$$

where $B_i^{[n]}$ is given by (13.1.2),

$$\tilde{B}_i^{[n]} = h\sum_{j=1}^{i}\sum_{l=1}^{m}\psi_j(c_i)\tilde{w}_{jl}k(t_{nj}, t_{n-1,l}Y_l^{[n]}), \tag{13.1.19}$$

and $F_j^{[n]}$, $\Phi_j^{[n+1]}$ are approximations of (13.1.15) by means of the quadrature formulae (13.1.13) and (13.1.16). The solution of the system (13.1.18) of dimension md can be obtained by solving m successive nonlinear systems of dimension d. Coherently with the case of ODEs, we denote the corresponding methods as diagonally implicit TSAC methods (DITSAC). If Ψ and W are lower triangular and, in addition, their product ΨW is one-point spectrum, i.e.

$$\Psi W = \begin{bmatrix} \lambda & & & \\ \mu_{21} & \lambda & & \\ \vdots & & \ddots & \\ \mu_{m1} & \mu_{m2} & \dots & \lambda \end{bmatrix}, \tag{13.1.20}$$

where $\lambda = \psi_i(c_i)w_{ii}$, $i = 1, 2, \dots, m$ then, in order to solve the system (13.1.18) by Newton-type iterations, we can repeatedly use the stored LU-factorization of the coefficient matrix

$$I - h\lambda\frac{\partial k}{\partial y}.$$

The related TSAC methods are then called singly diagonally implicit TSAC methods (SDITSAC).

If, in particular, Ψ and W are diagonal

$$\Psi = \begin{bmatrix} \psi_{11} & & & \\ & \psi_{22} & & \\ & & \ddots & \\ & & & \psi_{mm} \end{bmatrix}, \quad W = \begin{bmatrix} w_{11} & & & \\ & w_{22} & & \\ & & \ddots & \\ & & & w_{mm} \end{bmatrix}, \tag{13.1.21}$$

then the method (13.1.18) takes the form

$$Y_i^{[n+1]} - h\psi_i(c_i)w_{ii}k(t_{ni}, t_{ni}, Y_i^{[n+1]}) = B_i^{[n]} + \tilde{B}_i^{[n]}, \tag{13.1.22}$$

where $B_i^{[n]}$ and $\tilde{B}_i^{[n]}$ are given by (13.1.2) and (13.1.19), respectively. The nonlinear system (13.1.22) is then equivalent to m nonlinear systems of dimension d, which can be efficiently solved in a parallel environment. The corresponding methods are denominated diagonal TSAC methods (DTSAC). If ΨW is also one point spectrum, i.e.

$$\Psi W = \lambda I, \tag{13.1.23}$$

with $\lambda = \psi_i(c_i)w_{ii}$, it can be efficiently treated by means of Newton-type iterations, as in the case of SDITSAC methods: we denote these methods as singly diagonal TSAC methods (SDTSAC).

13.2 Linear stability analysis

We now focus our attention on the linear stability properties of TSAC methods (13.1.18) with respect to the basic test equation

$$y(t) = 1 + \lambda \int_0^t y(\tau)d\tau, \quad t \geq 0, \quad \mathrm{Re}(\lambda) \leq 0, \tag{13.2.1}$$

usually employed in the literature for the stability analysis of numerical methods for VIEs (see [27, 75, 77]). The following result holds.

Theorem 13.2.1 *The stability matrix associated to the two-step collocation method (13.1.18) takes the form*

$$R(z) = Q^{-1}(z)M(z), \tag{13.2.2}$$

where

$$Q(z) = \begin{bmatrix} 1 & -z\psi^T(1)W & -\psi^T(1) & 0 \\ 0 & I - z\Psi W & -\Psi & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{13.2.3}$$

is an invertible matrix for $z < \dfrac{1}{\|\Psi W\|}$ *(for some matrix norm) and*

$$M(z) = \begin{bmatrix} \varphi_1(1) + z\psi^T(1)w_0 & \chi^T(1) + z\psi^T(1)\tilde{W} & 0 & \varphi_0(1) \\ \varphi_1(c) + z\Psi w_0 & A + z\Psi\tilde{W} & 0 & \varphi_0(c) \\ zb_{m+1}u & zub^T & I & zb_0u \\ 1 & 0 & 0 & 0 \end{bmatrix}. \tag{13.2.4}$$

Proof: The method (13.1.18) applied to the test problem (13.2.1) assumes the form

$$\begin{aligned} y_{n+1} &= \varphi_0(1)y_{n-1} + (\varphi_1(1) + z\psi^T(1)w_0)y_n + (\chi^T(1) + z\psi^T(1)\tilde{W})Y^{[n]} \\ &\quad +\psi^T(1)F^{[n]} + z\psi^T(1)WY^{[n+1]}, \\ Y^{[n+1]} &= \varphi_0(c)y_{n-1} + (\varphi_1(c) + z\Psi w_0)y_n + (A + z\Psi\tilde{W})Y^{[n]} + \Psi F^{[n]} + z\Psi WY^{[n+1]}, \end{aligned} \tag{13.2.5}$$

where we define the column vectors $\psi(1) = (\psi_j(1))_{j=1}^m$, $\chi(1) = (\chi_j(1))_{j=1}^m$, $w_0 = (\tilde{w}_{j0})_{j=1}^m$, $\varphi_0(c) = (\varphi_0(c_j))_{j=1}^m$ and $\varphi_1(c) = (\varphi_1(c_j))_{j=1}^m$ and the matrix $\tilde{W} = (\tilde{w}_{ij})_{i,j=1}^m$. The lag term satisfies the following recurrence relation

$$F^{[n]} = F^{[n-1]} + zb_{m+1}uy_n + zub^TY^{[n]} + zb_0uy_{n-1}, \tag{13.2.6}$$

with $b = [b_1, b_2, \ldots, b_m]^T$ and $u = [1, \ldots, 1]^T \in \mathbb{R}^m$. By defining

$$v_n = \left[y_n, Y^{[n]}, F^{[n-1]}, y_{n-1}\right]^T,$$

from (13.2.5) and (13.2.6), we obtain the following recurrence relation

$$Q(z)v_{n+1} = M(z)v_n, \tag{13.2.7}$$

where $Q(z)$ and $M(z)$ take the form (13.2.3) and (13.2.4), respectively. The proof will be completed by showing the invertibility of the matrix $Q(z)$, by means of some algebraic tools, based on the Schür complement. It is well known (see [13]) that, given a block matrix of the type

$$\left[\begin{array}{c|c} Q_{11} & Q_{12} \\ \hline 0 & Q_{22} \end{array}\right]$$

with invertible blocks Q_{11} and Q_{22}, the inverse assumes the following form:

$$\left[\begin{array}{c|c} Q_{11} & Q_{12} \\ \hline 0 & Q_{22} \end{array}\right]^{-1} = \left[\begin{array}{c|c} Q_{11}^{-1} & -Q_{11}^{-1}Q_{12}Q_{22}^{-1} \\ \hline 0 & Q_{22}^{-1} \end{array}\right].$$

According to this result, the invertibility of the matrix $Q(z)$ follows from the invertibility of the matrix $I - z\Psi W$ thus, if $z\|\Psi W\| < 1$ for some matrix norm (see [13], p. 492), the inverse of $Q(z)$ can be computed as

$$Q^{-1}(z) = \begin{bmatrix} 1 & z\psi^T(1)W(I-z\Psi W)^{-1} & -\psi^T(1)(I+zW(I-z\Psi W)^{-1}) & 0 \\ 0 & (I-z\Psi W)^{-1} & -(I-z\Psi W)^{-1}\Psi & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{13.2.8}$$

□

The proof of Theorem 13.2.1 also provides the tools for an efficient inversion of the matrix $Q(z)$. In fact, using the Schür complement, we have reduced the problem of the inversion of $Q(z)$, i.e. a matrix of dimension $2m+2$, to the inversion of the lower triangular matrix $I - z\Psi W$, of dimension m, as results from (13.2.8). This allows an efficient computation of the inverse of $Q(z)$.

