Extending Convergence Theory for Nonlinear Stiff problems
Part II

Winfried Auzinger
Alexander Eder
Reinhard Frank

ANUM Preprint No. 22/01

Technische Universität Wien
Vienna University of Technology

Institute for Applied Mathematics
and Numerical Analysis

BIT 0006-3835/00/4004-0001 $15.00
xxxx, Vol. xx, No. x, pp. xxx–xxx

EXTENDING CONVERGENCE THEORY FOR NONLINEAR STIFF PROBLEMS PART II *

W. AUZINGER[1], A. EDER[1], and R. FRANK[1] †

[1] *Department of Applied Mathematics and Numerical Analysis, Vienna University of Technology, Wiedner Hauptstrasse 8-10/115,*
A-1040 Vienna, Austria, EU. email: w.auzinger@tuwien.ac.at

Abstract.

This paper deals with the convergence properties of high-order implicit Runge-Kutta methods applied to nonlinear stiff initial value problems. Earlier convergence concepts like the theory of B-convergence or singular perturbation analysis are not successfully applicable to stiff problems with a general 'geometry'. To overcome this drawback to a certain extent, in Part I of this paper a problem class was introduced, where the stiffness is axiomatically characterized in natural geometric terms, with a nonlinearly varying 'stiff eigendirection' corresponding to a 'stiff eigenvalue'. Furthermore, a convergence analysis for the implicit Euler scheme was presented.

In the meantime, some work has been done to extend this theory to general Runge-Kutta schemes. In particular, for the class of stiff problems considered in Part I, a complete analysis of Radau and Gauss schemes has been developed. In this paper these results are presented, commented, and the essential facts are proved.

Due to their excellent stability properties, the Radau IIa schemes are of particular interest. Therefore we will work out the theory for these methods in some details, and content ourselves with reporting the corresponding results for Radau Ia and Gauss schemes, which are obtained in an analogous way.

Lack of space prevents us from presenting all the technical details, which can be found in an underlying report. In the proofs given, straightforward but lengthy algebraic manipulations are not always worked out in full detail.

Our error estimates show that for certain stepsize ranges, the classical order of superconvergence can even be obtained in the stiff case. This is a further improvement compared to the B-convergence theory, where only the stage order is maintained in the error bounds.

AMS subject classification: 34A65, 65L05, 65L70.

Key words: Stiff differential equations, Runge-Kutta methods, convergence.

1 Introduction.

In Part I of this series of papers (see [6]) we have described in detail where the motivation for this work comes from. We have discussed the drawbacks

*Received xxxx 2001. Communicated by xxx.
†This work was supported by a grant from the Austrian Science Fund.

of standard approaches for the convergence analysis of discretization schemes applied to initial value problems for systems of *stiff* ODEs

$$y'(t) = f(y(t)), \quad t \in [0,T]. \tag{1.1}$$

In [6] it was argued that the theory of B-convergence (cf. [9],[10],[11]) suffers from the restriction that the resulting a-priori bounds become unrealistically large unless the Jacobian Df is very close to normal. Other conventional approaches like singular perturbation analysis ([12],[14]) assume a particular problem structure, which means that the phase portrait of transient solutions resembles that of a linear constant coefficient problem. In [5] the theory of B-convergence was extended to semilinear problems with a varying stiff spectrum and a smooth nonlinearity.

The approach taken in [6] and in the present paper tries to generalize these concepts and to overcome their drawbacks. Our approach is essentially 'geometric'. We generalize the notions 'stiff eigenvalue' and 'stiff eigendirection' to a nonlinear setting, at the same time allowing the stiff eigendirection to be varying. Under some further natural, geometrically motivated assumptions on the problem (see section 2) this enables us to develop a convergence theory for discretization schemes approximating a smooth solution $\tilde{y}(t)$ of (1.1), with initial value $\tilde{y}(0) = \tilde{y}_0 \in \tilde{\mathcal{M}}$, where $\tilde{\mathcal{M}}$ is a smooth invariant manifold of (1.1). Whereas in [6] the implicit Euler scheme has been considered, we now extend this theory to high-order implicit Runge-Kutta methods of Radau Ia, Radau IIa and Gauss type. Our error bounds are not affected by the stiffness of the problem but only depend on parameters like derivatives of the smooth solution to be followed, which are assumed to be moderate-sized. Note that, throughout, we use the term 'smooth' as a shortcut for 'sufficiently differentiable, with moderate-sized derivatives' not affected by the large size of the stiff eigenvalues $\lambda(\cdot) \ll 0$ in the sense of (2.6).

For a further motivating discussion and examples, cf. [6] (section 1) and [7].

The paper is organized as follows: Section 2 introduces the class of stiff problems considered, together with the necessary notation. In section 3 we review some basic properties of implicit Runge-Kutta schemes, and we represent a Runge-Kutta step in local coordinates. Section 4 is the heart of the paper and contains the convergence analysis. The global error induction is worked out in some technical detail. However, we make use of some a-priori estimates for the solution of the Runge-Kutta equations, for the precise derivation of which we refer to [3]. Section 5 gives an overview on further results from [2] and [3] which are not discussed in detail here. Section 6 contains a review of the stability concepts from the 'classical' theory of B-stability and B-convergence together with a strengthened B-stability estimate, which form an essential basis for our work.

2 Problem class and denotation.

Throughout, $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$ denote the Euclidean inner product and norm in $\mathbf{R}^n$, respectively. At several points in the text, $\| \cdot \|$ also denotes the Euclidean norm in $\mathbf{R}^{n \times s}$. Here, n is the dimension of the underlying ODE system (1.1), and s denotes the stage number of a Runge-Kutta scheme, cf. (3.1). For vectors $(x_1,\ldots,x_n)\in\mathbf{R}^n$ we sometimes use the shorthand notation x or $(x_k)_{k=1}^n$, and analogously for vectors $(X_1,\ldots,X_s)\in\mathbf{R}^{n\times s}$ where $X_i\in\mathbf{R}^n$. The symbol $|||\cdot|||$ represents the norm

$$|||X||| = |||(X_i)_{i=1}^s||| := \|(\|X_1\|,\ldots,\|X_s\|)\|_\infty. \tag{2.1}$$

Whenever the generic symbol $\mathcal{C}$ is used, it represents a quantity that depends in a moderate way on certain parameters of the method considered and on problem-dependent parameters like higher derivatives of a smooth solution of the given ODE, etc. Usually, the particular dependence can be seen from the proofs given, but it would be very inconvenient to write all this down in an explicit way. The essential point behind this generic notation is that, in any case, $\mathcal{C}$ is a quantity *independent of the stiffness*: It is not influenced by the large Lipschitz constant of the right hand side $f(y)$ of the given stiff problem, i.e., it is not affected by the large size of the stiff eigenvalue $\lambda(\cdot)\ll 0$ – cf. (2.3). At some points in the text we use the notation $\hat{\mathcal{C}}$ instead of $\mathcal{C}$ to indicate a particular value for $\mathcal{C}$.

We consider an autonomous ODE (1.1) with $f:\mathcal{G}\to\mathbf{R}^n$ ($\mathcal{G}\subset\mathbf{R}^n$) and make the following assumptions, formalizing the notion of stiffness in natural geometric terms. Some of the smoothness assumptions made below appear quite technical, but they are not really restrictive in the context of a theory of high-order schemes.

(i) *Existence of an invariant manifold* $\tilde{\mathcal{M}}$*:*

We assume that there exists a smooth invariant manifold $\tilde{\mathcal{M}}$ of the given ODE (1.1) containing smooth (non-transient) solutions.

In accordance with in Part I ([6]) we assume that $\tilde{\mathcal{M}}$ is $(n-1)$-dimensional. Concerning the more general case, where the dimension of $\tilde{\mathcal{M}}$ is $n-k$ with $k\geq 1$, cf. the remark at the end of section 5.

(ii) *Behavior of the flow of the ODE (1.1) on* $\tilde{\mathcal{M}}$*:*

$f|_{\tilde{\mathcal{M}}}$ is assumed to be smooth, i.e. there exist moderate constants $\tilde{M}_f$ and $\tilde{L}_f$ such that $\|f|_{\tilde{\mathcal{M}}}\|\leq\tilde{M}_f$, and $f|_{\tilde{\mathcal{M}}}$ is Lipschitz continuous with Lipschitz constant $\tilde{L}_f$. By $\tilde{m}_f$ we denote a one-sided Lipschitz constant for $f|_{\tilde{\mathcal{M}}}$, i.e. a constant $\tilde{m}_f$ such that[1]

$$\langle\check{y}-y, f(\check{y})-f(y)\rangle\leq\tilde{m}_f\,\|\check{y}-y\|^2 \qquad\text{for}\quad y,\check{y}\in\tilde{\mathcal{M}}. \tag{2.2}$$

This satisfies $|\tilde{m}_f|\leq\tilde{L}_f$, and $\tilde{m}_f$ may be nonpositive (if the flow $f|_{\tilde{\mathcal{M}}}$ is dissipative).

Our convergence theory is concerned with the approximation of a smooth solution $\tilde{y}(t)\subset\tilde{\mathcal{M}}$, with initial value $\tilde{y}_0\in\tilde{\mathcal{M}}$. For any smooth solution of (1.1) living in $\tilde{\mathcal{M}}$ we assume that the appropriate number of derivatives exists and is moderately bounded, as required in the local error analysis (see subsection 3.1).

[1] Note that for a stiff problem of the structure described here, the Lipschitz constant L_f *and also* the one-sided Lipschitz constant m_f of f with respect to the whole domain $\mathcal{G}\supset\tilde{\mathcal{M}}$ are very large and positive; cf. the discussion in [4], [6].

