Institut für Angewandte und
Numerische Mathematik

Technische Universität Wien

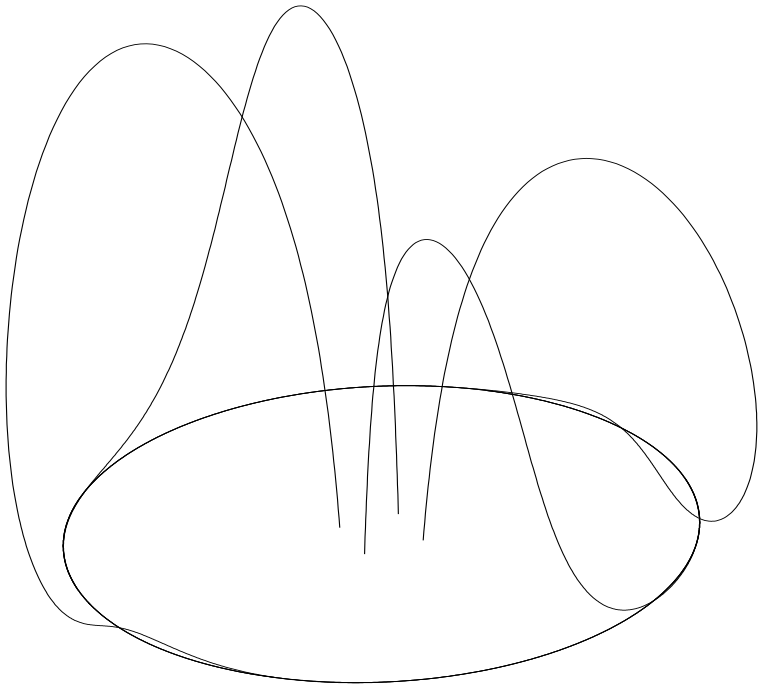
Report Nr. 123/98

Convergence Theory for Implicit Runge-Kutta Methods Applied to a One-Parameter Family of Stiff Autonomous Differential Equations

W. Auzinger
A. Eder
R. Frank

Convergence Theory for Implicit Runge-Kutta Methods Applied to a One-Parameter Family of Stiff Autonomous Differential Equations

W. Auzinger, R. Frank, A. Eder



The intention of this paper is to extend the convergence concepts for discretization methods applied to nonlinear stiff problems. First a one-parameter family of stiff autonomous differential equations is introduced, where stiffness is axiomatically characterized in geometric terms. Then the discretization methods are analyzed, where we restrict our considerations to the implicit Runge-Kutta methods of the type Radau Ia, IIa and Gauss. For these methods we prove the solvability of the algebraic equations and derive global error bounds.

Contents

Notation:

ODE	ordinary differential equation
$A\|_{\mathcal{M}}, f\|_{\mathcal{M}}, \ldots$	$A, f, \ldots$ restricted to $\mathcal{M}$
$\|\cdot\|_2$	Euclidean norm, also used as an operator norm
$\oplus$	direct sum
Re, Im	real resp. imaginary part of a complex number
ONS	orthonormal system
$\cdot^\top$	transposition of the matrix $\cdot$
$\langle\cdot,\cdot\rangle$	standard scalarproduct
$\|\cdot\|_V$	Lyapunov norm defined as $\langle V\cdot,\cdot\rangle^{\frac{1}{2}}$
$O(\cdot)$	Landau's order symbol
$M_{f\|_{\mathcal{M}}}, L_{f\|_{\mathcal{M}}}, \ldots$	constants (bounds, Lipschitz constants, ...) corresponding to $f\|_{\mathcal{M}}$
$Jf(y), J\varphi(x), \ldots$	Jacobian of $f(y), \varphi(x), \ldots$
$\mathcal{G} \setminus \mathcal{M}$	set $\mathcal{G}$ minus the set $\mathcal{M}$
$\overline{\mathcal{B}}_r$	closed ball with radius r and center 0
$\phi(\mathcal{U}), (x,d)(\mathcal{V}), \ldots$	image of a set under a mapping
$Hd_i(y)$	Hesse matrix of $d_i(y)$
$\nabla d_i(y), \nabla v_{ij}(y), \ldots$	gradient vector of $d_i(y), v_{ij}(y), \ldots$ (column vector)
$[\cdot]_j$	j-th component of a vector
$(\cdot)_{ij}$	matrix with the entries $\cdot$ in the i-th row and j-th column
$\|(\cdot,\cdot)\|_*$	norm defined as $\max\{\|\cdot\|_2, \|\cdot\|_2\}$
$I, I_k, \ldots$	identity matrix (k indicates the dimension)
$\mathcal{B}_r$	open ball with radius r and center 0
$\circ$	composition of mappings
$\mathcal{C}$	generic constant
$(\cdot)_{i=1}^k$	vector with the entries $\cdot$ in the i-th component
$\cdot^s$	applied to the set $\cdot$ is the set $\underbrace{\cdot \times \cdot \times \cdots \times \cdot}_{s \text{ times}}$
diag(...)	diagonal or block diagonal matrix with the diagonal resp. block diagonal entries ...