13.3 Derivation of A-stable (S)DITSAC and (S)DTSAC methods

The strategy we carry out in the construction of A-stable methods can be summarized as follows.

First of all we set the quadrature formulae (13.1.13) and (13.1.16) in such a way to enforce the desired structure on the matrix W and the order p of convergence. Moreover, we fix the polynomial $\varphi_0(s)$ and, eventually, also $\varphi_1(s)$, satisfying some of the interpolation/collocation conditions (13.1.7)-(13.1.8)-(13.1.9)-(13.1.10), e.g. the ones in (13.1.12). As a consequence some free parameters are available to be spent in order to enforce the desired structure on the matrix Ψ and to achieve A-stability. We next derive the remaining basis functions by solving the system of order conditions (12.3.3) up to p: as stated by Theorem 13.1.1, the same fixed interpolation/collocation conditions are inherited by the other basis functions. We next compute the stability polynomial $p(\omega, z)$ of the obtained methods, i.e. the characteristic polynomial of the stability matrix (13.2.2), which depends on the matrices $Q(z)$ in (13.2.3) and $M(z)$ in (13.2.4). In particular, in the construction process it is useful to consider the expression (13.2.8) of the inverse of the matrix $Q(z)$, which provides considerable simplifications in the computations. We next analyze

the stability properties of the polynomial $p(\omega, z)$: in particular, we aim to derive methods which are A-stable, i.e. the roots $\omega_1, \omega_2, \ldots, \omega_{2m+2}$ of the polynomial $p(\omega, z)$ lie in the unit circle, for all $z \in \mathbb{C}$ such that $\mathrm{Re}(z) \leq 0$. We investigate A-stability using the Schur criterion [184], similarly as it has already been done in [75, 92, 94, 95, 93, 138].

13.3.1 Examples of methods with $m = 2$ with Ψ and W lower triangular

We first show the construction of highly stable two-stage DITSAC methods (13.1.18), i.e. we require that the matrices Ψ and W are lower triangular. As a first attempt, we have derived and analyzed the stability properties of (13.1.18) with $m = 2$ and order $p = 2m + 1 = 5$, and discovered that no A-stable methods within this class exist. Therefore, we next relax one order condition ($r = 1$), and consider DITSAC methods (13.1.18) with $m = 2$ and order $p = 2m = 4$. We compute the weights of the quadrature formulae (13.1.13) and (13.1.16) according to the desired order $p = 4$, obtaining

$$
\begin{aligned}
b_0 &= -\frac{-6c_2c_1 + 2c_1 + 2c_2 - 1}{12c_1c_2}, \quad b = \left[\begin{array}{cc} -\frac{1-2c_2}{12(c_1-1)c_1(c_1-c_2)} & \frac{2c_1-1}{12(c_2-1)c_2(c_2-c_1)} \end{array} \right]^T, \\
b_3 &= -\frac{-6c_2c_1 + 4c_1 + 4c_2 - 3}{12(c_1-1)(c_2-1)}, \quad w_0 = \left[-\frac{c_1(c_1 - 3c_2 + 3)}{6(c_2-1)} - \frac{c_2^2 - 3c_1c_2}{6c_1} \right], \\
W &= \left[\begin{array}{cc} \frac{c_1(2c_1-3c_2+3)}{6(c_1-c_2+1)} & 0 \\ -\frac{c_2^3}{6c_1(c_1-c_2)} & -\frac{2c_2^2-3c_1c_2}{6(c_1-c_2)} \end{array} \right], \quad \tilde{W} = \left[\begin{array}{cc} 0 & \frac{c_1^3}{6(c_1-c_2+1)(c_2-1)} \\ 0 & 0 \end{array} \right].
\end{aligned}
$$

As in Remark 13.1.2, we assume

$$\varphi_0(s) = s(s - c_1)(s - c_2)(\alpha_0 + \alpha_1 s)$$

and, as a consequence, the matrix $\Psi = I$. Imposing such a factorization on the polynomial $\varphi_0(s)$ implies that it satisfies the interpolation condition $\varphi_0(0) = 0$ and the collocation conditions $\varphi_0(c_1) = \varphi_0(c_2) = 0$. We impose the condition $\alpha_0 = -\alpha_1$, in order to derive methods which do not depend on y_{n-1}: this choice, as also in the case of two-step Runge–Kutta methods for ODEs, is particularly suitable in order to improve the stability properties of the resulting methods (compare with [138, 140]). We next determine the remaining basis functions $\varphi_1(s), \chi_1(s), \chi_2(s), \psi_1(s), \psi_2(s)$ by imposing the

system of order conditions (12.3.3), which result to be

$$\begin{aligned}
\varphi_1(s) &= \frac{(c_1-s)(s-c_2)\big(c_2(c_2+1)\alpha_1(s-1)sc_1^2+(s+c_2((c_2+1)\alpha_1(s-1)s-1)+1)c_1+(c_2-s-1)(s+1)\big)}{(c_1-1)c_1(c_2-1)c_2},\\
\chi_1(s) &= \frac{(c_1-s)(c_2-s)(c_2((c_1+1)(c_2+1)\alpha_1(s-1)+1)-s-1)s}{(c_1-1)(c_1-c_2-1)(c_1-c_2)},\\
\chi_2(s) &= \frac{(c_1-s)(c_2-s)(c_1((c_1+1)(c_2+1)\alpha_1(s-1)+1)-s-1)s}{(c_1-c_2)(c_1-c_2+1)(c_2-1)},\\
\psi_1(s) &= \frac{(c_2-s)s\big(-c_2(c_2+1)\alpha_1(s-1)c_1^2+(s+c_2((c_2+1)\alpha_1(s-1)s-1)+1)c_1+(c_2-s-1)(s+1)\big)}{c_1(c_1-c_2)(c_1-c_2+1)},\\
\psi_2(s) &= \frac{(c_1-s)s\big(-c_2\alpha_1(c_2-s)(s-1)c_1^2+(c_2(-\alpha_1(c_2-s)(s-1)-1)+s+1)c_1+(c_2-s-1)(s+1)\big)}{(c_1-c_2-1)(c_1-c_2)c_2}.
\end{aligned}$$

The determined quadrature weights and basis functions now depend on the parameters α_1, c_1 and c_2, which can be regarded as degrees of freedom in order to enforce strong stability properties for the corresponding methods, such as A-stability. We also observe that, in force of Theorem 13.1.1, the interpolation/collocation conditions imposed on $\varphi_0(s)$ are automatically inherited by all the other basis functions and, *a fortiori*, on the whole collocation polynomial (12.3.1).

We next derive the stability polynomial $p(\omega, z)$ of order $2m + 2 = 6$ with respect to the variable ω. In force of the choices we have made, it takes the form

$$p(\omega, z) = \omega(p_0(z)+p_1(z)\omega+p_2(z)\omega^2+p_3(z)\omega^3+p_4(z)\omega^4+p_5(z)\omega^5), \quad (13.3.1)$$

where $p_j(z)$, $j = 0, 1, \ldots, 5$, are rational functions with respect to z, which do not depend on the value of the parameter α_1, but only the abscissae c_1 and c_2. In order to investigate on the stability properties of the polynomial (13.3.1), it is sufficient to consider the polynomial

$$\tilde{p}(\omega, z) = p_0(z) + p_1(z)\omega + p_2(z)\omega^2 + p_3(z)\omega^3 + p_4(z)\omega^4 + p_5(z)\omega^5,$$

of degree 5 with respect to ω. We apply the Schür criterion on the polynomial $\tilde{p}(\omega, z)$, in order to determine the values of the free parameters c_1 and c_2 corresponding to A-stable methods. The result of this analysis is reported in Figure 13.1.