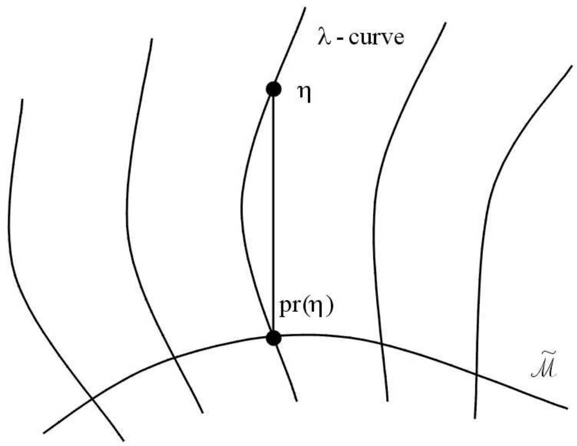

Figure 2.1: Representation (2.7) of $\eta \in \mathcal{G}$

(iii) *Transversality condition:*

The stiffness of the problem is axiomatically characterized in the following way: We assume that, for each point $\eta \in \mathcal{G}\backslash\tilde{\mathcal{M}}$, there exists a locally unique point $p = \mathrm{pr}(\eta) \in \tilde{\mathcal{M}}$ such that the 'transversality condition'

$$f(\eta) - f(p) = \lambda(\eta)(\eta - p) \tag{2.3}$$

holds with $\lambda(\eta) \ll 0$, where the corresponding projection

$$\mathrm{pr} : \mathcal{G}\backslash\tilde{\mathcal{M}} \to \tilde{\mathcal{M}}, \tag{2.4}$$

and the 'stiff eigenvalue'

$$\lambda : \mathcal{G}\backslash\tilde{\mathcal{M}} \to \mathbf{R}^-, \tag{2.5}$$

are continuous and continuously extend to $\tilde{\mathcal{M}}$. For the function $\lambda(\eta)$ we assume

$$\lambda(\eta) = -\frac{1}{\varepsilon} g(\eta) \tag{2.6}$$

with a Lipschitz continuous function $g: \ \mathcal{G} \to [1, \infty)$ and $0 < \varepsilon \ll 1$. Now the projection $\mathrm{pr}(\cdot)$ is used to represent a point $\eta \in \mathcal{G}$ in the form

$$\eta = \mathrm{pr}(\eta) + (\eta - \mathrm{pr}(\eta)) \tag{2.7}$$

with its 'foot' $\mathrm{pr}(\eta) \in \tilde{\mathcal{M}}$ and its 'transversal component' $\eta - \mathrm{pr}(\eta)$. Moreover we assume that, for each $p \in \tilde{\mathcal{M}}$, the set of all points η satisfying a relation of the form (2.3),(2.6) is a smooth curve which we call the 'λ-curve' associated with p:

$$\lambda\text{-curve} := \{p\} \cup \{\eta \in \mathcal{G} : \ f(\eta) - f(p) = \lambda(\eta)(\eta - p)\}. \tag{2.8}$$

(iv) *Transversal coordinate system:*

On the basis of (iii) it is natural to assume that $\mathcal{G}$ can be parametrized by a coordinate system including λ-curves. In particular, let

$$\pi(\eta) := sgn \cdot \frac{\eta - \mathrm{pr}(\eta)}{\|\eta - \mathrm{pr}(\eta)\|} \tag{2.9}$$

be the transversal unit vector, oriented through $sgn = \pm 1$ in a way such that the function $\pi(\eta)$ and its extension to $\tilde{\mathcal{M}}$ is Lipschitz continuous . We denote

$$\begin{aligned} \eta &= \mathrm{pr}(\eta) + d(\eta)\pi(\eta), \\ f(\eta) &= f(\mathrm{pr}(\eta)) + \lambda(\eta)d(\eta)\pi(\eta), \end{aligned} \tag{2.10}$$

where the function $d(\eta) = \langle \eta - \mathrm{pr}(\eta), \pi(\eta)\rangle$ (the transversal coordinate, $d|_{\tilde{\mathcal{M}}} \equiv 0$) is assumed to be injective on λ-curves. It can be shown that under this assumption the mapping $\eta \to (\mathrm{pr}(\eta), d(\eta))$ is also injective (cf. [3]).

(v) *Local parametrization of $\tilde{\mathcal{M}}$:*

For certain fixed $\bar{\eta} \in \mathcal{G}$ we shall use a local parametrization in the following sense:[2] Let $\bar{p} := \mathrm{pr}(\bar{\eta})$ and $\bar{d} := d(\bar{\eta})$, and extend $\bar{\pi} := \pi(\bar{\eta})$ to an orthonormal basis

$$[\hat{\pi}_1, \hat{\pi}_2, \ldots, \hat{\pi}_{n-1}, \bar{\pi}]. \tag{2.11}$$

We assume that, in a neighborhood of $\bar{p}$, the manifold $\tilde{\mathcal{M}}$ can be locally parametrized by a smooth function

$$u(x) := \bar{p} + x_1\hat{\pi}_1 + \cdots + x_{n-1}\hat{\pi}_{n-1} + v(x)\bar{\pi}. \tag{2.12}$$

Here, the coordinates $x = (x_1, \ldots, x_{n-1})$ vary in an appropriate neighborhood of $0 \in \mathbf{R}^{n-1}$, and the real-valued function $v(x)$ is assumed to be Lipschitz continuous.

REMARK 2.1. *The Lipschitz continuity of $v(x)$ is related to the following 'transversality property': Let L_v denote a Lipschitz constant for $v(x)$ and let e be an arbitrary unit vector in the hyperplane tangential to $\tilde{\mathcal{M}}$ at the point* $\mathrm{pr}(\bar{\eta})$. *Then it is easy to show that*

$$|\langle e, \pi(\bar{\eta})\rangle| \le \frac{L_v}{\sqrt{1 + L_v^2}}\,. \tag{2.13}$$

Therefore the angle α between $\pi(\bar{\eta})$ and $\tilde{\mathcal{M}}$ satisfies $\cos\alpha \le L_v/\sqrt{1 + L_v^2}$ and is therefore significantly away from 0 *for moderate L_v.*

Together with (2.10), $u(x)$ from (2.12) yields a local parametrization of $\mathcal{G}$, with the coordinates $(x, d) = (x_1, \ldots, x_{n-1}, d)$; in this sense we write $\eta = \eta(x, d)$, with the 'smooth coordinates' x and the 'stiff coordinate' d.

[2] In our global error induction, we use local parametrizations varying from step to step. We refrain from explicitly referencing the dependence of $u(x)$ and $v(x)$ on $\bar{\eta}$, as this will always be clear from the context.

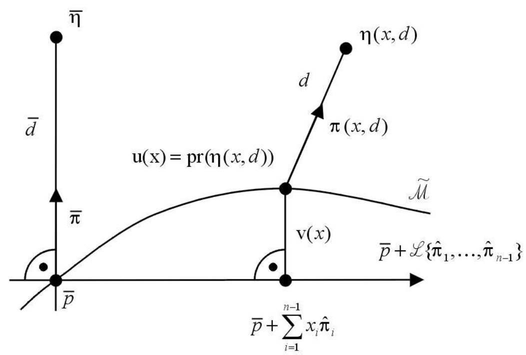

Figure 2.2: Local parametrization of $\tilde{\mathcal{M}}$ and $\mathcal{G}$

For the representation of values of $f|_{\tilde{\mathcal{M}}}$ in the basis (2.11) we introduce the functions $\hat\psi(x) = (\hat\psi_1(x), \ldots, \hat\psi_{n-1}(x))$ and $\psi(x)$ defined by

$$\hat\psi_k(x) := \langle f(u(x)), \hat\pi_k \rangle, \quad k = 1(1)n{-}1, \qquad \psi(x) := \langle f(u(x)), \bar\pi \rangle, \tag{2.14}$$

hence

$$f(u(x)) = \hat\psi_1(x)\hat\pi_1 + \ldots + \hat\psi_{n-1}(x)\hat\pi_{n-1} + \psi(x)\bar\pi. \tag{2.15}$$

Under our smoothness assumptions concerning $\tilde{\mathcal{M}}$ and $f|_{\tilde{\mathcal{M}}}$ it is easy to show that the $\hat\psi_k$ and ψ are Lipschitz continuous with a moderate Lipschitz constant L_ψ.

Finally, we represent the transversal unit vector $\pi(\eta) = \pi(\eta(x,d))$ (cf. (2.9)) with respect to the basis (2.11): With

$$\hat\theta_k(x,d) := \langle \pi(\eta(x,d)), \hat\pi_k \rangle, \ \ k = 1(1)n{-}1, \quad \theta(x,d) := \langle \pi(\eta(x,d)), \bar\pi \rangle, \tag{2.16}$$

we can write

$$\pi(\eta(x,d)) = \hat\theta_1(x,d)\hat\pi_1 + \ldots + \hat\theta_{n-1}(x,d)\hat\pi_{n-1} + \theta(x,d)\bar\pi. \tag{2.17}$$

3 Integration by an implicit Runge-Kutta method.

Now we consider Runge-Kutta methods applied on a grid $[0, \ldots, t_\nu \ldots]$, with stepsize h, i.e., $t_\nu = t_{\nu-1} + h$. For a Runge-Kutta method characterized by its coefficient scheme (the so-called Butcher array)

$$\begin{array}{c|c} c & A \\ \hline & b^T \end{array} = \begin{array}{c|ccc} c_1 & a_{11} & \cdots & a_{1s} \\ \vdots & \vdots & & \vdots \\ c_s & a_{s1} & \cdots & a_{ss} \\ \hline & b_1 & \cdots & b_s \end{array} \tag{3.1}$$

an integration step $\eta_{\nu-1} \to \eta_\nu$ consists in solving the algebraic system

$$Y_i = \eta_{\nu-1} + h \sum_{j=1}^{s} a_{ij} f(Y_j), \quad i = 1(1)s \tag{3.2}$$

for the stages Y_i, followed by

$$\eta_\nu := \eta_{\nu-1} + h \sum_{i=1}^{s} b_i f(Y_i). \tag{3.3}$$

In particular, we shall consider the following classical families of high-order Runge-Kutta schemes of collocation type, with collocation abscissae $t_{\nu-1} + c_i h$. If $P^*(t)$ denotes the Legendre polynomial of degree s transformed to $[0,1]$, the abscissas $c_i, i = 1(1)s$ in (3.1) are defined as follows (cf. [8]):

- *Radau Ia:* c_i = zeros of $P^*_{s-1}(t) + P^*_s(t)$ $(0 = c_1 < \ldots < c_s < 1)$
- *Radau IIa:* c_i = zeros of $P^*_{s-1}(1-t) + P^*_s(1-t)$ $(0 < c_1 < \ldots < c_s = 1)$
- *Gauss:* c_i = zeros of $P^*_s(t)$ $(0 < c_1 < \ldots < c_s < 1)$

For the Radau IIa schemes we have $b_i = a_{si},\ i = 1(1)s$ (cf. [8]).