$\otimes$	direct product (Kronecker product)
e	vector defined as $e := (1, \dots, 1)^\top$
$\|\cdot\|_\infty$	max-norm
$\|(\cdot, \dots, \cdot)^\top\|_\circ$	norm defined as $\|(\|\cdot\|_2, \dots, \|\cdot\|_2)^\top\|_\infty$
$\|(\cdot, \cdot)^\top\|_\diamond$	norm defined as $\|(\|\cdot\|_\circ, \|\cdot\|_\circ)^\top\|_\infty$
$\langle \cdot, \cdot \rangle_V$	inner product defined as $\langle V\cdot, \cdot\rangle$
$\|\|\|(\cdot, \dots, \cdot)^\top\|\|\|_V$	norm defined as $\|(\|\cdot\|_V, \dots, \|\cdot\|_V)^\top\|_2$
$\Phi_I(\cdot)$	BSI-stability function
$\Phi_B(\cdot)$	B-stability function
$\Phi_{BS}(\cdot)$	BS-stability function
$B(\cdot), C(\cdot), D(\cdot)$	simplifying conditions of Butcher
$\dfrac{\partial^{\|l\|}}{\partial y_1^{l_1} \cdots \partial y_n^{l_n}}$	partial derivative of order $\|l\| = l_1 + \cdots + l_n$ (w.r.t. the variables $y_1, \dots, y_n$)
$\mathcal{L}\{\dots\}$	linear hull of $\dots$
$\mathcal{T}(\cdot)$	tangential space to the manifold in the point $\cdot$
$\ell_j(\cdot)$	j-th Lagrange basis function
$\dfrac{\partial \tilde{x}}{\partial \check{x}}(t, \sigma, \check{x})$	is the matrix $\Big(\dfrac{\partial \tilde{x}_i}{\partial \check{x}_j}(t, \sigma, \check{x}_1, \dots, \check{x}_{n-k})\Big)_{ij}$
$\mathcal{B}_r(\xi_0)$	open ball with radius r and center ξ_0
$\overline{\mathcal{B}_r(\xi_0)}$	closed ball with radius r and center ξ_0
$\det(\cdot)$	determinant of a matrix
$\operatorname{sgn}(\cdot)$	is defined as -1 for $\cdot < 0$ and 1 else
IVP	initial value problem
$\langle \cdot, \cdot \rangle_\bullet$	arbitrary inner product
$\|\cdot\|_\bullet$	norm defined as $\langle \cdot, \cdot \rangle_\bullet^{\frac{1}{2}}$
$\|\|\|(\cdot, \dots, \cdot)^\top\|\|\|_\bullet$	norm defined as $\|(\|\cdot\|_\bullet, \dots, \|\cdot\|_\bullet)^\top\|_2$

Chapter 1

Introduction

As pointed out in recent papers by Auzinger et al. (cf. [1, 2, 3, 4, 5]) the existing convergence concepts for the analysis of discretizations of nonlinear stiff problems suffer from some limitations:

- The theory of B-convergence relies on the one-sided Lipschitz condition (cf. e.g., [7, 8, 9, 10, 11, 13, 21]). In this theory the one-sided Lipschitz constant m is the essential problem-characterizing parameter. If m is not strongly positive then satisfactory error bounds can be derived. Since for stiff problems the generic case is $m \gg 0$ there is a need for an extension of these concepts. For a detailed discussion cf. [1].

- Differential equations in standard singular perturbation form have also been analyzed (cf. [13, 14, 17, 18]). They have the special structure that the variation of the stiff eigendirections is only $O(\varepsilon), 0 < \varepsilon \ll 1$. This means that the phase portrait resembles that of a constant coefficient problem. For a more detailed discussion cf. [1].

These limitations motivate to consider a class of problems where stiffness is axiomatically described and which reflects the essence of stiffness:
Stiff differential equations admit smooth solutions with moderate derivatives, together with non-smooth solutions rapidly converging towards smooth solutions. Stiff problems are assumed to be well-conditioned in a global sense.[1)]
The problem class described in this paper contains nonlinear autonomous differential equations with one stiff parameter $0 < \varepsilon \ll 1$. In this class the smooth solutions of computational interest are assumed to lie in a smooth invariant manifold $\mathcal{M}$ of dimension $n-k$, where n is the dimension of the state space. The parameter ε describes how fast the nonsmooth solutions converge to the smooth manifold $\mathcal{M}$.
Based on a semi-global linearization a convergence theory for the implicit Runge-Kutta methods Radau Ia, IIa and Gauss is derived (including a proof for the solvability of the algebraic equations). Instead of considering the Jacobians in single points y, each point y is associated with another point $u \in \mathcal{M}$ such that $y - u$ is contained in a stiff eigenspace in a generalized sense. The details are specified in section 2.2. A natural splitting into stiff and smooth components arises, which leads to the following error bounds

[1)] Note that the bad local condition, which often occurs as a consequence of the transient behavior of nonsmooth solutions, does not contradict the fact that a stiff problem is typically well-conditioned in a global sense.

method	smooth error component	stiff error component
Radau Ia	$O(h^{s-1})$	$O(h^{s-1})$
Radau IIa	$O(\varepsilon h^s + h^{2s-1})$	$O(\varepsilon h^s)$
Gauss	$O(\min\{\varepsilon, h\}h^{s-1} + h^{2s})$	$O(h^s)$

where s is the number of the stages of the method. The constants in the $O(\cdot)$-terms are ε-independent.

Chapter 2

A One-Parameter Family of Stiff Autonomous Differential Equations

We consider initial value problems for autonomous stiff ODEs

$$y'(t) = f(y(t)), \quad t \in [0, T], \tag{2.1}$$

where $f : \mathcal{G} \to \mathbb{R}^n$, $\mathcal{G} \subseteq \mathbb{R}^n$.
In the following sections the notion of stiffness is formalized in geometric terms and via a real stiff parameter $0 < \varepsilon \ll 1$.[1)]
As an introduction to our notion of stiffness we present a linear homogeneous autonomous system of stiff differential equations with one stiff parameter (cf. also [6]).