We next derive A-stable two-stage SDITSAC methods within the class (13.1.18). In this case, by using the Schür criterion, we did not find A-stable methods with $m = 2$ and $p = 4$ and, therefore, we focus our attention on methods with $m = 2$ and $p = 3$, by relaxing two order conditions. We determine the weights of the quadrature formulae (13.1.13) and (13.1.16)

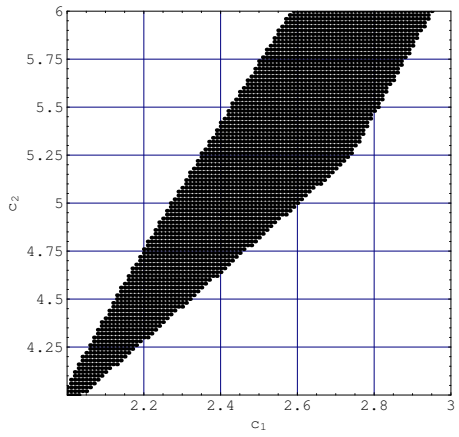

Figure 13.1: Region of A-stability in the parameter space (c_1, c_2) for DITSAC methods (13.1.18), with $m = 2$ and $p = 4$, for any value of the parameter α_1

corresponding to the lower triangular case, obtaining

$$b_0 = -\frac{1-3c_2}{6c_2}, \quad b = \begin{bmatrix} 0 & -\frac{1}{6(c_2-1)c_2} \end{bmatrix}^T, \quad b_3 = -\frac{2-3c_2}{6(c_2-1)},$$
$$w_0 = \begin{bmatrix} \frac{c_1}{2} & \frac{c_2}{2} \end{bmatrix}^T, \quad W = \begin{bmatrix} \frac{c_1}{2} & 0 \\ 0 & \frac{c_2}{2} \end{bmatrix}, \quad \tilde{W} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

We next impose

$$\varphi_0(s) = s(\alpha_0 + \alpha_1 s)(s - \alpha_2),$$
$$\varphi_1(s) = s(\beta_0 + \beta_1 s)(s - \beta_2)$$

and, as a consequence, the interpolation condition $\varphi_0(0) = \varphi_1(0) = 0$ holds. In addition, we also set $\alpha_2 = 1$ in order to enforce the independency on y_{n-1}. We next determine the remaining basis functions $\chi_1(s), \chi_2(s), \psi_1(s), \psi_2(s)$ by imposing the system of order conditions (12.3.3) and which inherit the interpolation condition in 0. The computed basis functions and quadrature weights now depend on the free parameters $\alpha_0, \alpha_1, \beta_0, \beta_1, \beta_2, c_1$ and c_2. First of all, we spend α_0 and β_0 in order to enforce ΨW being lower triangular and one point spectrum. We next enforce some further simplifying assumptions

on the basis functions, using the values of α_1, α_2 and β_1, obtaining a three-parameter family of methods to be investigated. We next derive the stability polynomial $p(\omega, z)$ of order $2m+2=6$ with respect to the variable ω, which assumes the same form (13.3.1), where now $p_j(z)$, $j = 0, 1, \ldots, 5$, depend on the free parameters c_1, c_2 and β_2. As in the previous case, we focus our attention on a polynomial $\tilde{p}(\omega, z)$ of degree 5 with respect to ω. We apply the Schür criterion on $\tilde{p}(\omega, z)$, in order to determine the values of the free parameters c_1, c_2 and β_2 achieving A-stability. The results are shown in Figure 13.2.

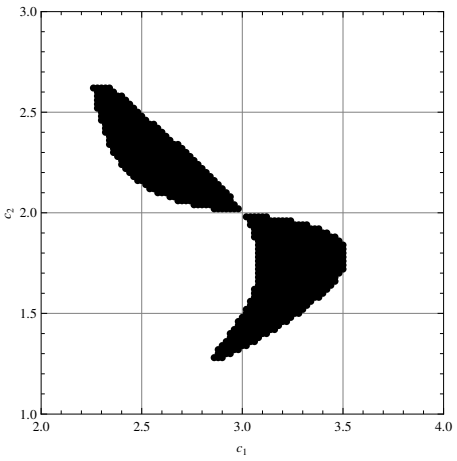

Figure 13.2: Region of A-stability in the parameter space (c_1, c_2) for SDITSAC methods (13.1.18), with $m = 2$, $p = 3$ and $\beta_2 = \frac{1}{4}$

13.3.2 Examples of methods with $m = 2$ with Ψ and W diagonal

We now present the construction of highly stable two-stage DTSAC methods (13.1.22), i.e. we require that the matrices Ψ and W are diagonal. We first observe that, among the examples of A-stable methods provided in [75], the one reported in Figure 5 belongs to the class of DTSAC methods with $m = 2$ and $p = 3$. In this chapter, we present examples of two-stage A-stable SDTSAC methods, requiring that the matrix ΨW is diagonal and one-point spectrum. First of all, we did not find A-stable SDTSAC methods

with $m = 2$ and $p = 4, 5$ exist and, therefore, we relax two order conditions ($r = 2$), and consider SDTSAC methods (13.1.22) with $m = 2$ and order $p = 3$. We compute the weights of the quadrature formulae (13.1.13) and (13.1.16) corresponding to the diagonal case, obtaining

$$\begin{aligned} b_0 &= -\frac{1-3c_2}{6c_2}, \quad b = \begin{bmatrix} 0 & -\dfrac{1}{6(c_2-1)c_2} \end{bmatrix}^T, \quad b_3 = -\frac{2-3c_2}{6(c_2-1)}, \\ w_0 &= \begin{bmatrix} \frac{c_1}{2} & \frac{c_2}{2} \end{bmatrix}^T, \quad W = \begin{bmatrix} \frac{c_1}{2} & 0 \\ 0 & \frac{c_2}{2} \end{bmatrix}, \quad \tilde{W} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

We next impose

$$\begin{aligned} \varphi_0(s) &= s(\alpha_0 + \alpha_1 s)(s - \alpha_2), \\ \varphi_1(s) &= s(\beta_0 + \beta_1 s)(s - \beta_2) \end{aligned}$$

and, as a consequence, the interpolation condition $\varphi_0(0) = \varphi_1(0) = 0$ holds. In addition, we also set $\alpha_2 = 1$ in order to enforce the independency on y_{n-1}. We next determine the remaining basis functions $\chi_1(s), \chi_2(s), \psi_1(s), \psi_2(s)$ by imposing the system of order conditions (12.3.3) up to $p = 3$, transferring to them the interpolation condition in 0. Then, At this point, everything depends on the values of $\alpha_0, \alpha_1, \beta_0, \beta_1, \beta_2, c_1$ and c_2. We spend α_0, α_1 and β_0 in order to obtain ΨW being diagonal and one point spectrum. We next enforce some further simplifying assumptions on the basis functions, using the values of α_2 and β_1, obtaining a three-parameter family of methods, depending on c_1, c_2 and β_2. We next derive the stability polynomial $p(\omega, z)$ of order 6 with respect to the variable ω, which assumes the form (13.3.1), where now $p_j(z)$, $j = 0, 1, \ldots, 5$, are rational functions with respect to z depending on c_1, c_2 and α_2. We apply the Schür criterion, in order to determine the values of the free parameters c_1, c_2 and β_2 achieving A-stability. The results are shown in Figure 13.3.