All these methods satisfy the stability properties in the spirit of the theory of B-convergence (cf. [8],[9],[10]). In particular, they are B-, BS- and BSI-stable (cf. section 6), and we shall make use of these stability properties in section 4.

3.1 Local truncation error, classical order, and stage order.

For a point $\bar{p} \in \tilde{\mathcal{M}}$, we denote by $\tilde{p}(t)$ the solution of (1.1) in $\tilde{\mathcal{M}}$ with initial value[3] $\tilde{p}(0) = \bar{p}$. For such a smooth solution, let $\mathcal{T}_i, i = 1(1)s$, and τ denote its truncation errors (residuals) with respect to the Runge-Kutta equations (3.2),(3.3):

$$\mathcal{T}_i \ := \ \tilde{p}(c_i h) - \bar{p} - h \sum_{j=1}^{s} a_{ij} \underbrace{f(\tilde{p}(c_j h))}_{\tilde{p}'(c_j h)}, \quad i = 1(1)s, \tag{3.4}$$

$$\tau \ := \ \tilde{p}(h) - \bar{p} - h \sum_{i=1}^{s} b_i \underbrace{f(\tilde{p}(c_i h))}_{\tilde{p}'(c_i h)}. \tag{3.5}$$

The classical theory of implicit Runge-Kutta methods (cf. [8],[14]) provides us with asymptotic estimates of these quantities, namely:

$$\|\mathcal{T}_i\| \ \le \ \mathcal{C}(M_\ell) h^{\sigma+1}, \quad i = 1(1)s, \tag{3.6}$$

$$\|\tau\| \ \le \ \mathcal{C}(M_\ell) h^{\rho+1}. \tag{3.7}$$

[3] Here, t is not the original independent variable but has to be understood in a local sense; the initial condition for $\tilde{p}(t)$ is posed at $t = 0$, with the aim to simplify the notation. This use of the symbol t is somewhat sloppy but unproblematic since $f = f(y)$ is autonomous.

Here, $\mathcal{C}(M_\ell)$ denotes certain quantities depending on bounds for a number of derivatives of the smooth solution in $\tilde{\mathcal{M}}$ under consideration:

$$\left\| \frac{d^\ell \tilde{p}(t)}{dt^\ell} \right\| \le M_\ell, \quad \ell = 1(1)\rho+1. \tag{3.8}$$

(Note that these quantities do not depend on derivatives of f like its Jacobian Df, which take large values also on $\tilde{\mathcal{M}}$.) The exponents ρ and σ in the local error estimates (3.6) are called the *classical order* and the *stage order* of the method, respectively. For the methods under consideration we have

$$\rho = \begin{cases} 2s, & \text{Gauss}, \\ 2s-1, & \text{Radau Ia, IIa}, \end{cases} \qquad \sigma = \begin{cases} s, & \text{Gauss, Radau IIa}, \\ s-1, & \text{Radau Ia}. \end{cases} \tag{3.9}$$

(A proof of (3.6)–(3.9) can also be found in [3], p. 48.)

In our convergence theory, we naturally assume that all solutions living in $\tilde{\mathcal{M}}$ satisfy the smoothness requirements (3.8).

3.2 Local parametrization of a smooth solution

Now we use the local parametrization of $\tilde{\mathcal{M}}$ introduced in section 2 (assumption (v)), i.e., a parametrization corresponding to a fixed $\bar{p} = \mathrm{pr}(\bar{\eta})$. In particular, let the function $\tilde{x}(t) = (\tilde{x}_1(t), \ldots, \tilde{x}_{n-1}(t))$ be defined by the relation $u(\tilde{x}(t)) = \tilde{p}(t)$, with $u(x)$ from (2.12) and $\tilde{p}(t)$ as defined in subsection 3.1. With (2.15), we have

$$\begin{aligned} \tilde{p}'(t) \;&=\; \tilde{x}_1'(t)\hat{\pi}_1 + \ldots + \tilde{x}_{n-1}'(t)\hat{\pi}_{n-1} + (v\circ\tilde{x})'(t)\bar{\pi} \\ = f(\tilde{p}(t)) \;&=\; \hat{\psi}_1(\tilde{x}(t))\hat{\pi}_1 + \ldots + \hat{\psi}_{n-1}(\tilde{x}(t))\hat{\pi}_{n-1} + \psi(\tilde{x}(t))\bar{\pi}. \end{aligned} \tag{3.10}$$

In particular, the $\tilde{x}_k(t)$ satisfy the differential equations

$$\tilde{x}_k'(t) = \hat{\psi}_k(\tilde{x}(t)), \quad k = 1(1)n-1, \tag{3.11}$$

or, in vector form, $\tilde{x}'(t) = \hat{\psi}(\tilde{x}(t))$.

3.3 Local parametrization of the Runge-Kutta equations

The Runge-Kutta equations (3.2) are now formulated in terms of local stiff and smooth coordinates in the sense of the splitting (2.10), and for the analysis of a Runge-Kutta step $\eta_{\nu-1} \to \eta_\nu$ we shall use the parametrization (2.12) with $\bar{\eta} := \eta_{\nu-1}$. As far as we are concerned with local considerations, we shall not explicitly write down the indices $\nu-1, \nu$, but use the shorthand notation

$$\begin{aligned} \bar{p} &:= \mathrm{pr}(\bar{\eta}), & \bar{d} &:= d(\bar{\eta}), & \bar{\pi} &:= \pi(\bar{\eta}), \\ P_i &:= \mathrm{pr}(Y_i), & D_i &:= d(Y_i), & \Pi_i &:= \pi(Y_i). \end{aligned} \tag{3.12}$$

Furthermore let $\Lambda_i := \lambda(Y_i)$ for $i = 1(1)s$. The result of a Runge-Kutta step starting from $\bar{\eta}$ will simply be written as η, and we denote

$$p := \mathrm{pr}(\eta), \quad d := d(\eta), \quad \pi := \pi(\eta). \tag{3.13}$$

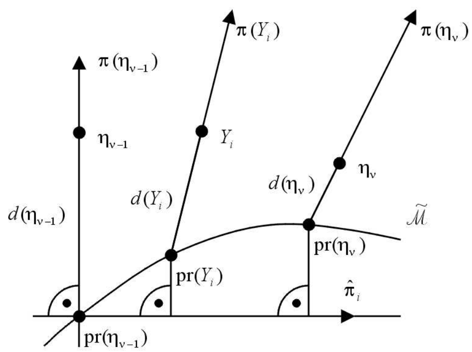

Figure 3.1: A Runge-Kutta step

With this notation, we have

$$\bar\eta = \bar p + \bar d\,\bar\pi, \qquad \eta = p + d\pi, \qquad Y_i = P_i + D_i\Pi_i, \qquad f(Y_i) = f(P_i) + \Lambda_i D_i \Pi_i, \tag{3.14}$$

and the Runge-Kutta equations (3.2),(3.3) take the form

$$P_i + D_i\Pi_i = \bar p + \bar d\,\bar\pi + h\sum_{j=1}^{s} a_{ij}(f(P_j) + \Lambda_j D_j \Pi_j), \quad i = 1(1)s, \tag{3.15}$$

$$p + d\pi = \bar p + \bar d\bar\pi + h\sum_{i=1}^{s} b_i(f(P_i) + \Lambda_i D_i \Pi_i). \tag{3.16}$$

Now, using local coordinates, we write the $P_i \in \tilde{\mathcal{M}}$ as

$$P_i = u(X_i) = \bar p + X_{i,1}\hat\pi_1 \ldots + X_{i,n-1}\hat\pi_{n-1} + V_i\bar\pi, \quad i = 1(1)s, \tag{3.17}$$

with the coordinate vector $X_i = (X_{i,1}, \ldots, X_{i,n-1})$ in the sense of (2.12), and $V = (V_1, \ldots, V_s)$ with $V_i := v(X_i)$. The coordinates of $f(P_i) = f(u(X_i))$ with respect to the basis (2.11) in the sense of (2.14),(2.15) are denoted as $\hat\Psi_i := \hat\psi(X_i)$, $\Psi_i := \psi(X_i)$, and the coordinates of the $\Pi_i = \pi(\eta(X_i, D_i))$ according to (2.16),(2.17) are denoted as $\hat\Theta_i = (\hat\Theta_{i,1}, \ldots, \hat\Theta_{i,n-1}) := \hat\theta(X_i, D_i)$ and $\Theta_i := \theta(X_i, D_i)$. Now the Runge-Kutta equations (3.15) can be written in local coordinate form as

$$X_i + D_i\hat\Theta_i = h\sum_{j=1}^{s} a_{ij}(\hat\Psi_j + \Lambda_j D_j \hat\Theta_j), \tag{3.18}$$

$$V_i + D_i\Theta_i = \bar d + h\sum_{j=1}^{s} a_{ij}(\Psi_j + \Lambda_j D_j \Theta_j), \tag{3.19}$$

with the unknowns (X_i, D_i), $i = 1(1)s$, or equivalently,

$$X + (\mathrm{diag}(D) \otimes I_{n-1})\hat\Theta = h(A \otimes I_{n-1})(\hat\Psi + (\Lambda\,\mathrm{diag}(D) \otimes I_{n-1})\hat\Theta), \tag{3.20}$$

$$V + \Theta D = \bar d\,\mathbf{1} + hA(\Psi + \Lambda\Theta D). \tag{3.21}$$

(Here, $\otimes$ is the Kronecker product, and we denote $\mathbf{1} := (1,\dots,1) \in \mathbf{R}^s$, $\Lambda := \operatorname{diag}(\Lambda_1,\dots,\Lambda_s) \in \mathbf{R}^{s\times s}$ and $\Theta := \operatorname{diag}(\Theta_1,\dots,\Theta_s) \in \mathbf{R}^{s\times s}$.) Relation (3.16) transforms into

$$x + d\,\hat{\theta} = h\sum_{i=1}^{s} b_i(\hat{\Psi}_j + \Lambda_i D_i \hat{\Theta}_i), \tag{3.22}$$

$$v + d\,\theta = \bar{d} + h\sum_{i=1}^{s} b_i(\Psi_i + \Lambda_i D_i \Theta_i), \tag{3.23}$$

where $\hat{\theta} := \hat{\theta}(x,d)$, $\theta := \theta(x,d)$, $v := v(x)$.