2.1 A Linear Problem Class with One Stiff Parameter

Given a linear homogeneous autonomous system of ODEs

$$y'(t) = Ay(t), \quad t \geq 0, \tag{2.2}$$

where A is a real $n \times n$-matrix, we define stiffness in the following way:

Existence of an invariant manifold $\mathcal{M}$ containing smooth solutions

- Let $\mathcal{M} \subseteq \mathbb{R}^n$ be an invariant linear space of dimension $n–k$ of the linear mapping $A : \mathbb{R}^n \to \mathbb{R}^n$. Then $\mathcal{M}$ is an invariant manifold of the differential equation (2.2). Solutions $u(t)$ of the differential equation (2.2) with initial values $u(0) = u_0 \in \mathcal{M}$ are of the form

$$u(t) = \exp(A|_{\mathcal{M}} t)u_0, \quad t \geq 0. \tag{2.3}$$

- We assume that $\|A|_{\mathcal{M}}\|_2$ is moderate, which means that the manifold $\mathcal{M}$ contains smooth solutions. A moderate growth of the solutions in $\mathcal{M}$ is allowed.

Remark: The moderateness of $\|A|_{\mathcal{M}}\|_2$ implies the moderateness of $n-k$ eigenvalues of A, the so-called *smooth eigenvalues*. The remaining k eigenvalues $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$ of A are called the *stiff*

[1)] Since our proceeding is first to consider a function $f(y)$ and then to identify the parameter ε we prefer the notation $f(y(t))$ instead of $f(y(t), \varepsilon)$.

eigenvalues of A. For these eigenvalues and the corresponding real invariant linear space we make the following assumptions:

Transversality condition and a natural splitting into smooth and stiff components

- Let $\mathcal{E} \subseteq \mathbb{R}^n$ be the k-dimensional invariant linear space of the mapping A such that the state space is the direct sum $\mathbb{R}^n = \mathcal{M} \oplus \mathcal{E}$.
- It is due to the splitting $\mathbb{R}^n = \mathcal{M} \oplus \mathcal{E}$ that to each point $y \in \mathbb{R}^n$ there exists a unique point $u \in \mathcal{M}$ such that $y - u \in \mathcal{E}$. This defines a linear projection $p\colon \mathbb{R}^n \to \mathcal{M}$, $p(y) := u$. The projection $p(y)$ is called the *smooth component* of y and $y - p(y)$ is called the *stiff component* of y. The affine spaces $u + \mathcal{E}$, $u \in \mathcal{M}$ are called λ-*planes*, i.e. a λ-plane through $u \in \mathcal{M}$ consists of all points $y \in \mathbb{R}^n$ for which $p(y) = u$. Compare figure 2.1.

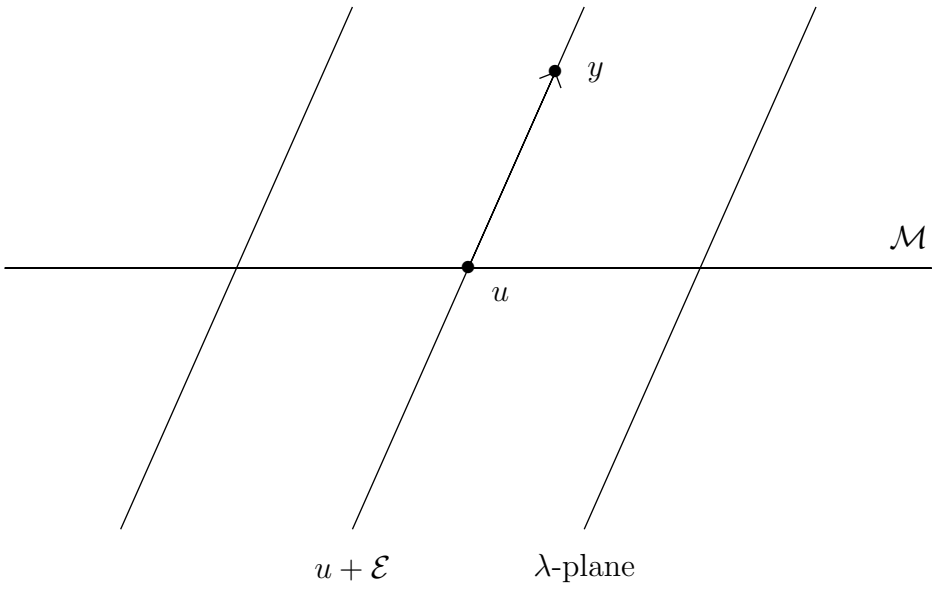


Figure 2.1: Transversality condition and λ-planes

- The stiff eigenvalues $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$ of $A|_{\mathcal{E}}$ are assumed to be of the form

$$\lambda_i = -\frac{c_i}{\varepsilon}, \quad i = 1, \ldots, k, \tag{2.4}$$

  where $0 < \varepsilon \ll 1$ is the so-called *stiff parameter*. The constants c_i are assumed to satisfy

$$\operatorname{Re} c_i \geq 1, \tag{2.5}$$

  with c_i moderate.

Remark: We restrict our considerations to problems which are globally well-conditioned concerning the stiff component,[2)] i.e. problems for which there exist moderate real constants $\mathcal{K}_0, \mathcal{K}_1 \geq 1$ such that there holds

$$\|y(t) - p(y(t))\|_2 \leq \mathcal{K}_0 \exp(-\frac{1}{\varepsilon \mathcal{K}_1} t) \|y_0 - p(y_0)\|_2, \quad t \geq 0 \tag{2.6}$$

[2)] Stiff problems are typically ill conditioned in a local sense, compare [3] for more details.

for all solutions $y(t)$ of the differential equation (2.2) with initial values y_0. Since stiff eigenvalues alone do not in all cases imply globally well-conditionedness concerning the stiff component we need further assumptions for the geometry of the ODE:[3)]

The stiff coordinates and a corresponding Lyapunov norm

- Let $\pi_1, \ldots, \pi_k$ be an ONS of $\mathcal{E}$. To each point $y \in \mathbb{R}^n$ the coordinate vector of $y - p(y)$ in the ONS of $\mathcal{E}$ is denoted by $d(y) = (d_1(y), \ldots, d_k(y))^\top$. The components $d_i(y)$ are called the *stiff coordinates* of y.