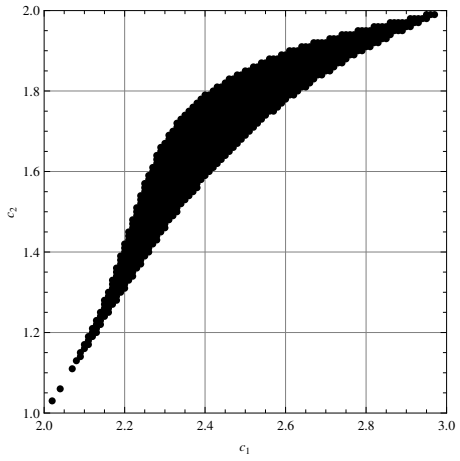

Figure 13.3: Region of A-stability in the parameter space (c_1, c_2) for SDTSAC methods (13.1.22), with $m = 2$, $p = 3$ and $\beta_2 = \frac{1}{4}$

Part IV

Conclusions and future research

This dissertation has pointed out the attention on the derivation, the analysis and the implementation of highly stable methods for the numerical integration of first and second order Ordinary Differential Equations and Volterra Integral Equations. The main point we have addressed is the requirement of strong stability, which is an essential property in order to carry out an accurate and efficient integration of stiff system. We have derived A-stable, L-stable and algebraically stable methods within the class of TSRK methods (2.2.1), mostly obtained by suitable modifications of the collocation technique we have named *two-step (almost) collocation technique*: the spirit of these modifications is the combination of high stability with other desired properties, e.g. structured coefficient matrices, one-point spectrum matrices, quadratic stability, narrowed contribution of high order terms in the local error. One of the peculiarity of the derived methods is that they are *continuous*, i.e. the approximation is provided by means of a continuous function (in our case an algebraic polynomial) which can suitably be exploited in the variable stepsize implementation of the methods, without the need of deriving continuous extensions for the analyzed methods, as it happens for discrete numerical methods (for instance, see [138]).

The experiments we have carried out up to now show that two-step almost collocation methods can provide building blocks for the design of a modern solver for first order ODEs. In fact, they do not suffer from order reduction in the integration of stiff systems as the case of Runge–Kutta methods. Moreover, the local error estimation we have derived in Chapter 6 is absolutely reliable and particularly cheap if compared, for instance, with Richardson extrapolation. Of course, in order to carry out serious comparisons between our methods and the classical methods already considered and implemented in the literature, it is necessary to derive high order formulae, for instance up to order 8, and design a variable stepsize-variable order strategy: these aspects will be addressed in future works. Further developments of this research will address, together with the design of a variable step-variable order solver for two-step almost collocation methods, the construction and the implementation of algebraically stable two-step almost collocation methods and the analysis of the G-simplecticity properties of TSRK methods, which have not yet been considered in the literature up to now.

Concerning second order ODEs, we have mainly addressed our attention on the analysis of the family of two-step hybrid methods (9.1.1)-(9.1.2), in order to provide their adapted versions in order to efficiently approach problems exhibiting an oscillatory behaviour, by means of trigonometrical/exponential

fitting arguments. Concerning exponentially-fitted multistage methods, we are investigating on the possibility to consider how the error in the internal stages cumulates and is inherited by the external ones since, up to now, the EF-based multistage methods presented in the literature have been derived under the unrealistic hypothesis that the values provided by the internal stages are exact.

Further developments of the research on second order ODEs mainly address the importance to derive an unifying framework of the numerical methods for their integration, i.e. the construction of General Linear Methods for second order ODEs. This analysis is certainly nontrivial, but can benefit of the lines drawn in the construction of the theory of GLMs for first order ODEs. Moreover, our first steps in this direction and related to possible desirable stability features, such as the ones presented in Chapter 11, could provide an useful tool to derive highly stable GLMs, e.g. P-stable formulae.

Concerning Volterra Integral Equations, our interest in this area is quite recent. We aim to address in future works the variable step-variable order implementation of high order two-step almost collocation methods for VIEs. This is a nontrivial task and, moreover, quite an open problem in this area since, up to now, only few attempts have been provided in the literature, such as the solver COLVI2 (by Blom and Brunner, 1991). In order to succeed in the design of variable step-variable order solver for VIEs, we need to address various implementation issues such as the choice of appropriate starting procedures, stepsize and order changing strategy, the derivation of a proper local error estimation.

Bibliography

[1] Abdou, M. A.; Badr, A. A. On a method for solving an integral equation in the displacement contact problem, *Appl. Math. Comput.* 127 (2002), no. 1, 65–78.

[2] Acary, V.; Brogliato, B. *Numerical Methods for Nonsmooth Dynamical Systems. Applications in Mechanics and Electronics. Springer Verlag Berlin Heidelberg*, Lecture Notes in Applied and Computational Mechanics, vol. 35, 2008.

[3] Agarwal, R. P.; O'Regan, D. *Integral and Integrodifferential Equations. Theory, Methods and Applications*, Series in Mathematical Analysis and Applications, vol. 2, Gordon and Breach, Amsterdam, 2000.

[4] Albert A. Conditions for positive and non-negative definiteness in terms of pseudoinverses, *SIAM J. Appl. Math.* 17 (1969), 434-440.

[5] Albrecht, P. Numerical treatment of ODEs: the theory of A-methods, *Numer. Math.* 47 (1985), no. 1, 59–87.

[6] Albrecht, P. A new theoretical approach to Runge–Kutta methods, *SIAM J. Numer. Anal.* 24 (1987), no. 2, 391–406.

[7] 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.

[8] Albrecht, P. The Runge-Kutta theory in a nutshell, *SIAM J. Numer. Anal.* 33 (1996), no. 5, 1712–1735.

[9] Alekseev, V.M. An estimate for the perturbations of the solution of ordinary differential equations (Russian), *Vestn. Mosk. Univ., Ser.I, Math. Meh*, 2 (1961), 28-36.

[10] Alexander, R. Diagonally implicit Runge–Kutta methods for stiff o.d.e.'s, *SIAM J. Numer. Anal.* 14 (1977), no. 6, 1006–1021.

[11] Alfieri, R. ; Mosca, E.; Merelli, I.; Milanes, L. Parameter Estimation for Cell Cycle Ordinary Differential Equation (ODE) Models using a Grid Approach, *Stud. Health Technol. Inform.* 126 (2007), 93-102.

[12] 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 Schrödinger equation, *J. Math. Chem.* 41 (2007), no. 1, 79–100.

[13] Atkinson, K.E. *Introduction to Numerical Analysis*, Wiley, New York, 1989.

[14] Bartoszewski, Z.; Jackiewicz, Z. Construction of two-step Runge-Kutta methods of high order for ordinary differential equations, *Numer. Algorithms* 18 (1998), 51-70.

[15] Bartoszewski, Z.; Jackiewicz, Z. Stability analysis of two-step Runge-Kutta methods for delay differential equations, *Comput. Math. Appl.* 44 (2002), 83–93.

[16] Bartoszewski, Z.; Jackiewicz, Z. Nordsieck representation of two-step Runge-Kutta methods for ordinary differential equations, *Appl. Numer. Math.* 53 (2005), 149–163.

[17] Bartoszewski, Z.; Jackiewicz, Z. Derivation of continuous explicit two-step Runge-Kutta methods of order three, *J. Comput. Appl. Math.* 205 (2007), 764–776.

[18] 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.

[19] 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).

[20] Bellen, A.; Jackiewicz, Z.; Vermiglio, R.; Zennaro, M. Stability analysis of Runge-Kutta methods for Volterra integral equations of the second kind, *IMA J. Numer. Anal.* 10 (1990), no. 1, 103–118.

[21] Bellen, A.; Jackiewicz, Z.; Vermiglio, R.; Zennaro, M. Natural continuous extensions of Runge-Kutta methods for Volterra integral equations of the second kind and their applications, *Math. Comp.* 52 (1989), no. 185 ,49–63.