REMARK 3.1. *On the basis of the above transformation of a Runge-Kutta step into local coordinates, the solvability of the algebraic Runge-Kutta equations was studied in [3]. Locally unique solvability was proven under a mild stepsize restriction, together with a-priori estimates of the form* $\|(X,D)\|, \|(x,d)\| \leq \mathcal{C}\max\{h, |\bar{d}|\}$.

4 Convergence analysis for Radau and Gauss schemes.

For the global error induction, the Runge-Kutta scheme is written as

$$e_\nu := \eta_\nu - \tilde{y}(t_\nu) = (\eta_\nu - \operatorname{pr}(\eta_\nu)) + (\operatorname{pr}(\eta_\nu) - \tilde{y}(t_\nu)), \tag{4.1}$$

with

$$\eta_\nu - \operatorname{pr}(\eta_\nu) = d_\nu \pi(\eta_\nu) \ \dots \ \text{'stiff error component', with } d_\nu := d(\eta_\nu), \tag{4.2}$$

$$\operatorname{pr}(\eta_\nu) - \tilde{y}(t_\nu) \ \dots \ \text{'smooth error component'}. \tag{4.3}$$

The smooth error component will further be split according to

$$\operatorname{pr}(\eta_\nu) - \tilde{y}(t_\nu) = (\operatorname{pr}(\eta_\nu) - \tilde{p}_{\nu-1}(t_\nu)) + (\tilde{p}_{\nu-1}(t_\nu) - \tilde{y}(t_\nu)), \tag{4.4}$$

where $\tilde{p}_{\nu-1}(t)$ denotes a smooth solution of the given ODE living in $\tilde{\mathcal{M}}$, namely that with initial value $\tilde{p}_{\nu-1}(t_{\nu-1}) = \operatorname{pr}(\eta_{\nu-1})$.

We consider a situation where the Runge-Kutta scheme follows a smooth solution $\tilde{y}(t)$ living in $\tilde{\mathcal{M}}$. Sufficiently small initial perturbations are admitted.[4] In particular, we assume that the initial error satisfies

$$\|\eta_0 - \tilde{y}(0)\| \leq \mathcal{C}h^q \tag{4.5}$$

with a certain exponent q, and for the stiff error component d_0 an assumption will be made which allows the separate inductive estimation of the d_ν for $\nu > 0$ (for the details see subsection 4.4). As always in the context of implicit methods, there is a restriction on the size of stepsizes admitted, $h \leq h_{max}$. In our error analysis, such restrictions appear at several places, but with a bound h_{max} which is not restrictively small and not affected by the stiffness. We call this a 'mild stepsize restriction'.

In the sequel, the notation in the spirit of section 6, $\Phi_B(\cdot)$, $\Phi_{BS}(\cdot)$, $\Phi_{BSI}(\cdot)$ is used for the stability functions of the underlying Runge-Kutta scheme.

[4] In particular, this allows an inductive application of the convergence argument over several intervals with different stepsizes, such that the case of nonequidistant grids is covered; cf [6].

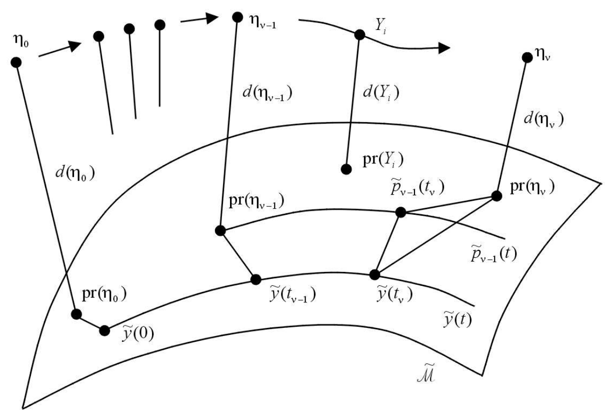

Figure 4.1: Induction for the global error of a Runge-Kutta scheme

4.1 Estimation of the stiff error component.

As a first step in our global error analysis we provide an inductive estimate for $|d(\eta_\nu)|$, the norm of the stiff error component (4.2). We use the local parametrization (2.12) with respect to $\bar{\eta} = \eta_{\nu-1}$ and the local denotation from subsection 3.3 for a Runge-Kutta step $\bar{\eta} \to \eta$, omitting indices $\nu-1, \nu$ (cf. (3.12)–(3.14)).

Proceeding from (3.15),(3.16), we write

$$D_i\Pi_i = \bar{d}\,\bar{\pi} + h\sum_{j=1}^{s} a_{ij}\Lambda_j D_j \Pi_j - \Xi_i, \quad i = 1(1)s, \tag{4.6}$$

$$\text{with} \qquad \Xi_i := P_i - \bar{p} - h\sum_{j=1}^{s} a_{ij} f(P_j),$$

and

$$d\,\pi = \bar{d}\,\bar{\pi} + h\sum_{i=1}^{s} b_i\Lambda_i D_i\Pi_i - \chi, \tag{4.7}$$

$$\text{with} \qquad \chi := p - \bar{p} - h\sum_{i=1}^{s} b_i f(P_i).$$

Taking the inner product of (4.6) and (4.7) with π yields

$$D_i\langle\Pi_i,\pi\rangle = \bar{d}\langle\bar{\pi},\pi\rangle + h\sum_{j=1}^{s} a_{ij}\Lambda_j D_j\langle\Pi_j,\pi\rangle - \langle\Xi_i,\pi\rangle, \quad i = 1(1)s, \tag{4.8}$$

$$d = \bar{d}\langle\bar{\pi},\pi\rangle + h\sum_{i=1}^{s} b_i\Lambda_i D_i\langle\Pi_i,\pi\rangle - \langle\chi,\pi\rangle. \tag{4.9}$$

(4.8),(4.9) represent a perturbed Runge-Kutta step $\bar{d}\langle\bar{\pi},\pi\rangle \to d$, with internal stages $D_i\langle\Pi_i,\pi\rangle$, for a scalar differential equation of the type (6.7), with $\lambda(\cdot) \leq -\frac{1}{\varepsilon}$; cf. assumption (iii), (2.6).

In order to estimate $|d|$, we consider the auxiliary schemes

$$\check{D}_i = \bar{d}\langle\bar{\pi},\pi\rangle + h\sum_{j=1}^{s} a_{ij}\Lambda_j\check{D}_j, \quad 0 = 0 + h\sum_{j=1}^{s} a_{ij}\Lambda_j 0, \quad i = 1(1)s, \tag{4.10}$$

$$\check{d} = \bar{d}\langle\bar{\pi},\pi\rangle + h\sum_{i=1}^{s} b_i\Lambda_i\check{D}_i, \qquad 0 = 0 + h\sum_{i=1}^{s} b_i\Lambda_i 0. \tag{4.11}$$

Together with the B- and BS-stability of the underlying Runge-Kutta scheme (cf. section 6), (4.8),(4.9) and (4.10),(4.11) lead us to the following estimate for d in terms of $\bar{d}$ and the Ξ_i and χ:

$$\begin{aligned} |d| &\leq |d-\check{d}| + |\check{d}-0| \\ &\leq \Phi_{BS}(-\tfrac{h}{\varepsilon})(|\langle\chi,\pi\rangle| + \|(\langle\Xi_i,\pi\rangle)_{i=1}^{s}\|) + \Phi_B(-\tfrac{h}{\varepsilon})\,\bar{d}\langle\bar{\pi},\pi\rangle \\ &\leq \Phi_B(-\tfrac{h}{\varepsilon})\,|\bar{d}| + \Phi_{BS}(-\tfrac{h}{\varepsilon})(\|\chi\| + \sqrt{s}\,|||\Xi|||). \end{aligned} \tag{4.12}$$

In order to proceed, estimates for the auxiliary quantities Ξ_i and χ are required.

Lemma 4.1. *The quantities Ξ_i and χ from (4.6),(4.7) satisfy the inequalities*

$$\|\chi\| \leq \mathcal{C}\,\|x-\tilde{x}(h)\| + \mathcal{C}h\,|||(X_i-\tilde{x}(c_ih))_{i=1}^{s}||| + \mathcal{C}h^{\rho+1}, \tag{4.13}$$

$$\|\Xi_i\| \leq \mathcal{C}\,|||(X_j-\tilde{x}(c_jh))_{j=1}^{s}||| + \mathcal{C}h^{\sigma+1}, \quad i = 1(1)s, \tag{4.14}$$

where ρ resp. σ are the classical order and stage order of the underlying Runge-Kutta scheme (cf. subsection 3.1).

Proof. Using the definition of the local truncation error τ (cf. (3.5)) we estimate $\|\chi\|$ in the following way:

$$\begin{aligned} \|\chi\| &= \|\,p-\tilde{p}(h) - h\sum_{i=1}^{s} b_i(f(P_i)-f(\tilde{p}(c_ih))) + \tau\,\| \\ &\leq \|p-\tilde{p}(h)\| + h\|b\|_1\,|||(f(P_i)-f(\tilde{p}(c_ih)))_{i=1}^{s}||| + \|\tau\|. \end{aligned}$$

Now (4.13) follows from (3.7) and from the estimates

$$\begin{aligned} \|p-\tilde{p}(h)\| &\leq \mathcal{C}\,\|x-\tilde{x}(h)\|, \\ |||(f(P_i)-f(\tilde{p}(c_ih)))_{i=1}^{s}||| &\leq \mathcal{C}\,|||(X_i-\tilde{x}(c_ih))_{i=1}^{s}|||, \end{aligned}$$

which are a consequence of our smoothness assumptions concerning the function $u(x)$ (see section 2, assumption (v)) and of the smoothness of $f|_{\tilde{\mathcal{M}}}$.