- The matrix representation of $A|_{\mathcal{E}}$ corresponding to the ONS $\pi_1, \ldots, \pi_k$ of $\mathcal{E}$ is denoted by Λ. We call Λ the *stiff eigenmatrix* of A.

- The unique positive definite solution $V \in \mathbb{R}^{k\times k}$ of the Lyapunov equation[4)]

$$\Lambda^\top V + V\Lambda \;=\; -\frac{1}{\varepsilon} I. \tag{2.7}$$

  is assumed to have moderate $\|V\|_2$ and $\|V^{-1}\|_2$. The solution V of (2.7) induces an elliptic vector norm

$$\|d\|_V \;:=\; \langle Vd, d\rangle^{\frac{1}{2}} \tag{2.8}$$

  for $d \in \mathbb{R}^k$, which we call the *Lyapunov norm*.

Remarks:

- The choice of an ONS implies norm-invariance, i.e. $\|y - p(y)\|_2 = \|d(y)\|_2$ as well as $\|\exp(A|_{\mathcal{E}} t)\|_2 = \|\exp(\Lambda t)\|_2$.

- The moderateness of $\|V\|_2, \|V^{-1}\|_2$ is indepedent of which ONS in $\mathcal{E}$ is chosen.

- The condition that $\|V\|_2$ and $\|V^{-1}\|_2$ are moderate means that the distortion of the corresponding elliptic vector norm $\|\cdot\|_V$ is moderate.

Globally well-conditionedness concerning the stiff component

The Lyapunov norm $\|\cdot\|_V$ can be used to derive[5)]

$$\|d(y(t))\|_V \;\le\; \exp(-\frac{1}{\varepsilon}\frac{1}{2\|V\|_2}\, t)\|d(y_0)\|_V, \quad t \ge 0 \tag{2.9}$$

for all solutions $y(t)$ of (2.2) with initial values y_0. The estimate (2.9) yields

$$\begin{aligned} \|y(t) - p(y(t))\|_2 \;&\le \\ &\le\; \left(\|V\|_2\|V^{-1}\|_2\right)^{\frac{1}{2}} \exp(-\frac{1}{\varepsilon}\frac{1}{2\|V\|_2}\, t)\|y_0 - p(y_0)\|_2 \end{aligned} \tag{2.10}$$

for $t \ge 0$, which together with the moderateness assumption for $\|V\|_2$ and $\|V^{-1}\|_2$ implies that (2.6) is fulfilled. This means that for the stiff components of a solution $y(t)$ of the ODE (2.2) locally a

[3)] Compare [6] for more details concerning globally well-conditioned stiff linear problems.
[4)] For more information concerning the Lyapunov equation see [15].
[5)] Compare [6], as well as for the estimate (2.10).

small growth on a short intervall is allowed, but globally there is an exponential decrease towards the smooth manifold. Compare [6] for more details concerning linear problems.

In the next section these concepts are extended to the nonlinear autonomous case. A semi-global linearization with the so-called generalized Jacobian and additional smoothness assumptions again lead to well-conditionedness concerning the stiff components in a global sense.

2.2 A Nonlinear Problem Class with One Stiff Parameter

Now we consider nonlinear homogeneous autonomous ODEs

$$y'(t) \;=\; f(y(t)), \quad t \geq 0, \tag{2.11}$$

where $f : \mathcal{G} \to \mathbb{R}^n$, $\mathcal{G} \subseteq \mathbb{R}^n$. The right hand side f is assumed to be as often differentiable as the analysis requires.
The notion of stiffness is formalized in the following way:

Existence of an invariant manifold $\mathcal{M}$ containing smooth solutions

- Let $\mathcal{M} \subseteq \mathcal{G}$ be a $(n-k)$-dimensional invariant manifold of the differential equation (2.11) on which the smooth solutions persist. It is assumed that $\mathcal{G}$ forms at least a $O(1)$-neighbourhood of $\mathcal{M}$. Compare figure 2.2.

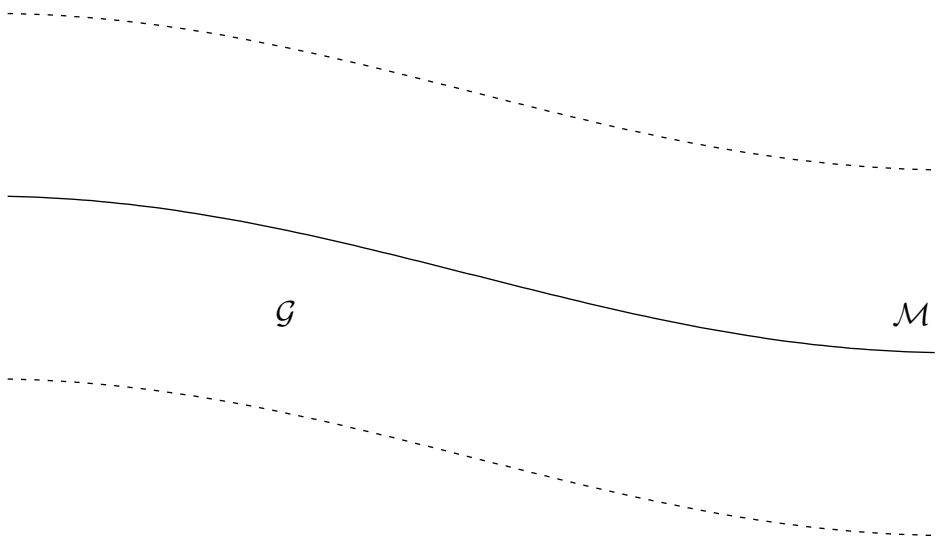


Figure 2.2: The invariant manifold $\mathcal{M}$

- We assume that there exist moderate real constants $M_{f|\mathcal{M}}, L_{f|\mathcal{M}} \geq 0$ such that

$$\|f(u)\|_2 \;\leq\; M_{f|\mathcal{M}}, \tag{2.12}$$

$$\|f(u) - f(\tilde{u})\|_2 \;\leq\; L_{f|\mathcal{M}}\|u - \tilde{u}\|_2, \tag{2.13}$$

for all $u, \tilde{u} \in \mathcal{M}$.[6] The constants $M_{f|\mathcal{M}}$, $L_{f|\mathcal{M}}$ depend only on the behavior of the differential equation on $\mathcal{M}$.