[22] Beuter, A.; Glass, L.; Mackey, M. C.; Titcombe, M. S. *Nonlinear Dynamics in Physiology and Medicine*, Springer, 2002.

[23] Bother, P.; De Meyer, H.; Vanden Berghe, G. Numerical solution of Volterra equations based on mixed interpolation, *Comput. Math. Appl.* 27 (1994), 1–11.

[24] Brunner, H. *Collocation Methods for Volterra Integral and Related Functional Equations*, Cambridge University Press, 2004.

[25] Brunner, H.; Makroglou, A.; Miller, R. K. On mixed collocation methods for Volterra integral equations with periodic solution, *Appl. Numer. Math.* 24 (1997), 115-130.

[26] Brunner, H.; Makroglou, A.; Miller, R. K. Mixed interpolation collocation methods for first and second order Volterra integro-differential equations with periodic solution, *Appl. Numer. Math.* 23 (1997), 381–402.

[27] Brunner, H.; van der Houwen, P. J. *The Numerical Solution of Volterra Equations*, CWI Monographs, 3, North-Holland Publishing, Amsterdam, 1986.

[28] Burrage, K. A special family of Runge-Kutta methods for solving stiff differential equations, *BIT* 18 (1978), 22-41.

[29] Burrage, K.; Butcher, J. C.; Chipman, F. H. An implementation of singly–implicit Runge–Kutta methods, *BIT* 20 (1980), 326-340.

[30] Burrage, K.; Butcher, J. C. Stability criteria for implicit Runge-Kutta methods. *SIAM J. Numer. Anal.* 16(1979), no. 1, 46–57.

[31] Burrage, K.; Butcher, J. C. Nonlinear stability of a general class of differential equation methods, *BIT* 20 (1980), no.2, 185–203.

[32] Butcher, J. C. Coefficients for the study of Runge-Kutta integration processes, *J. Austral. Math. Soc.* 3 (1963), 185–201.

[33] Butcher, J. C. On the integration process of A. Huta, *J. Austral. Math. Soc.* 3 (1963), 202–206.

[34] Butcher, J. C. A transformed implicit Runge-Kutta method, *J. Assoc. Comput. Mach.* 26 (1979), 731–738.

[35] Butcher, J. C. A modified multistep method for the numerical integration of ordinary differential equations, *J. Assoc. Comput. Mach.* 12 (1965), 124–135.

[36] Butcher, J. C. *The Numerical Analysis of Ordinary Differential Equations. Runge-Kutta and General Linear Methods*, John Wiley & Sons, Chichester, New York, 1987.

[37] Butcher, J. C. Linear and nonlinear stability for general linear methods, *BIT* 27 (1987), no. 2, 182–189.

[38] Butcher, J. C. The equivalence of algebraic stability and AN-stability, *BIT* 27 (1987), 510-533.

[39] Butcher, J.C. Diagonally-implicit multi-stage integration methods, *Appl. Numer. Math.* 11 (1993), 347–363.

[40] Butcher, J.C. Thirty years of G-stability, *BIT* 46 (2006), 479–489.

[41] Butcher, J.C. Forty-five years of A-stability, *J. Numer. Anal. Ind. Appl. Math* 4 (2009), no. 1-2, 1–9.

[42] Butcher, J.C. *Numerical Methods for Ordinary Differential Equations*, 2nd Edition, John Wiley & Sons, Chichester, 2008.

[43] Butcher, J. C.; Jackiewicz, Z. Diagonally implicit general linear methods for ordinary differential equations, *BIT* 33(1993), no.3, 452–472.

[44] Butcher, J. C.; Jackiewicz, Z. Construction of diagonally implicit general linear methods of type 1 and 2 for ordinary differential equations, *Appl. Numer. Math.* 21 (1996), n0. 4, 385–415.

[45] Butcher, J. C.; Jackiewicz, Z. Implementation of diagonally implicit multistage integration methods for ordinary differential equations, *SIAM J. Numer. Anal.* 34 (1997), no. 6, 2119–2141.

[46] Butcher, J. C.; Jackiewicz, Z. A new approach to error estimation for general linear methods, *Numer. Math.* 95 (2003), no. 3, 487–502.

[47] Butcher, J. C.; Jackiewicz, Z.; Wright, W.M. Error propagation of general linear methods for ordinary differential equations, *J. Complexity* 23 (2007), no. 4-6, 560–580.

[48] Butcher, J. C.; Tracogna, S. Order conditions for two-step Runge-Kutta methods, *Appl. Numer. Math.* 24 (1997), 351–364.

[49] Butcher, J. C.; Wright, W.M. A transformation relating explicit and diagonally-implicit general linear methods, *Appl. Numer. Math.* 44 (2003), 313–327.

[50] Butcher, J. C.; Wright, W.M. The construction of practical general linear methods, *BIT* 43 (2003), 695–721.

[51] Butcher, J. C.; Wright, W.M. Applications of doubly companion matrices, *Appl. Numer. Math.* 56 (2006), 358–373.

[52] Byrne, G. D.; Lambert, R. J. Pseudo-Runge-Kutta methods involving two points, *J. Assoc. Comput. Mach.* 13 (1966), 114–123.

[53] 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.

[54] Capobianco, G.; Conte, D.; Del Prete, I.; Russo, E. Fast Runge–Kutta Methods for nonlinear convolution systems of Volterra Integral Equations, *BIT* 47 (2007), no. 2, 259–275.

[55] Capobianco, G.; Conte, D.; Del Prete, I.; Russo, E. Stability analysis of fast numerical methods for Volterra Integral Equations, *Electron. Trans. Numer. Anal.* 30 (2008), 305–322.

[56] Cardone, A.; Ixaru, L. Gr.; Paternoster, B. Exponential fitting Direct Quadrature methods for Volterra integral equations,*Numer. Algorithms*, doi: 10.1007/s11075-010-9365-1.

[57] Carr, J.; Duncan, D. B.; Walshaw, C. H. Numerical approximation of a metastable system, *IMA J. Numer. Anal.* 15 (1995), no. 4, 505-521.

[58] 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.

[59] Chan, R. P. K.; Leone, P.; Tsai, A. Order conditions and symmetry for two-step hybrid methods, *Int. J. Comput. Math.* 81 (2004), no. 12, 1519–1536.

[60] 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.

[61] Coleman, J. P. Rational approximations for the cosine function; P-acceptability and order, *Numer. Algorithms* 3 (1992), no. 1-4, 143-158.

[62] Coleman, J. P. Mixed interpolation methods with arbitrary nodes, *J. Comput. Appl. Math.* 92 (1998), 69-83.

[63] 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.

[64] 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.

[65] Coleman, J. P.; Ixaru, L.Gr. P-stability and exponential-fitting methods for $y'' = f(x, y)$, *IMA J. Numer. Anal.* **16**, 2 (1996), 179–199.

[66] Coleman, J. P.; Ixaru, L.Gr. Truncation errors in exponential fitting for oscillatory problems, *SIAM J. of Numerical Analysis* 44 (2006), 1441-1465.

[67] 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.

[68] Cong, N. H.; Mitsui, T. Collocation-based two-step Runge-Kutta methods, *Japan J. Indust. Appl. Math.* 13 (1996), no. 1, 171–183.

[69] Conte, D.; D'Ambrosio, R.; Ferro, M.; Paternoster, B. Piecewise-polynomial approximants for solutions of Functional Equations, in press on *Collana Scientifica di Ateneo, Univ. degli Studi di Salerno*.

[70] Conte, D.; D'Ambrosio, R.; Ferro, M.; Paternoster, B. Practical construction of Two-Step Collocation Runge–Kutta methods for Ordinary Differential Equations, in *Applied and Industrial Mathematics in Italy III*, 278-288, World Scientific Publishing, Singapore, 2009.