With the definition of the $\mathcal{T}_i$ (cf. (3.4)) and (3.6), the $\|\Xi_i\|$ can be estimated in an analogous way, yielding (4.14). □

As the next step in the estimation of $|d|$, bounds for $x-\tilde{x}(h)$ and the $X_i-\tilde{x}(c_ih)$ are required. To this end we need the following a-priori bounds (cf. the remark at the beginning of section 5):

LEMMA 4.2. *The quantities*[5] $D = (D_1, \ldots, D_s)$, $\hat\Theta = (\hat\Theta_1, \ldots, \hat\Theta_s)$ *and* $\hat\theta$ *satisfy the estimates*

$$\|D\|_\infty \le \mathcal{C}\,\Phi_{BSI}(-\tfrac{h}{\varepsilon})(|\bar d| + |||\Xi|||), \tag{4.15}$$

$$\|h\Lambda D\|_\infty \le \mathcal{C}(|\bar d| + |||\Xi|||), \tag{4.16}$$

$$\left.\begin{matrix} |||\hat\Theta||| \\ \|\hat\theta\| \end{matrix}\right\} \le \mathcal{C}\max\{h, |\bar d|\}. \tag{4.17}$$

PROOF. *(sketch)* We proceed from the local parametrization of the Runge-Kutta equations and the notation introduced in subsection 3.3. In [3] it was shown that under a mild a-priori restriction on $|\bar d|$ the $s \times s$-matrices $\Theta = \mathrm{diag}(\Theta_1, \ldots, \Theta_s)$ and $\Gamma := I_s - hA\Lambda$ (with $\Lambda = \mathrm{diag}(\Lambda_1, \ldots, \Lambda_s)$) are invertible and satisfy

$$\|\Theta\|_\infty \le \mathcal{C}, \quad \|\Theta^{-1}\|_\infty \le \mathcal{C}, \qquad \|\Gamma^{-1}\|_\infty \le \mathcal{C}\,\Phi_{BSI}(-\tfrac{h}{\varepsilon}).$$

Therefore we may rewrite (3.21) as

$$D = \Theta^{-1}\Gamma^{-1}(\bar d\,\mathbf{1} - V + hA\Psi),$$

and together with

$$\|V - hA\Psi\|_\infty = \|(\langle \Xi_i, \bar\pi\rangle)_{i=1}^s\|_\infty \le |||\Xi|||$$

(which follows from (3.21) and (4.6)) we obtain (4.15).

Furthermore, we write $h\Lambda D$ in the form[6]

$$h\Lambda D = \Theta^{-1}A^{-1}(\Theta D - (\bar d\,\mathbf{1} - V + hA\Psi)),$$

and together with (4.15) and (6.8) the estimate (4.16) easily follows.

(4.17) can be shown on the basis of the a-priori bounds

$$\|(X, D)\| \le \mathcal{C}, \qquad \|(x, d)\| \le \mathcal{C}\max\{h, |\bar d|\},$$

cf. [3]. □

LEMMA 4.3. *The quantities* $X_i - \tilde x(c_i h)$ *and* $x - \tilde x(h)$ *satisfy the estimates*

$$\|X_i - \tilde x(c_i h)\| \le \mathcal{C}(|\bar d| + |||\Xi|||)\cdot\max\{h, |\bar d|\} + \mathcal{C}h^{\sigma+1}, \quad i = 1(1)s, \tag{4.18}$$

$$\|x - \tilde x(h)\| \le \mathcal{C}(|\bar d| + |d| + |||\Xi|||)\cdot\max\{h, |\bar d|\} + \mathcal{C}(h^{\rho+1} + h^{\sigma+2}). \tag{4.19}$$

PROOF. From (3.10),(3.11) it follows that

$$\|\tilde x(c_i h) - h\sum_{j=1}^s a_{ij}\hat\Psi_j(\tilde x(c_j h))\| \le \|\mathcal{T}_i\|, \quad i = 1(1)s,$$

[5] For the definition of these quantities cf. subsection 3.3.

[6] For the Gauss and Radau Ia, IIa methods the coefficient matrix A is invertible; cf. [14], section IV.14.

($\mathcal{T}_i$ from (3.4)), and together with (3.18) and (4.15),(4.16) we can estimate:

$$\begin{aligned}
\|X_i - \tilde{x}(c_i h)\| &= \| \, h \sum_{j=1}^{s} a_{ij} \hat{\Psi}_j - D_i \hat{\Theta}_i + h \sum_{j=1}^{s} a_{ij} \Lambda_j D_j \hat{\Theta}_j - \tilde{x}(c_i h) \, \| \\
&= \| \, h \sum_{j=1}^{s} a_{ij} (\hat{\Psi}_j - \hat{\psi}(\tilde{x}(c_j h)) - (\tilde{x}(c_i h) - h \sum_{j=1}^{s} a_{ij} \hat{\psi}(\tilde{x}(c_j h))) \\
&\qquad - D_i \hat{\Theta}_i + h \sum_{j=1}^{s} a_{ij} \Lambda_j D_j \hat{\Theta}_j \, \| \\
&\leq h \|A\|_\infty L_\psi |||(X_j - \tilde{x}(c_j h))_{j=1}^s||| + |||\mathcal{T}||| \\
&\qquad + (\, \|D\|_\infty + \|A\|_\infty \|h \Lambda D\|_\infty \,) |||\hat{\Theta}|||.
\end{aligned}$$

Under the mild stepsize restriction $h \leq 1/(2\|A\|_\infty L_\psi)$ this yields

$$\begin{aligned}
\|X_i - \tilde{x}(c_i h)\| &\leq 2 \, (\, |||\mathcal{T}||| + (\|D\|_\infty + \|A\|_\infty \|h \Lambda D\|_\infty) h |||\hat{\Theta}||| \,) \\
&\leq 2 \, \mathcal{C} h^{\sigma+1} + 2 \, (\, \mathcal{C} \, \Phi_{BSI}(-\tfrac{h}{\varepsilon})(|\bar{d}| + |||\Xi|||) + \\
&\qquad + \|A\|_\infty \mathcal{C}(|\bar{d}| + |||\Xi|||) \,) \cdot \mathcal{C} \max\{h, |\bar{d}|\}.
\end{aligned}$$

Together with $\Phi_{BSI}(-\frac{h}{\varepsilon}) \leq \mathcal{C}$ (cf. (6.8)) this implies (4.18).

An analogous reasoning, with

$$\|\tilde{x}(h) - h \sum_{i=1}^{s} b_i \hat{\Psi}(\tilde{x}(c_i h))\| \leq \|\tau\|,$$

(τ from (3.5)), yields

$$\begin{aligned}
\|x - \tilde{x}(h)\| &\leq h \|b\|_1 L_\psi |||(X_i - \tilde{x}(c_i h))_{i=1}^s||| + \mathcal{C} h^{\rho+1} \\
&\qquad + |d| \|\hat{\theta}\| + \|b\|_1 \|h \Lambda D\|_\infty |||\hat{\Theta}|||,
\end{aligned}$$

and together with (4.17) and (4.18) we obtain

$$\begin{aligned}
\|x - \tilde{x}(h)\| &\leq h \|b\|_1 L_\psi (\, \mathcal{C}(|\bar{d}| + |||\Xi|||) \cdot \max\{h, |\bar{d}|\} + \mathcal{C} h^{\sigma+1} \,) + \mathcal{C} h^{\rho+1} \\
&\qquad + \mathcal{C} |d| \, \max\{h, |\bar{d}|\} + \|b\|_1 \, \mathcal{C}(|\bar{d}| + \|\Xi|||) \max\{h, |\bar{d}|\},
\end{aligned}$$

which yields (4.19). ☐

Now we can provide an induction for the size of the stiff error component.

LEMMA 4.4. *The stiff error component d satisfies the inductive estimate*

$$\begin{aligned}
|d| &\leq (1 + \mathcal{C} \, \Phi_{BS}(-\tfrac{h}{\varepsilon}) h) \cdot \Phi_B(-\tfrac{h}{\varepsilon}) |\bar{d}| \\
&\qquad + \Phi_{BS}(-\tfrac{h}{\varepsilon})(\mathcal{C} \, |\bar{d}|^2 + \mathcal{C} \, |\bar{d}| h + \mathcal{C} h^{\sigma+1}).
\end{aligned} \tag{4.20}$$

PROOF. Inserting (4.18) into (4.14) we obtain

$$\begin{aligned}
\|\Xi_i\| &\leq \mathcal{C}(\mathcal{C}(|\bar{d}| + |||\Xi|||) \cdot \max\{h, |\bar{d}|\} + \mathcal{C} h^{\sigma+1}) + \mathcal{C} h^{\sigma+1} \\
&\leq \hat{\mathcal{C}}((|\bar{d}| + |||\Xi|||) \cdot \max\{h, |\bar{d}|\} + h^{\sigma+1}), \quad i = 1(1)s.
\end{aligned}$$

Under the mild a priori restriction $h, |\bar{d}| \le 1/(2\hat{\mathcal{C}})$ this yields

$$\|\!|\Xi|\!\|\le \mathcal{C}(|\bar{d}|\max\{h,|\bar{d}|\}+h^{\sigma+1}). \tag{4.21}$$

Furthermore, inserting (4.18),(4.19) into (4.13), using (4.21) and rearranging terms leads us to

$$\|\chi\| \le \mathcal{C}(|d|+|\bar{d}|+\|\!|\Xi|\!\|)\cdot\max\{h,|\bar{d}|\}+\mathcal{C}(h^{\rho+1}+h^{\sigma+2}). \tag{4.22}$$