Remark: The smoothness assumptions (2.12), (2.13) imply that solutions $u(t)$ of the differential equation (2.11) in $\mathcal{M}$ have moderate bounds for the first and second derivatives, i.e. there holds

$$\|u'(t)\|_2 \leq M_{f|\mathcal{M}}, \quad \|u''(t)\|_2 \leq L_{f|\mathcal{M}} M_{f|\mathcal{M}}. \tag{2.14}$$

Transversality condition and a natural splitting into smooth and stiff components

- We consider the identity
$$f(y) - f(u) = J(y,u)(y-u), \tag{2.15}$$
where
$$J(y,u) := \int_0^1 Jf(u + \sigma(y-u))\, d\sigma \tag{2.16}$$
is called the *generalized Jacobian* at the points $y, u \in \mathcal{G}$.

- We assume that to each point $y \in \mathcal{G} \setminus \mathcal{M}$ there exists a locally unique point $u \in \mathcal{M}$, such that $y - u \in \mathcal{E}(y,u)$, where $\mathcal{E}(y,u)$ is assumed to be a k-dimensional real stiff invariant linear space of $J(y,u)$. The terminology *stiff* concerning the invariant space $\mathcal{E}(y,u)$ means that the eigenvalues $\lambda_1(y,u), \ldots, \lambda_k(y,u) \in \mathbb{C}$ of $J(y,u)|_{\mathcal{E}(y,u)}$ are of the form
$$\lambda_i(y,u) = -\frac{1}{\varepsilon} c_i(y,u), \quad i = 1, \ldots, k, \tag{2.17}$$
where $0 < \varepsilon \ll 1$ is the so-called *stiff parameter* and the functions $c_i(y,u)$ are assumed to satisfy
$$\operatorname{Re} c_i(y,u) \geq 1, \quad i = 1, \ldots, k, \tag{2.18}$$
with $c_i(y,u)$ moderate.

  This defines a nonlinear projection $p\colon \mathcal{G} \setminus \mathcal{M} \to \mathcal{M}$, $p(y) := u$ which is assumed to be extendable to $\mathcal{G}$. The extension of p to $\mathcal{G}$ is also denoted by p. The projection $p(y)$ is called the *smooth component* of y and $y - p(y)$ is called the *stiff component* of y. The manifolds $\{y\colon p(y) = u\}$, $u \in \mathcal{M}$ are called λ-*planes*. Compare figure 2.3.

Remark: It seems to be natural that the λ-planes are k-dimensional smooth manifolds which together with the $(n-k)$-dimensional invariant manifold $\mathcal{M}$ form a nonlinear coordinate system. The details are specified in the following assumptions.

The stiff coordinates and a corresponding Lyapunov norm

- Let $\pi_1(y), \ldots, \pi_k(y)$ be an ONS of $\mathcal{E}(y) := \mathcal{E}(y, p(y))$ for $y \in \mathcal{G}$. Concerning the smoothness of the ONS we assume that there exists a moderate real constant $L_\pi \geq 0$ such that
$$\|\pi_i(y) - \pi_i(\tilde{y})\|_2 \leq L_\pi \|y - \tilde{y}\|_2, \quad i = 1, \ldots, k \tag{2.19}$$

[6] We restrict our considerations to a Lipschitz condition for f on $\mathcal{M}$ instead of a one sided Lipschitz condition
$$\langle f(u) - f(\tilde{u}), u - \tilde{u} \rangle \leq m_{f|\mathcal{M}} \|u - \tilde{u}\|_2^2, \quad \forall u, \tilde{u} \in \mathcal{M}$$
with a moderate one sided Lipschitz constant $m_{f|\mathcal{M}} \in \mathbb{R}$.