[71] Conte, D.; D'Ambrosio, R.; Ferro, M.; Paternoster, B. Modified Collocation Techniques for Volterra Integral Equations, in *Applied and Industrial Mathematics in Italy III*, 268-277, World Scientific Publishing, Singapore, 2009.

[72] Conte, D.; D'Ambrosio, R.; Jackiewicz, Z. Two-step Runge-Kutta methods with quadratic stability functions, *submitted*.

[73] Conte, D.; D'Ambrosio, R.; Jackiewicz, Z.; Paternoster, B. Algebraically stable two-step collocation methods for ordinary differential equations, *in preparation*.

[74] Conte, D.; Del Prete, I. Fast collocation methods for Volterra integral equations of convolution type, *J. Comput. Appl. Math.* 196 (2006), no. 2, 652–663.

[75] Conte, D.; Jackiewicz, Z.; Paternoster, B. Two-step almost collocation methods for Volterra integral equations, *Appl. Math. Comput.* 204 (2008), 839–853.

[76] Conte, D.; Paternoster, B. A Family of Multistep Collocation Methods for Volterra Integral Equations, in *Numerical Analysis and Applied Mathematics* ed. by T. E. Simos, G. Psihoyios, Ch. Tsitouras; *AIP Conference Proceedings* 936 (2007), 128–131.

[77] Conte, D.; Paternoster, B., Multistep collocation methods for Volterra Integral Equations, *Appl. Numer. Math.* 59 (2009), 1721–1736.

[78] Corduneanu, C.; Sandberg, I. W. *Volterra Equations and Applications*, Stability Control, Theory, Methods Appl. 10, Gordon and Breach, Amsterdam, 2000.

[79] Crisci, M. R.; Russo, E.; Vecchio, A. Stability results for one-step discretized collocation methods in the numerical treatment of Volterra integral equations, *Math. Comp.* 58 (1992), no. 197, 119–134.

[80] Crouzeix, M. Sur la B-stabilité des méthodes de Runge-Kutta, *Numer. Math.* 32 (1979), no. 1, 75–82.

[81] Curtiss, C. F.; Hirschfelder, J. O., Integration of stiff equations, *Proc. Nat. Acad. Sci. U. S. A.* 38 (1952), 235–243.

[82] Dahlquist, G. A special stability property for linear multistep methods, *BIT* 3 (1963), 27-43.

[83] Dahlquist, G. Error analysis for a class of methods for stiff nonlinear initial value problems, in *Proc. Numer. Anal. Conf.* (Dundee, Scotland, 1975), *Lecture Notes Math.* 506, 60–74, Springer, New York, 1976.

[84] Dahlquist, G. G-stability is equivalent to A-stability, *BIT* 18 (1978), no. 4, 384–401.

[85] Dahlquist, G. On one-leg multistep methods, *SIAM J. Numer. Anal.* 20 (1983), no. 6, 1130–1138.

[86] Dahlquist, G.; Liniger, W.; Nevanlinna, O. Stability of two-step methods for variable integration steps. *SIAM J. Numer. Anal.* 20 (1983), no. 5, 1071–1085.

[87] D'Ambrosio, R.; Esposito, E.; Paternoster, B. Exponentially fitted two-step hybrid methods for $y'' = f(x, y)$, *submitted*.

[88] D'Ambrosio, R.; Ferro, M.; Paternoster, B. A General Family of Two Step Collocation Methods for Ordinary Differential Equations, in *Numerical Analysis and Applied Mathematics* ed. by T. E. Simos, G. Psihoyios, Ch. Tsitouras; *AIP Conference Proceedings* 936 (2007) 45–48.

[89] D'Ambrosio, R.; Ferro, M.; Paternoster, B. Collocation based two step Runge–Kutta methods for Ordinary Differential Equations, *Lecture Notes in Comput. Sci.* 5073, Part II, 736–751, Springer, New York, 2008.

[90] D'Ambrosio, R.; Ferro, M.; Paternoster, B. Two-step hybrid collocation methods for $y'' = f(x, y)$, *Appl. Math. Lett.* 22 (2009), 1076–1080.

[91] D'Ambrosio, R.; Ferro, M.; Paternoster, B. Trigonometrically fitted two-step hybrid methods for second order ordinary differential equations with one or two frequencies, accepted for publication on *Math. Comput. Simulat.*

[92] 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.

[93] D'Ambrosio, R.; Izzo, G.; Jackiewicz, Z. Search for highly stable two-step Runge–Kutta methods for ordinary differential equations, accepted for publication on *Appl. Numer. Math.*

[94] D'Ambrosio, R.; Jackiewicz, Z. Continuous Two–Step Runge–Kutta Methods for Ordinary Differential Equations, *Numer. Algorithms* (doi: 10.1007/ s11075-009-9329-5).

[95] D'Ambrosio, R.; Jackiewicz, Z., Construction and implementation of highly stable two-step collocation methods for stiff differential systems, *submitted.*

[96] D'Ambrosio, R.; Izzo, G.; Jackiewicz, Z. Highly Stable General Linear Methods for Differential Systems, in *Numerical Analysis and Applied Mathematics* ed. by T. E. Simos, G. Psihoyios, Ch. Tsitouras; *AIP Conference Proceedings* 1168 (2009), no. 1, 21–24.

[97] 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.

[98] D'Ambrosio, R.; Paternoster, B. Two-step modified collocation methods with structured coefficient matrices for ordinary differential equations, *submitted.*

[99] Dekker, K.; Verwer, J. G. *Stability of Runge-Kutta Methods for Stiff Nonlinear Differential Equations*, CWI Monographs, vol. 2, North-Holland Publishing, Amsterdam, 1984.

[100] De Meyer, H.; Vanthournout, J.; Vanden Berghe, G. On a new type of mixed interpolation, *J. Comput. Appl. Math.* 30 (1990), 55-69.

[101] Demmel, J. W. *Applied Numerical Linear Algebra*, SIAM, Philadelphia, 1997.

[102] Donelson, J. III; Hansen, E. Cyclic composite multistep predictor-corrector methods, *SIAM J. Numer. Anal.* 8 (1971), 137–157.

[103] Duffing, G. *Erzwungene Schwingungen bei Vernderlicher Eigenfrequenz.*, F. Vieweg u. Sohn, Braunschweig, 1918.

[104] 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

[105] Engquist, B.; Fokas, A. S.; Hairer, E.; Iserles, A. *Highly Oscillatory Problems*, London Mathematical Society Lecture Note Series, Cambridge University Press, 2009.

[106] 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.

[107] Franco, J. M. Exponentially fitted explicit Runge-Kutta Nyström methods, *J. Comput. Appl. Math.* 167 (2004), 1-19.

[108] 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.

[109] Franco, N. B. A Volterra integral equation arising from the propagation of nonlinear waves, *Rev. Mat. Estat.* 17 (1999), 35–49.

[110] Gautschi, W. Numerical integration of ordinary differential equations based on trigonometric polynomials, *Numer. Math.* **3** (1961), 381-397.

[111] Gear, C. W., Hybrid methods for initial value problems in ordinary differential equations, *J. Soc. Indust. Appl. Math. Ser. B Numer. Anal.* 2, 69–86 (1965).

[112] Golub, G. H.; Van Loan, C. F. *Matrix Computations*, The Johns Hopkins University Press, Baltimore, London, 1996.

[113] Gragg, W. B.; Stetter, H. J. Generalized multistep predictor-corrector methods, *J. Assoc. Comput. Mach.* 11 (1964), 188-209.

[114] Gröbner, W. Die Liereihen und ihre Anwendungen. VEB Deutscher Verlag derWiss., Berlin, Second edition, 1967.

[115] Guardiola, J.; Izzo, G.; Vecchio, A. Simulating the effect of vaccine-induced immune responses on HIV infection, *Hum. Immunol.* 64 (2003), 840-851.