Now we insert (4.21),(4.22) into (4.12) and obtain

$$|d| \le \Phi_B(-\tfrac{h}{\varepsilon})\,|\bar{d}| + \Phi_{BS}(-\tfrac{h}{\varepsilon})\,\hat{\mathcal{C}}\,((|d|+|\bar{d}|)\cdot\max\{h,|\bar{d}|\}+h^{\sigma+1}).$$

Under the mild a priori restriction $h, |\bar{d}| \le 1/(2\hat{\mathcal{C}}\Phi_{BS}(-\frac{h}{\varepsilon}))$ this yields

$$\begin{aligned} |d| \quad \le \quad & (\,\Phi_B(-\tfrac{h}{\varepsilon})\,|\bar{d}| + \Phi_{BS}(-\tfrac{h}{\varepsilon})(\mathcal{C}|\bar{d}|\max\{h,|\bar{d}|\}+\mathcal{C}h^{\sigma+1})\,)\\ & \cdot(1+2\Phi_{BS}(-\tfrac{h}{\varepsilon})\,\mathcal{C}\,\max\{h,|\bar{d}|\}). \end{aligned}$$

Together with $\max\{h,|\bar{d}|\}\le h+|\bar{d}|$, this leads to (4.20). □

4.2 The stiff error component for the case Radau IIa.

In the following we restrict our considerations to the Radau IIa scheme. In the estimates derived in subsection 4.1, the stability functions $\Phi_{BS}(-\frac{h}{\varepsilon})$ and $\Phi_{BSI}(-\frac{h}{\varepsilon})$ have played an essential role. Now we also need the special form of the B-stability function $\Phi_B(-\frac{h}{\varepsilon})$ from (6.10), which we show to be applicable for the inductive estimation of the d_ν.

LEMMA 4.5. *Assume that the initial stiff error component $d_0 = d(\eta_0)$ satisfies*

$$|d_0| \le K\varepsilon h^s \tag{4.23}$$

with a constant $K \in [K_{min}, \frac{1}{\varepsilon}]$ and $h \le h_{max}$. Then the inequality (4.23) is valid for all ν, i.e.,

$$|d_\nu| \le K\varepsilon h^s, \quad \nu > 0 \tag{4.24}$$

with the same constant K as in (4.23). (The admitted range for K_{min} is seen in the proof below.)

PROOF. The proof is by induction $\bar{d}\to d$ (i.e., $d_{\nu-1}\to d_\nu$).

First, if we define a function $\gamma(t)$ by interpolation such that

$$\gamma(t_{\nu-1}) = g(\bar{\eta}) \quad \text{and} \quad \gamma(t_{\nu-1}+c_i h) = g(Y_i), \quad i = 1(1)s,$$

then (4.8),(4.9) can be interpreted as a perturbed Runge-Kutta step for a linear scalar ODE $y'(t) = -(\gamma(t)/\varepsilon)\,y(t)$, with $\gamma(t)\ge 1$ for $t = t_{\nu-1}$ and $t = t_{\nu-1}+c_i h$. Furthermore we can estimate

$$\begin{aligned} &\left|\frac{\gamma(t_{\nu-1}+c_i h)-\gamma(t_{\nu-1})}{\gamma(t_{\nu-1})}\right| = \left|\frac{g(Y_i)-g(\bar{\eta})}{g(\bar{\eta})}\right|\\ \le\;& \mathcal{C}\|(X_i,D_i)-(0,\bar{d})\| \le \mathcal{C}\max\{h,|\bar{d}|\}+|\bar{d}|\} \le \mathcal{C}\max\{h,|\bar{d}|\}, \end{aligned}$$

and therefore, under a mild stepsize restriction and the inductive assumption (4.24) for $\bar{d} = d_{\nu-1}$, the assumptions of Lemma 6.4 are satisfied. Thus, $\Phi_{BS}(-\frac{h}{\varepsilon})$ and $\Phi_B(-\frac{h}{\varepsilon})$ satisfy estimates of the form (6.9) and (6.10), and together with Lemma 4.4, (4.20), this easily leads us to

$$|d| \leq \frac{1}{1+\phi\frac{h}{\varepsilon}}\,(\,(1+\mathcal{C}h)|\bar{d}| + \mathcal{C}|\bar{d}|^2 + Ch^{s+1}\,)$$

with $\phi := \min\{\phi_B, \phi_{BS}\}$. Now we use the inductive assumption $|\bar{d}| \leq K\varepsilon h^s$ and obtain (again under a mild stepsize restriction)

$$|d| \leq \frac{K\varepsilon h^s}{1+\phi\frac{h}{\varepsilon}}\Big(1+\mathcal{C}h+\mathcal{C}\underbrace{\varepsilon K}_{\leq 1}h^s + \frac{\mathcal{C}}{K}\frac{h}{\varepsilon}\Big) \leq \frac{K\varepsilon h^s}{1+\phi\frac{h}{\varepsilon}}\Big(1+\phi\frac{h}{\varepsilon}\,(\hat{\mathcal{C}}\varepsilon + \frac{\hat{\mathcal{C}}}{K})\Big)$$

with a certain moderate-sized constant $\hat{\mathcal{C}}$. If K satisfies $2\hat{\mathcal{C}} \leq K \leq 1/\varepsilon$, we have $\hat{\mathcal{C}}\varepsilon \leq 1/2$ and $\hat{\mathcal{C}}/K \leq 1/2$ and obtain

$$|d| \leq K\varepsilon h^s.$$

Therefore the estimate (4.24) is satisfied if we choose K_{min} as $2\times$ the maximum of the values for $\hat{\mathcal{C}}$ occurring for $\nu = 1, 2, \ldots$. □

The proof shows that K is not an arbitrary constant but lower and upper bounds are required.

4.3 Estimation of the smooth error component for Radau IIa.

After the preparations in the previous subsections, it is now easy to derive global error bounds. At first we prove:

LEMMA 4.6. *For the s-stage Radau IIa scheme, the smooth error component* $\mathrm{pr}(\eta_\nu) - \tilde{y}(t_\nu)$ *satisfies the inductive estimate*

$$\|\mathrm{pr}(\eta_\nu) - \tilde{y}(t_\nu)\| \leq e^{h\tilde{m}_f}\,\|\mathrm{pr}(\eta_{\nu-1}) - \tilde{y}(t_{\nu-1})\| + \mathcal{C}(h+\varepsilon)h^{s+1}. \tag{4.25}$$

PROOF. Lemma 4.3 provides a local estimate for the x-coordinates of the smooth error component, $x - \tilde{x}(h)$. This depends on the quantities $\bar{d}$, d and Ξ, for which estimates are now available: Due to Lemma 4.5 we know that $|\bar{d}|$ and $|d|$ are bounded by $\mathcal{C}\varepsilon h^s$, and therefore (4.21) (from the proof of Lemma 4.4) yields

$$\begin{aligned} |||\Xi||| &\leq \mathcal{C}(|\bar{d}|\max\{h, |\bar{d}|\} + h^{\sigma+1}) \\ &\leq \mathcal{C}(\,\mathcal{C}\varepsilon h^s \max\{h, \mathcal{C}\varepsilon h^s\} + h^{s+1}\,) \leq \mathcal{C}h^{s+1}. \end{aligned}$$

Inserting these estimates into (4.19) (Lemma 4.3) we obtain (cf. (3.9))

$$\begin{aligned} \|x - \tilde{x}(h)\| &\leq \mathcal{C}(|\bar{d}| + |d| + |||\Xi|||)\cdot\max\{h, |\bar{d}|\} + \mathcal{C}(h^{s+2} + h^{2s}) \\ &\leq \mathcal{C}(\,\mathcal{C}\varepsilon h^s + \mathcal{C}\varepsilon h^s + \mathcal{C}h^{s+1}\,)\cdot\max\{h, \mathcal{C}\varepsilon h^s\} + \mathcal{C}(h^{s+2} + h^{2s}) \\ &\leq \mathcal{C}(h+\varepsilon)h^{s+1}, \end{aligned}$$

and thus,

$$\|p - \tilde{p}(h)\| \le \mathcal{C}(h+\varepsilon)h^{s+1}, \quad \text{i.e.:} \quad \|\mathrm{pr}(\eta_\nu) - \tilde{p}_{\nu-1}(t_\nu)\| \le \mathcal{C}(h+\varepsilon)h^{s+1}.$$

Now we use the representation (4.4) and apply the one-sided Lipschitz continuity of $f|_{\tilde{\mathcal{M}}}$ in the usual way to estimate the norm of the difference of the smooth solutions $\tilde{p}_{\nu-1}(t) \subset \tilde{\mathcal{M}}$ and $\tilde{y}(t) \subset \tilde{\mathcal{M}}$ along the time interval $[t_{\nu-1}, t_\nu]$. This yields

$$\begin{aligned} \|\mathrm{pr}(\eta_\nu) - \tilde{y}(t_\nu)\| &\le \|\mathrm{pr}(\eta_\nu) - \tilde{p}_{\nu-1}(t_\nu))\| + \|\tilde{p}_{\nu-1}(t_\nu) - \tilde{y}(t_\nu))\| \\ &\le \mathcal{C}(h+\varepsilon)h^{s+1} + e^{h\tilde{m}_f}\|\tilde{p}_{\nu-1}(t_{\nu-1}) - \tilde{y}(t_{\nu-1}))\|. \end{aligned}$$

Since, by definition of $\tilde{p}_{\nu-1}(t)$, we have $\tilde{p}_{\nu-1}(t_{\nu-1}) = \mathrm{pr}(\eta_{\nu-1})$, (4.25) follows. □

4.4 Convergence results.

As a consequence of the above considerations we can now formulate the following theorem, which generalizes the results from [6]:

THEOREM 4.7. *Let the assumptions (i)–(v) (cf. section 2) be satisfied and assume that $\tilde{y}(t) \subset \tilde{\mathcal{M}}$ is a sufficiently smooth solution of (1.1). Let the initial error satisfy*

$$\begin{aligned} \|\mathrm{pr}(\eta_0) - \tilde{y}_0\| &\le \mathcal{C}(h+\varepsilon)h^s, \\ |d(\eta_0)| &\le K\varepsilon h^s, \end{aligned}$$

with a constant[7] $K \in [K_{min}, \frac{1}{\varepsilon}]$. *Then, under a mild stepsize restriction $h \le h_{max}$, the Radau IIa method is convergent and satisfies the global error estimate*

$$\begin{aligned} \|\eta_\nu - \tilde{y}(t_\nu)\| &\le \mathcal{C}(h+\varepsilon)h^s, \\ \|\mathrm{pr}(\eta_\nu) - \tilde{y}(t_\nu)\| &\le \mathcal{C}(h+\varepsilon)h^s, \\ |d(\eta_\nu)| &\le K\varepsilon h^s, \end{aligned} \tag{4.26}$$

with the same constant K, for $t_\nu \in [0,T]$.