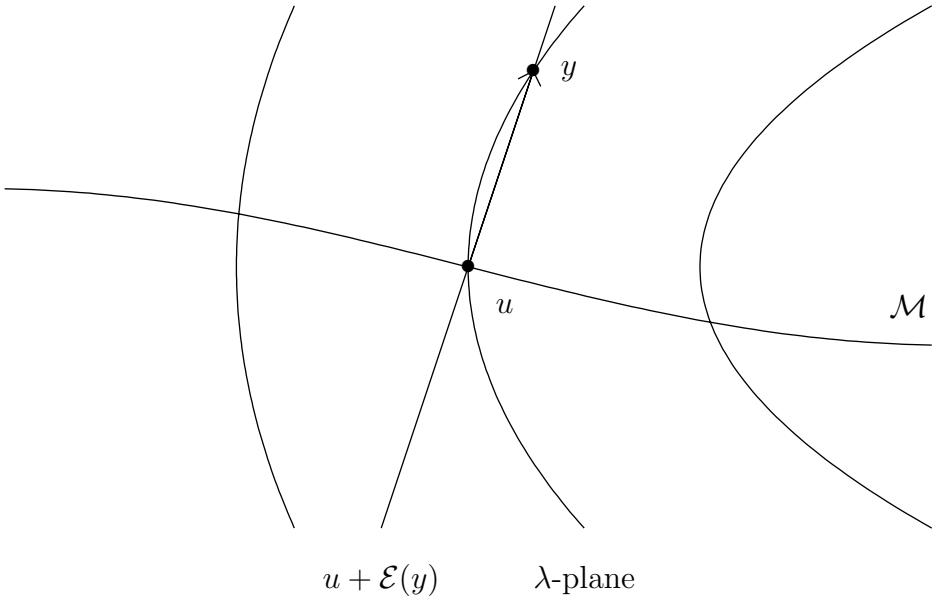


Figure 2.3: Transversality condition and λ-planes

holds for all $y, \tilde{y} \in \mathcal{G}$. To each point $y \in \mathcal{G}$ the coordinate vector of $y - p(y)$ in the ONS of $\mathcal{E}(y)$ is denoted by $d(y) = (d_1(y), \ldots, d_k(y))^\top$. The components $d_j(y)$ are called the *stiff coordinates* of y. We assume that the stiff coordinate $d(\cdot)$ is injective on each λ-plane and that there exists $r > 0$ such that the image $\mathcal{D} := \{d(y) : y \in \lambda\text{-line}\}$ of each λ-line contains the closed ball $\overline{\mathcal{B}}_r := \{d \in \mathbb{R}^k : \|d\|_2 \leq r\}$. Compare figure 2.4.

- Let $J(y) := J(y, p(y))$. The matrix representation of $J(y)|_{\mathcal{E}(y)}$ corresponding to the ONS $\pi_1(y), \ldots, \pi_k(y)$ is denoted by $\Lambda(y)$. The matrix $\Lambda(y)$ is called the *stiff eigenmatrix* of $J(y)$. We assume that the matrix $G(y)$ defined via the identity

$$\Lambda(y) \;=\; \frac{1}{\varepsilon} G(y) \tag{2.20}$$

is moderate, i.e. there exists a moderate real constant $M_G \geq 0$ such that there holds

$$\|G(y)\|_2 \;\leq\; M_G \tag{2.21}$$

for all $y \in \mathcal{G}$. Concerning the smoothness of $G(y)$ we assume that there exists a moderate real constant $L_G \geq 0$ such that there holds

$$|g_{ij}(y) - g_{ij}(\tilde{y})| \;\leq\; L_G \|y - \tilde{y}\|_2, \quad i, j = 1, \ldots, k \tag{2.22}$$

for all $y, \tilde{y} \in \mathcal{G}$, where $g_{ij}(y)$ denote the entries of the matrix $G(y)$.

- Let $V(y) \in \mathbb{R}^{k \times k}$ be the unique positive definite solution of the Lyapunov equation[7)]

$$\Lambda(y)^\top V(y) + V(y)\Lambda(y) \;=\; -\frac{1}{\varepsilon} I. \tag{2.23}$$

[7)] See [15] concerning the Lyapunov equation.

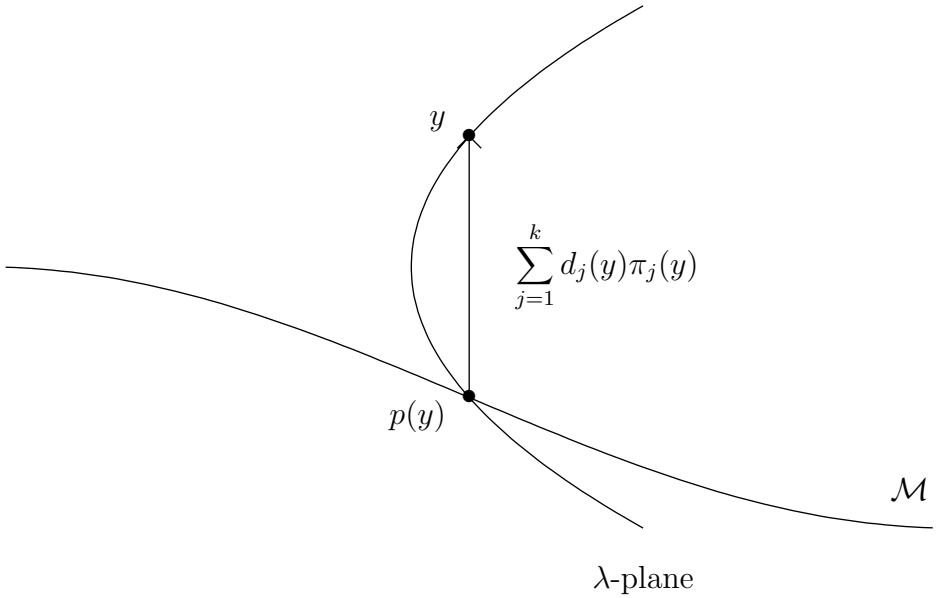


Figure 2.4: The stiff coordinate $d(\cdot)$

The solution $V(y)$ of the Lyapunov equation (2.23) induces a state dependent elliptic vector norm

$$\|d\|_{V(y)} := \langle V(y)d, d\rangle^{\frac{1}{2}} \tag{2.24}$$

for $d \in \mathbb{R}^k$, which we call the *Lyapunov norm*. We assume that there exist moderate real constants $M_V, M_{V^{-1}} \geq 0$ such that there holds

$$\|V(y)\|_2 \leq M_V, \quad \|V(y)^{-1}\|_2 \leq M_{V^{-1}} \tag{2.25}$$

for all $y \in \mathcal{G}$. Furthermore we assume that there exists a moderate real constant $L_V \geq 0$ such that there holds

$$|v_{ij}(y) - v_{ij}(\tilde{y})| \leq L_V \|y - \tilde{y}\|_2, \quad i,j = 1,\ldots,k \tag{2.26}$$

for all $y, \tilde{y} \in \mathcal{G}$, where $v_{ij}(y)$ denote the entries of $V(y)$.

Now we introduce a special family of local parametrizations of the smooth manifold $\mathcal{M}$ which leads to a family of local coordinate transformations in $\mathcal{G}$. These local parametrizations and coordinate transformations are suiteable for further considerations.