[116] Guillou, A.; Soulé, F. L. La résolution numérique des problèmes differentiels aux conditions par des méthodes de collocation, *RAIRO Anal. Numér. Ser. Rouge* R-3 (1969), 17–44.

[117] Gustafsson, K.; Lundh, M.; Söderlind, G. A PI stepsize control for the numerical solution of ordinary differential equations, *BIT* 28 (1988), no. 2, 270–287.

[118] Hairer, E.; Lubich, C.; Wanner, G. *Geometric numerical integration. Structure-preserving algorithms for ordinary differential equations*, Second edition, Springer Series in Computational Mathematics 31, Springer-Verlag, Berlin, 2006.

[119] Hairer, E.; Norsett, S. P.; Wanner, G. *Solving Ordinary Differential Equations I – Nonstiff Problems*, Springer Series in Computational Mathematics 8, Springer–Verlag, Berlin, 2000.

[120] Hairer, E.; Wanner, G. On the Butcher group and general multi-value methods. *Computing* 13 (1974), 1-15.

[121] Hairer, E.; Wanner, G. Order conditions for general two-step Runge-Kutta methods, *SIAM J. Numer. Anal.* 34(1997), 2087-2089.

[122] Hairer, E.; Wanner, G. *Solving Ordinary Differential Equations II – Stiff and Differential–Algebraic Problems*, Springer Series in Computational Mathematics 14, Springer–Verlag, Berlin, 2002.

[123] Henrici, P. *Discrete variable methods in ordinary differential equations*, John Wiley & Sons, New York-London, 1962.

[124] Hetcote, H. W.; Tudor, D. W. Integral equation models for endemic infectious diseases, *J. Math. Biol.*, 9 (1980), 37–47.

[125] Hewitt L. L.; Hill, A. T. Algebraically stable general linear methods and the G-matrix, *BIT* 49 (2009), 93–111.

[126] Hewitt L. L.; Hill, A. T. Algebraically stable diagonally implicit general linear methods, *submitted*.

[127] Hill, A. T. Nonlinear stability of general linear methods, *Numer. Math.* 103 (2006), 611–629.

[128] 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.

[129] Hoppensteadt, F. C.; Jackiewicz, Z.; Zubik-Kowal, B. Numerical solution of Volterra integral and integro-differential equations with rapidly vanishing convolution kernels, *BIT* 47 (2007), no. 2, 325–350.

[130] Huang, W.; Li, Y.; Chen, W. Analysis of the dynamic response of a fluid-supported circular elastic plate impacted by a low-velocity projectile, Proceedings of the Institution of Mechanical Engineers. Part C, Journal of mechanical engineering science 214 (2000), no. 5, 719-727.

[131] Iannelli, M.; Martcheva, M.; Milner, F. A. *Gender-structured Population Modeling : Mathematical Methods, Numerics, and Simulations*, SIAM Frontiers in Applied Mathematics, 2005.

[132] Ixaru, L.Gr. Operations on oscillatory functions, *Comput. Phys. Comm.* 105 (1997), 1–19.

[133] 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.

[134] Ixaru, L. Gr.; Vanden Berghe, G.; De Meyer, H. Frequency evaluation in exponential fitting multistep methods for ODEs, *J. Comput. Appl. Math.* 140 (2002), 423–435.

[135] Ixaru, L. Gr.; Vanden Berghe, G. *Exponential Fitting*, Kluwer Academic Publishers, Dordrecht, 2004.

[136] Iyengar, R. Quantitative models of Mammalian Cell Signaling Pathways, *Sci. Signal.* 1 (2008), no. 7.

[137] Izzo, G.; Messina, E.; Jackiewicz, Z.; Vecchio, A. General linear methods for Volterra integral equations, *J. Comput. Appl. Math.*, doi:10.1016/j.cam.2010.01.023.

[138] Jackiewicz, Z. *General Linear Methods for Ordinary Differential Equations*, John Wiley & Sons, Hoboken, New Jersey, 2009.

[139] Jackiewicz, Z.; Renaut, R.; Feldstein, A. Two-step Runge-Kutta methods, *SIAM J. Numer. Anal.* 28 (1991), no. 4, 1165–1182.

[140] 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.

[141] Jackiewicz, Z.; Tracogna, S. Variable stepsize continuous two-step Runge-Kutta methods for ordinary differential equations, *Numer. Algorithms* 12 (1996), 347–368.

[142] Jackiewicz, Z.; Verner, J. H. Derivation and implementation of two-step Runge-Kutta pairs, *Japan J. Indust. Appl. Math.* 19 (2002), 227–248.

[143] 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.

[144] Janestad, H.; Wendin, K.; Ruheb, A.; Halla, G. Modelling of dynamic flavour properties with ordinary differential equations, *Food Qual. Prefer.* 11 (2000), no. 4, 323-329.

[145] 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, 547554.

[146] Kramarz, L. Stability of collocation methods for the numerical solution of $y'' = f(t, y)$, *BIT* 20 (1980), no. 2, 215–222.

[147] Lambert, J. D. *Numerical methods for ordinary differential systems: The initial value problem*, John & Wiley, Chichester, 1991.

[148] Lambert, J. D.; Watson, I. A. Symmetric multistep methods for Periodic Initial Value Problems, *J. Inst. Maths. Applics.* 18 (1976), 189–202.

[149] Lang, S. *Linear Algebra*, Springer, 1987.

[150] Larter, R.; Speelman, B.; Worth, R. M. A coupled ordinary differential equation lattice model for the simulation of epileptic seizures, *Chaos* 9 (1999), no. 3, 795-804.

[151] 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.* 97 (1997), no. 2, 165-172.

[152] Lefschetz, S. *Differential Equations: Geometric Theory*, Interscience Publishers, New York, 1957.

[153] Lie, I. The stability function for multistep collocation methods, *Numer. Math.* 57 (1990), no. 8, 779–787.

[154] Lie, I.; Norsett, S. P. Superconvergence for Multistep Collocation, *Math. Comp.* 52 (1989), no. 185, 65–79.

[155] Lopez-Fernandez, M.; Lubich, C.; Schädle, A. Adaptive, fast, and oblivious convolution in evolution equations with memory, *SIAM J. Sci. Comp.* 30 (2008), 1015–1037.

[156] Lopez-Fernandez, M.; Lubich, C.; Schädle, A. Fast and oblivious convolution quadrature, *SIAM J. Sci. Comp.* 28 (2006), 421–438.

[157] Lyche, T. Chebyshevian multistep methods for ordinary differential equations, *Numer. Math.* 19 (1972), 65–75.

[158] 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 *et al.*; *Lecture Notes in Comput. Sci.* 3039, Part IV, Springer Verlag, 2004, 418–425.

[159] Martucci, S.; Paternoster, B., General two step collocation methods for special second order Ordinary Differential Equations, Paper Proceedings of the 17th IMACS World Congress Scientific Computation, *Applied Mathematics and Simulation*, Paris, 2005.

[160] Nerukh, A. G.; Sewell, P.; Benson T. M. Volterra Integral Equations for Nonstationary Electromagnetic Processes in Time-Varying Dielectric Waveguides, *J. Lightwave Technol.* 22 (2004), no. 5.

[161] Noble, D.; Varghese, A.; Kohl, P.; Noble, P. Improved guinea-pig ventricular cell model incorporating a diadic space, I_{Kr} and I_{Ks}, length- and tension-dependent processes, *Can. J. Cardiol.* 14 (1998), 123134.

[162] 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).

[163] Norsett, S. P. Semi–explicit Runge–Kutta methods, Report No. 6/74, Dept. of Math., Univ. of Trondheim, 1974.