PROOF. The estimate for the stiff error component has $d(\eta_\nu)$ has already been proved in Lemma 4.5. The other estimates follow from Lemma 4.6 by induction over ν in the usual way. □

REMARK 4.1. *Concerning the choice of the constant K, the same remarks apply as for the implicit Euler scheme in Part 1 ([6]). The strong stability of the Radau IIa schemes guarantees that the size of the stiff error component does not grow. In particular, if the integration initially starts with the exact initial value, $\eta_0 = y(0)$, the stiff error component d_ν remains at the $O(\varepsilon h^s)$-level.*

Moreover, the assertion of Theorem 4.7 can be extended in the following way: If $|d(\eta_{\nu-1})| \le K_{\nu-1}\varepsilon h^s$ holds with $K_{min} < K_{\nu-1} \le \frac{1}{\varepsilon}$, then it follows that $|d(\eta_\nu)| \le K_\nu \varepsilon h^s$ holds with

$$K_\nu = \frac{1 + \phi\frac{h}{\varepsilon}\frac{K_{min}}{K_{\nu-1}}}{1+\phi\frac{h}{\varepsilon}} K_{\nu-1}, \tag{4.27}$$

[7] Cf. Lemma 4.5.

which implies $\lim_{\nu\to\infty} K_\nu = K_{min}$*. Thus, the stiff error component is damped down to the* $O(\varepsilon h^s)$*-level with increasing* ν*, even if it is not* $O(\varepsilon h^s)$ *at the beginning. (Details of this argument are given in [3].)*

For the Radau Ia and Gauss schemes, analogous results can be proved. However, due to the lack of strong stability of these methods, there is no 'ε-factor' in the estimate for the stiff error component $d(\eta_\nu)$.

THEOREM 4.8. *Let the assumptions (i)–(v) (cf. section 2) be satisfied and assume that* $\tilde{y}(t) \subset \tilde{\mathcal{M}}$ *is a sufficiently smooth solution of (1.1). Let the initial error satisfy*

$$\begin{aligned} \|\mathrm{pr}(\eta_0) - \tilde{y}_0\| &\le \mathcal{C}h^\sigma, \\ |d(\eta_0)| &\le Kh^\sigma, \end{aligned}$$

Then, for $\sigma \ge 2$ *and under a mild stepsize restriction* $h \le h_{max}$*, the Radau Ia and Gauss methods are convergent and satisfy the global error estimate*

$$\begin{aligned} \|\eta_\nu - \tilde{y}(t_\nu)\| &\le \mathcal{C}h^\sigma, \\ \|\mathrm{pr}(\eta_\nu) - \tilde{y}(t_\nu)\| &\le \mathcal{C}h^\sigma, \\ |d(\eta_\nu)| &\le \mathcal{C}h^\sigma. \end{aligned} \tag{4.28}$$

(Note that σ *denotes the stage order,* $\sigma = s-1$ *for Radau Ia,* $\sigma = s$ *for Gauss, cf. (3.9).)*

PROOF. Given in [3]. □

4.5 The strongly stiff case.

The following sharpened global error estimate for the Radau IIa and Gauss methods has been obtained in [3]. It is valid under the same assumptions as before, plus certain differentiability requirements concerning the data functions $\mathrm{pr}(\eta), \pi(\eta), d(\eta)$ and $\hat{\psi}(x)$ (cf. section 2). Again, $\tilde{y}(t) \subset \tilde{\mathcal{M}}$ denotes a sufficiently smooth solution of (1.1).

THEOREM 4.9. *If, under the assumptions just mentioned, the initial error satisfies*

$$\begin{aligned} \|p(\eta_0) - \tilde{y}_0\| &\le \begin{cases} \mathcal{C}(\varepsilon + h^{s-1})h^s, & \textit{Radau IIa}, \\ \mathcal{C}(\min\{\varepsilon, h\} + h^{s+1})h^{s-1}, & \textit{Gauss}, \end{cases} \\ |d(\eta_0)| &\le \begin{cases} K\varepsilon h^s, & \textit{Radau IIa} \\ \mathcal{C}h^s, & \textit{Gauss}, \end{cases} \end{aligned}$$

then the Radau IIa and Gauss methods satisfy the global error estimates

$$\begin{aligned} \|\eta_\nu - \tilde{y}(t_\nu)\| &\le \begin{cases} \mathcal{C}(\varepsilon + h^{s-1})h^s, & \textit{Radau IIa}, \\ \mathcal{C}h^s, & \textit{Gauss}, \end{cases} \\ \|p(\eta_\nu) - \tilde{y}(t_\nu)\| &\le \begin{cases} \mathcal{C}(\varepsilon + h^{s-1})h^s, & \textit{Radau IIa}, \\ \mathcal{C}(\min\{\varepsilon, h\} + h^{s+1})h^{s-1}, & \textit{Gauss}, \end{cases} \\ |d(\eta_\nu)| &\le \begin{cases} K\varepsilon h^s, & \textit{Radau IIa}, \\ \mathcal{C}h^s, & \textit{Gauss}. \end{cases} \end{aligned} \tag{4.29}$$

PROOF. Given in [3]. The proof is based on the Gröbner-Alekseev Theorem (cf. [13]), on the fact that the methods Radau IIa and Gauss are collocation methods, and on the estimates of Theorem 4.7 . □

The results of Theorem 4.9 are of particular interest for the case of the Radau IIa method: For $\varepsilon \leq \mathcal{C}h^{s-1}$, (4.29) shows that the classical order $p = 2s-1$ is obtained. An analogous result for problems in standard singular perturbation form has been presented in [12] (see also [14], section VI.3).

5 Comments and further results.

In this section we give a brief overview of further results from [2] and [3] which are not discussed in detail here.

- The unique solvability of the Runge-Kutta equations (3.2) is shown under the present assumptions (assumption (i)–(v) from cf. section 2), for $\eta_{\nu-1}$ in an appropriate neighborhood of $\tilde{\mathcal{M}}$. This analysis also provides certain a-priori estimates for $\|(X, D)\|$ and $\|(x, d)\|$ which play a role in the error induction (cf., in particular, the proof of Lemma 4.2) which have not been discussed here in full detail.

- Our assumptions axiomatically characterize a class of stiff problems. The existence of a smooth invariant manifold $\tilde{\mathcal{M}}$ (assumptions (i) and (ii)) and the transversality condition (2.3) (assumption (iii)) are used to describe a 'stiff flow'. However, the transient behavior of nonsmooth solutions of (1.1) starting outside $\tilde{\mathcal{M}}$ does not trivially follow from these assumptions. Such a result has been proven in [3]; in particular, it is shown that $\tilde{\mathcal{M}}$ is strongly exponentially attractive.

- Stiff problems in standard singular perturbation form,

  $$\begin{aligned} x' &= f(x, y), \\ \varepsilon y' &= g(x, y) \end{aligned}$$

  (cf. [12],[17]) have a special geometry. In [3] it is shown that stiff problems of this type form a subclass of the class considered here, with a small, $O(\varepsilon)$-variation of the stiff eigendirection (i.e., the λ-lines are nearly parallel).

- A possible way to generalize the problem class considered here was discussed in [2], assuming that $\tilde{\mathcal{M}}$ is an $(n-k)$-dimensional invariant manifold of (1.1) with $k \geq 1$. The idea is to represent f-differences in the form

  $$f(\eta) - f(p) = J(\eta, p) \cdot (\eta - p) = \int_0^1 Df(p + \xi(\eta - p))d\xi \cdot (\eta - p)$$

  ($\eta \in \mathcal{G}$, $p \in \tilde{\mathcal{M}}$), and the transversality condition is expressed in terms of the eigensystem of the 'generalized Jacobian' $J(\eta, p)$ and an associated elliptic vector norm (see also [1]).

6 Appendix: Underlying stability concepts and results.

In this section we review the stability concepts and results from the 'B-theory' (see [8],[9],[10],[14],[16]) which have been used in the above convergence argument.

In this context, the parameter m denotes a one-sided Lipschitz constant for the function[8] $f(t,y)$, i.e. a parameter $m \in \mathbf{R}$ such that

$$\langle \check{y} - y, f(t,\check{y}) - f(t,y) \rangle \leq m \, \|\check{y} - y\|^2 \tag{6.1}$$

for $(t,y),(t,\check{y})$ in an appropriate domain $[0,T] \times \mathcal{D} \subset \mathbf{R} \times \mathbf{R}^n$.

The notions of BSI- and BS-stability refer to a perturbed Runge-Kutta step

$$\check{Y}_i \;=\; \eta_{\nu-1} + h \sum_{j=1}^{s} a_{ij} f(t_{\nu-1} + c_j h, \check{Y}_j) + \Delta_i, \quad i = 1(1)s, \tag{6.2}$$

$$\check{\eta}_\nu \;:=\; \eta_{\nu-1} + h \sum_{i=1}^{s} b_i f(t_{\nu-1} + c_i h, \check{Y}_i) + \delta. \tag{6.3}$$

6.1 BSI-stability.

A Runge-Kutta scheme is called *BSI-stable* if there exists a continuous monotonously increasing function $\Phi_{BSI}(z)$ such that for a Runge-Kutta scheme (3.2) and a perturbed scheme (6.2) the estimate

$$\|\check{Y} - Y\| \leq \Phi_{BSI}(hm) \|\Delta\| \tag{6.4}$$

is valid for $hm \leq z_{BSI}$.