A family of parametrizations and the corresponding smooth coordinates

- Let $\bar{\eta} \in \mathcal{G}$ be arbitrary. Then the corresponding smooth component is denoted by $\bar{p} := p(\bar{\eta})$ and the corresponding stiff coordinates are denoted by $\bar{d} := (\bar{d}_1,\ldots,\bar{d}_k)^\top := d(\bar{\eta})$. We extend the ONS $\bar{\pi}_1 := \pi_1(\bar{\eta}),\ldots,\bar{\pi}_k := \pi_k(\bar{\eta})$ of $\mathcal{E}(\bar{\eta})$ by $n-k$ vectors $\hat{\pi}_1,\ldots,\hat{\pi}_{n-k} \in \mathbb{R}^n$ to an ONS

$$B := \{\hat{\pi}_1,\ldots,\hat{\pi}_{n-k},\bar{\pi}_1,\ldots,\bar{\pi}_k\} \tag{2.27}$$

of $\mathbb{R}^n$.

- Now we assume that there exists a neighbourhood $\mathcal{U} \subseteq \mathbb{R}^{n-k}$ of $0 \in \mathbb{R}^{n-k}$ and a function $\varphi\colon \mathcal{U} \to \mathbb{R}^k$ with the component vector $\varphi := (\varphi_1,\ldots,\varphi_k)^\top$ such that the function $\phi\colon \mathcal{U} \to \mathcal{M}$

defined via

$$\phi(x) := \bar{p} + \sum_{j=1}^{n-k} x_j \hat{\pi}_j + \sum_{j=1}^{k} \varphi_j(x) \bar{\pi}_j \tag{2.28}$$

for all $x := (x_1, \ldots, x_{n-k})^\top \in \mathcal{U}$ is a local parametrization of $\mathcal{M}$ in a neighbourhood of $\bar{p}$. In this way a family of parametrizations is defined, where each parametrization corresponds to a point $\bar{\eta} \in \mathcal{G}$.[8] Note that there holds $\varphi(0) = 0$ and $\phi(0) = \bar{p}$. Compare figure 2.5.

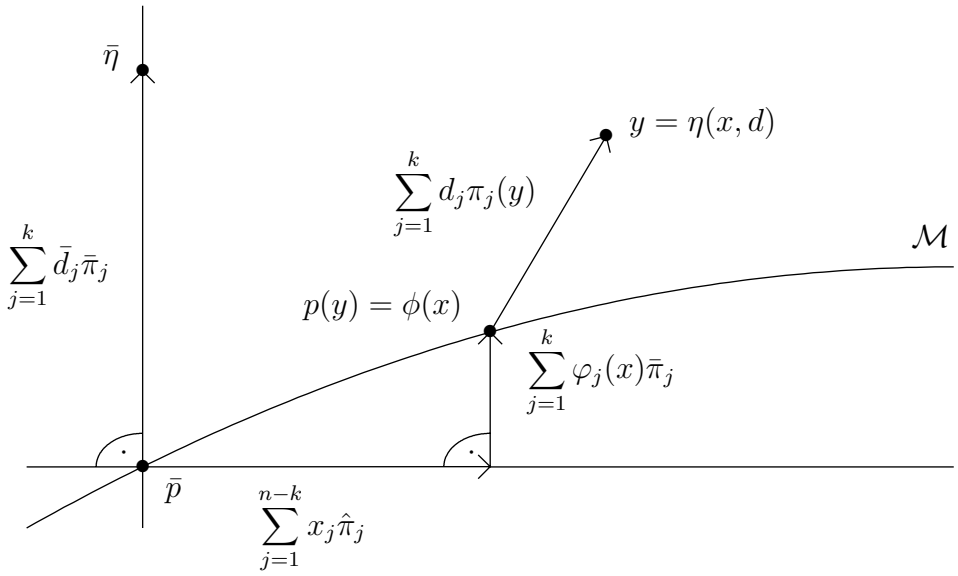


Figure 2.5: A local parametrization of $\mathcal{M}$

- We assume that there exists a moderate real constant $L_\varphi \geq 0$ such that for all parametrizations and all $x, \tilde{x} \in \mathcal{U}$ there holds

$$\|\varphi(x) - \varphi(\tilde{x})\|_2 \leq L_\varphi \|x - \tilde{x}\|_2. \tag{2.29}$$

  Here L_φ has to be interpreted as the supremum of the Lipschitz constants of the functions φ.[9]

- A local coordinate transformation is introduced as follows. In a neighbourhood of each $\bar{\eta} \in \mathcal{G}$ we define the local coordinates $(x, d)(y) := (\phi^{-1}(p(y)), d(y))$ for $y \in \mathcal{V} := \{y \in \mathcal{G} : p(y) \in \phi(\mathcal{U})\}$. Since $(x, d)(\cdot)$ is injective the inverse function $(x, d) \mapsto \eta(x, d)$ can be defined in $\mathcal{W} := (x, d)(\mathcal{V})$. This is the desired local coordinate transformation in a neighbourhood of $\bar{\eta} \in \mathcal{G}$. We call $x(y)$ the *smooth coordinate* of y.

Remark: There holds $\eta(0, 0) = \bar{p}$ and $\eta(0, \bar{d}) = \bar{\eta}$. Furthermore there exists $r > 0$ such that $\mathcal{U} \times \overline{\mathcal{B}}_r \subset \mathcal{W}$.

The next goal is to show the global well-conditionedness concerning the stiff coordinates. Therefore we derive a differential equation for the stiff components $d(y(t))$ corresponding to a solution $y(t)$ of the differential equation (2.11). In the following sections it is assumed that the appearing functions are as often differentiable as the analysis requires.

[8] These parametrizations are suiteable for further numerical analysis. In later sections $\bar{\eta}$ will play the role of the numerical approximations $\eta_{\nu-1}, \nu \in \mathbb{N}$ generated by the underlying method.

[9] The index φ does not mean that L_φ depends on φ!