[164] Norsett, S. P. Collocation and perturbed collocation methods, in *Numerical analysis, Proc. 8th Biennial Conf.* (Univ. Dundee, Dundee, 1979), 119–132; *Lecture Notes in Math.* 773, Springer, Berlin, 1980.

[165] Norsett, S. P.; Wanner, G. Perturbed collocation and Runge Kutta methods, *Numer. Math.* 38 (1981), no. 2, 193–208.

[166] Oppenheim, A. V.; Schafer, R. W.; Buck, J. R. *Discrete–Time Signal Processing*, 2nd edn., Prentice Hall, Upper Saddle River, 1999.

[167] Owren, B.; Zennaro, M. Order barriers for continuous explicit Runge-Kutta methods, *Math. Comput.* 56 (1991), 645–661.

[168] Owren, B.; Zennaro, M. Derivation of efficient, continuous, explicit Runge-Kutta methods, *SIAM J. Sci. Statist. Comput.* 13 (1992), 1488–1501.

[169] Owren, B.; Zennaro, M. Continuous explicit Runge-Kutta methods. Computational ordinary differential equations (London, 1989), 97–105, *Inst. Math. Appl. Conf. Ser. New Ser.* 39, Oxford University Press, New York, 1992.

[170] Paternoster, B. Runge-Kutta(-Nyström) methods for ODEs with periodic solutions based on trigonometric polynomials, *Appl. Numer. Math.* 28 (1998), 401-412.

[171] Paternoster, B. A phase–fitted collocation–based Runge–Kutta–Nyström methods, *Appl. Numer. Math.* 35 (2000), no. 4, 339–355.

[172] Paternoster, B. General two-step Runge-Kutta methods based on algebraic and trigonometric polynomials, *Int. J. Appl. Math.* 6 (2001), no. 4, 347–362.

[173] Paternoster, B. Two step Runge-Kutta-Nystrom 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.

[174] 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.

[175] Paternoster, B. Two step Runge-Kutta-Nystrom methods based on algebraic polynomials, *Rendiconti di Matematica e sue Applicazioni* 23 (2003), Serie VII, 277–288.

[176] Paternoster, B. Two step Runge-Kutta-Nystrom methods for oscillatory problems based on mixed polynomials, in *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.

[177] 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.

[178] Petzold, L. R.; Jay, L. O.; Yen, J. Numerical solution of highly oscillatory ordinary differential equations, *Acta Numer.* 6 (1997), 437–483.

[179] Podhaisky, H.; Schmitt B. A.; Weiner, R. Design, analysis and testing of some parallel two-step W-methods for stiff systems, *Appl. Numer. Math.* 42 (2002), 381–395.

[180] Prothero, A.; Robinson, A. On the stability and the accuracy of one-step methods for solving stiff systems of ordinary differential equations, *Math. Comput.* 28 (1974), 145–162.

[181] Raptis, A. D.; Allison, A. C. Exponential-fitting methods for the numerical solution of the Schrödinger equation, *Computational Physics Communications* 14 (1978), 1-5.

[182] Richtmyer, R. D.; Morton, K. W. *Difference methods for initial-value problems*, Second edition, Interscience Tracts in Pure and Applied Mathematics 4, Interscience Publishers John Wiley & Sons, Inc., New York-London-Sydney, 1967.

[183] Schädle, A.; López-Fernández, M.; Lubich, C. *Fast and oblivious convolution quadrature*, *SIAM J. Sci. Comput.* 28 (2006), no. 2, 421–438.

[184] Schur, J. Über Potenzreihen die im Innern des Einheitskreises beschränkt sind, *J. Reine Angew. Math.* 147 (1916), 205–232.

[185] Shampine, L. F. *Numerical Solution of Ordinary Differential Equations*, Chapman & Hall, New York, London, 1994.

[186] Shampine, L. F. *Computer Solution of Ordinary Differential Equations*, W.H. Freeman and Company, San Francisco, 1975.

[187] Shampine, L. F.; Baca L. S. Error estimators for stiff differential equations, *J. Comput. Appl. Math.* 11 (1984), no. 2, 197–207.

[188] Shampine, L. F.; Gordon, M. K. *Computer Solution of Ordinary Differential Equations. The Initial Value Problem*, W.H. Freeman and Company, San Francisco, 1975.

[189] Simos, T. E. Dissipative trigonometrically-fitted methods for linear second-order IVPs with oscillating solution, *Appl. Math. Lett.* **17**, 5 (2004), 601–607.

[190] 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).

[191] Söderlind, G. Automatic control and adaptive time-stepping, *Numer. Algorithms* 31 (2002), no. 1-4, 281–310.

[192] Söderlind, G. Digital filters in adaptive time-stepping, *ACM Trans. Math. Software* 29 (2003), no. 1, 1–26.

[193] Southern, J.; Pitt-Francis, J.; Whiteley, J.; Stokeley, D.; Kobashi, H.; Nobes, R.; Kadooka, Y.; Gavaghan, D. Multi-scale computational modelling in biology and physiology, *Prog. Biophys. Mol. Bio.* 96 (2008), 60-89.

[194] 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.

[195] Stiefel, E.; Bettis, D. G. Stabilization of Cowell's method, *Numer. Math.* 13 (1969), 154–175.

[196] Svirezhev, Y. M.; von Bloh, W. A minimal model of interaction between climate and vegetation: qualitative approach, *Ecol. Mod.* 92 (1996), 89-99.

[197] Tracogna, S. Implementation of two-step Runge-Kutta methods for ordinary differential equations, *J. Comput. Appl. Math.* 76 (1997), 113–136.

[198] Tracogna, S.; Welfert, B. Two-step Runge-Kutta: Theory and practice, *BIT* 40 (2000), 775–799.

[199] Ueda, Y. Explosion of strange attractors exhibited by Duffing's equation, in *Nonlinear Dynamics*, R.H.G. Helleman (ed.), 422-434, New York Academy of Science, 1980.

[200] Ueda, Y. *The Road to Chaos*, Aerial Press, 1992.

[201] Vanden Berghe, G.; De Meyer, H.; Van Daele, M.; Van Hecke, T. Exponentially fitted Runge-Kutta methods, *Comput. Phys. Comm.* 123 (1999), 7-15.

[202] Vanden Berghe, G.; Van Daele, M. Exponentially-fitted Numerov methods, *J. Comput. Appl. Math.* 200 (2007), 140-153.

[203] van der Houwen, P. J.; Sommeijer, B. P. Diagonally implicit Runge–Kutta Nyström methods for oscillatory problems, *SIAM J. Num. Anal.* 26 (1989), 414–429.

[204] van der Houwen, P. J.; Sommeijer, B. P.; Cong, N. H. Stability of collocation–based Runge–Kutta–Nyström methods, *BIT* 31 (1991), 469–481.

[205] Van de Vyver, H. Phase-fitted and amplification-fitted two-step hybrid methods for $y'' = f(x, y)$, *J. Comput. Appl. Math.* 209 (2007), no. 1, 33–53.

[206] Van de Vyver, H. A phase-fitted and amplification-fitted explicit two-step hybrid method for second-order periodic initial value problems, *Internat. J. Modern Phys.* C 17 (2006), no. 5, 663–675.

[207] 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.

[208] Wright, K. Some relationships between implicit RungeKutta, collocation and Lanczos τ-methods, and their stability properties, *BIT* 10 (1970), 217–227.

[209] Wright, W. M. General linear methods with inherent Runge-Kutta stability, Ph.D. thesis, The University of Auckland, New Zealand, 2002.

[210] Wright, W. M. Explicit general linear methods with inherent Runge-Kutta stability, *Numer. Algorithms* 31 (2002), 381–399.

[211] Zhao, X. Q. *Dynamical Systems in Population Biology*, CMS Books in Mathematics, New York, Springer–Verlag, 2003.