THEOREM 6.1. *The Gauss and Radau Ia, IIa schemes are BSI-stable with*

$$\Phi_{BSI}(z) = \max_{i,j} \sqrt{\frac{d_i}{d_j}} \, \frac{\|A\|_D}{z_{BSI} - z}.$$

Here, D denotes a diagonal matrix, $D = \operatorname{diag}(d_1, \ldots, d_s)$, and $\|A\|_D$ denotes the matrix norm associated with the vector norm $\|\xi\|_D := \sqrt{\sum_{j=1}^{s} d_j \xi_j^2}$.

PROOF. Given in [8], where also the values D and z_{BSI} specific to the particular schemes are specified. ◻

6.2 BS-stability.

A Runge-Kutta scheme is called *BS-stable* if there exists a continuous monotonously increasing function $\Phi_{BS}(z)$ such that for a Runge-Kutta step (3.2),(3.3) and a perturbed step (6.2),(6.3) the estimate

$$\|\check{\eta}_\nu - \eta_\nu\| \leq \Phi_{BS}(hm) \left(\|\Delta\| + \|\delta\| \right) \tag{6.5}$$

[8] Here we also admit $f = f(t,y)$ (nonautonomous case). The notation f for the underlying ODE $y' = f(t,y)$ is generic; it is not to be directly identified with the function f from (1.1).

is valid for $hm \le z_S$.

THEOREM 6.2. *The Gauss and Radau Ia, IIa schemes are BS-stable.*

PROOF. Given in [10]. □

6.3 B-stability.

A Runge-Kutta scheme (3.2),(3.3) is called *B-stable* if there exists a continuous monotonously increasing function $\Phi_B(z)$ with $\Phi_B(0) = 1$, such that for a pair of 'parallel' Runge-Kutta steps $\eta_{\nu-1} \to \eta_\nu$ and $\check{\eta}_{\nu-1} \to \check{\eta}_\nu$ the estimate

$$\|\check{\eta}_\nu - \eta_\nu\| \le \Phi_B(hm)\|\check{\eta}_{\nu-1} - \eta_{\nu-1}\| \tag{6.6}$$

is valid for $hm \le z_B$.

THEOREM 6.3. *The Gauss and Radau Ia, IIa schemes are B-stable with* $z_B = 0$*, i.e. for* $hm \le 0$*, with*

$$\Phi_B(z) = \begin{cases} 1, & \textit{Gauss}, \\ \frac{1}{\sqrt{1-2z}}, & \textit{Radau Ia}, \\ \frac{1}{\sqrt{1-2b_s z}}, & \textit{Radau IIa}. \end{cases}$$

PROOF. Given in [10]. □

REMARK 6.1. *A computable bound for the B-stability function of arbitrary irreducible, algebraically stable Runge-Kutta methods has been obtained in [16] (Theorem 3.1); cf. also [14], section IV.12.*

6.4 The case $m = -1/\varepsilon$, $0 < \varepsilon \ll 1$.

In section 4, B- and BS-stability estimates are required for the special case of scalar ODEs of the form

$$y' = \lambda(\cdot)y \tag{6.7}$$

with $\lambda(\cdot) \le -\frac{1}{\varepsilon}$ ($\Rightarrow m = -\frac{1}{\varepsilon}$). Due to the representation for $\Phi_{BSI}(z)$ given in Theorem 6.1, the BSI-stability function for $m = -\frac{1}{\varepsilon}$ is of the form

$$\Phi_{BSI}(-\tfrac{h}{\varepsilon}) = \frac{\mathcal{C}}{1 + \phi_{BSI}\frac{h}{\varepsilon}} \le \mathcal{C} \tag{6.8}$$

with a certain moderate constant $\mathcal{C}$ and $\phi_{BSI} := 1/z_{BSI}$.

6.5 The Radau IIa schemes.

For the Radau IIa schemes we have $b_i \equiv a_{si}$, and therefore the BSI-stability function yields also a BS-stability function via $\Phi_{BS}(z) = \Phi_{BSI}(z)$ for $\delta := \Delta_s$, and with (6.8) we also have

$$\Phi_{BS}(-\tfrac{h}{\varepsilon}) = \frac{\mathcal{C}}{1 + \phi_{BS}\frac{h}{\varepsilon}} \tag{6.9}$$

with $\phi_{BS} = \phi_{BSI}$.

Concerning B-stability for Radau IIa, there is a subtle point. The estimate for the stiff error component in section 4 (Lemma 4.5) relies on a sharp stability estimate for the special case [9] (6.7), namely

$$\Phi_B(-\tfrac{h}{\varepsilon}) \le \frac{1}{1+\phi_B \frac{h}{\varepsilon}} \tag{6.10}$$

with a constant $\phi_B > 0$. For $s = 1, 2$ we have

$$s=1: \quad \Phi_B(-\tfrac{h}{\varepsilon}) = \frac{1}{1+\frac{h}{\varepsilon}}, \qquad \text{(implicit Euler)},$$

$$s=2: \quad \Phi_B(-\tfrac{h}{\varepsilon}) = \frac{4}{5+2\frac{h}{\varepsilon}}, \qquad \text{(cf. [16])},$$

leading to the conjecture that the desired form (6.10) also holds for general $s > 2$. However, such a result appears not to be available, and the estimates from [16] are too weak for this purpose; they only yield the weaker bound

$$\Phi_B(-\tfrac{h}{\varepsilon}) \le \frac{1}{\sqrt{1+\phi_B \frac{h}{\varepsilon}}} \tag{6.11}$$

in accordance with Theorem 6.3. It would be possible to prove Lemma 4.5 using this weaker estimate, but only under the extreme stepsize restriction $h \le \mathcal{C}\varepsilon$. Motivated by this observation, the stability of Radau IIa methods applied to (6.7) was studied in [3] in more detail.

LEMMA 6.4. *For scalar ODEs of the type (6.7) with* $\lambda(\cdot) = \lambda(t) = -\gamma(t)/\varepsilon$, *with* $\gamma(t) \ge 1$ *at* $t = t_{\nu-1}$ *and* $t = t_{\nu-1} + c_i h$, *and under the weak smoothness requirement* $(\gamma(t_{\nu-1}+c_i h) - \gamma(t_{\nu-1})) \le \mathcal{C}\gamma(t_{\nu-1})$ *and a mild stepsize restriction* $0 < h \le h_{max}$, *the Radau IIa schemes are B-stable with a stability function of the form (6.10).*

PROOF. Given [3]. This result is considered for separate publication; it is an interesting contribution to the question to what extent the functions describing B-stability and 'AN-stability', i.e., B-stability for the special case of a scalar nonautonomous ODE, are related (cf. [15] and the discussion in [16]). □

REFERENCES

1. W. Auzinger, A. Eder, *A note on Lyapunov transformation and exponential decay in linear ODE systems*, Mathematical Models and Methods in Applied Sciences, 11:1 (2001), pp. 23–31.
2. W. Auzinger, A. Eder, and R. Frank, *Convergence Theory for Implicit Runge-Kutta Methods Applied to a One-Parameter Family of Stiff Autonomous Differential Equations*, Tech. Report 123/98, Institut für Angewandte und Numerische Mathematik, Technische Universität Wien, 1998.

[9] In [14],[16], such a special stability function is denoted by $\varphi_K(\cdot)$.

3. W. Auzinger, A. Eder, and R. Frank, *Eine erweiterte Konvergenztheorie impliziter Runge-Kutta Verfahren für steife Anfangswertprobleme*, Tech. Report 130/2000, Institut für Angewandte und Numerische Mathematik, Technische Universität Wien, 2000. Available under `http://www.math.tuwien.ac.at/~winfried/ber130.ps.gz`.
4. W. Auzinger, R. Frank, and G. Kirlinger, *A note on convergence concepts for stiff problems*, Computing, 44 (1990), pp. 197–208.
5. W. Auzinger, R. Frank, and G. Kirlinger, *An extension of B-convergence for Runge-Kutta methods*, Appl. Numer. Math., 9 (1992), pp. 91–109.
6. W. Auzinger, R. Frank, and G. Kirlinger, *Extending convergence theory for nonlinear stiff problems, Part I*, BIT, 36:4 (1996), pp. 635–652.
7. W. Auzinger, R. Frank, and H. J. Stetter, *Vienna contributions to the development of RK methods*, Appl. Numer. Math., 22 (1996), pp. 35–49.
8. K. Dekker and J. G. Verwer, *Stability of Runge-Kutta Methods for Stiff Nonlinear Differential Equations*, North-Holland, 1984.
9. R. Frank, J. Schneid, and C. W. Ueberhuber, *The concept of B-convergence*, SIAM J. Numer. Anal., 18 (1981), pp. 753–780.
10. R. Frank, J. Schneid, and C. W. Ueberhuber, *Stability properties of implicit Runge-Kutta methods*, SIAM J. Numer. Anal., 22 (1985), pp. 497–515.
11. R. Frank, J. Schneid, and C. W. Ueberhuber, *Order results for implicit Runge-Kutta methods applied to stiff systems*, SIAM J. Numer. Anal., 22 (1985), pp. 515–534.
12. E. Hairer, E. Lubich, and M. Roche, *Error of Runge-Kutta methods for stiff problems s tudied via differential-algebraic equations*, BIT, 28 (1988), pp. 678–700.
13. E. Hairer and G. Wanner, *Solving Ordinary Differential Equations I. Nonstiff Problems*, Springer-Verlag, Berlin, Heidelberg, 1991.
14. E. Hairer and G. Wanner, *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems*, 2nd ed., Springer-Verlag, Berlin, Heidelberg, 1996.
15. E. Hairer and G. Wanner, *On a generalization of a theorem of von Neumann*, Z. Angew. Math. Mech. 76 (1996), suppl. 1, pp. 95–98.
16. E. Hairer and M. Zennaro, *On error growth functions of Runge-Kutta methods*, Appl. Numer. Math., 22 (1996), pp. 205–216.
17. K. Nipp and D. Stoffer, *Invariant manifolds and global error estimates of numerical integration schemes applied to stiff systems of singular perturbation type*, Numer. Math., 70 (1995), pp. 245–257.