2.3 The Decrease of the Stiff Components

First we assume the following smoothness concerning the stiff coordinate:

Smoothness assumption concerning the stiff coordinate

- We assume that there exists a moderate real constant $M_{Hd} \geq 0$ such that there holds

$$\|Hd_i(y)\|_2 \leq M_{Hd}, \quad i = 1, \ldots, k \tag{2.30}$$

  for all $y \in \mathcal{G}$.

Remark: This smoothness assumption can be replaced by suiteable smoothness assumptions on $\varphi(\cdot)$, $\pi_i(\cdot), i = 1, \ldots, k$, such that (2.30) is a consequence of these assumptions.

Now let $y(t)$ be a solution of the differential equation (2.11). Differentiation of the components of $d(y(t))$ yields

$$\frac{d}{dt} d_i(y(t)) = \nabla d_i(y(t))^\top y'(t) = \nabla d_i(y(t))^\top f(y(t)), \quad i = 1, \ldots, k.$$

The assumptions on the generalized Jacobian imply the identity [10)]

$$\begin{aligned} f(y) &= f(p(y)) + J(y)(y - p(y)) = \\ &= f(p(y)) + \sum_{j=1}^{k} [\Lambda(y)d(y)]_j \pi_j(y), \quad \forall y \in \mathcal{G}, \end{aligned} \tag{2.31}$$

(compare the definition of $\Lambda(y)$, $d(y)$ in section 2.2, page 11), which leads to

$$\frac{d}{dt} d_i(y(t)) = \nabla d_i(y(t))^\top f\Big(p(y(t))\Big) + \sum_{j=1}^{k} [\Lambda(y(t))d(y(t))]_j \nabla d_i(y(t))^\top \pi_j(y(t)) \tag{2.32}$$

for $i = 1, \ldots, k$. Now the first term of (2.32) is analyzed.

Proposition 2.3.1 *For $i = 1, \ldots, k$ there holds*

$$\nabla d_i(u)^\top f(u) = 0, \quad \forall u \in \mathcal{M}. \tag{2.33}$$

Proof: Let $\tilde{u}(t)$ be a solution of (2.11) in $\mathcal{M}$ with $\tilde{u}(0) = u \in \mathcal{M}$. Differentiation of the identity $d_i(\tilde{u}(t)) \equiv 0$ yields

$$0 \equiv \frac{d}{dt} d_i(\tilde{u}(t)) \equiv \nabla d_i(\tilde{u}(t))^\top \tilde{u}'(t) \equiv \nabla d_i(\tilde{u}(t))^\top f(\tilde{u}(t))$$

for $i = 1, \ldots, k$. Setting $t = 0$ completes the proof. □

[10)] The brackets $[\cdot]_j$ denote the j-th component of a vector.

Equation (2.33) from proposition 2.3.1 leads to

$$\begin{aligned}\nabla d_i(y)^\top f(p(y)) &= f(p(y))^\top \Big(\nabla d_i(y) - \nabla d_i(p(y))\Big) = \\ &= f(p(y))^\top \int_0^1 Hd_i\Big(p(y) + \sigma(y - p(y))\Big)\, d\sigma\, (y - p(y)) = \\ &= \sum_{j=1}^{k} f(p(y))^\top \int_0^1 Hd_i\Big(p(y) + \sigma(y - p(y))\Big)\, d\sigma\, \pi_j(y) d_j(y) \qquad (2.34)\end{aligned}$$

for $i = 1, \ldots, k$ and $y \in \mathcal{G}$. In order to formulate (2.32) in the matrix vector notation we define the $k \times k$-matrix $K(y)$ elementewise as

$$k_{ij}(y) := \sum_{j=1}^{k} f(p(y))^\top \int_0^1 Hd_i\Big(p(y) + \sigma(y - p(y))\Big)\, d\sigma\, \pi_j(y)$$

and the $k \times k$-matrix $L(y)$ with the entries

$$l_{ij}(y) := \nabla d_i(y)^\top \pi_j(y),$$

where $i, j = 1, \ldots, k$ and $y \in \mathcal{G}$. Insertion of (2.34) into (2.32) leads to a linear system of differential equations for $d(y(t))$, which in the vector form reads as

$$\frac{d}{dt} d(y(t)) = \Big(K(y(t)) + L(y(t))\Lambda(y(t))\Big) d(y(t)). \qquad (2.35)$$

First we analyze the marix $K(y)$. The smoothness assumptions (2.12) and (2.30) imply that there exists a moderate real constant $M_K \geq 0$ such that

$$\|K(y)\|_2 \leq M_K, \quad \forall y \in \mathcal{G}. \qquad (2.36)$$

Note that M_K depends on $M_{f|_{\mathcal{M}}}, M_{Hd}$ and on k.
Second we analyze the matrix $L(y)$:

Proposition 2.3.2 *The matrix $L(y)$ can be written in the form*

$$L(y) = I + B(y)d(y), \qquad (2.37)$$

where $B(y) \colon \mathbb{R}^k \to \mathbb{R}^{k \times k}$ is a linear matrix valued mapping which depends on the parameter $y \in \mathcal{G}$. There exists a moderate real constant $M_B \geq 0$ such that

$$\|B(y)\|_2 \leq M_B, \quad \forall y \in \mathcal{G}. \qquad (2.38)$$

The constant M_B depends on L_φ, L_π and on k.

Proof: We choose $\bar{\eta} := y$ and consider the corresponding local coordinates $(x, d)(\cdot)$ in a neighbourhood of $\bar{\eta}$. The entries of $L(\bar{\eta})$ can be computed as the following limit

$$l_{ij}(\bar{\eta}) = \nabla d_i(\bar{\eta})^\top \bar{\pi}_j = \lim_{s \to 0} \frac{d_i(\bar{\eta} + s\bar{\pi}_j) - \bar{d}_i}{s}, \quad i, j = 1, \ldots, k. \qquad (2.39)$$

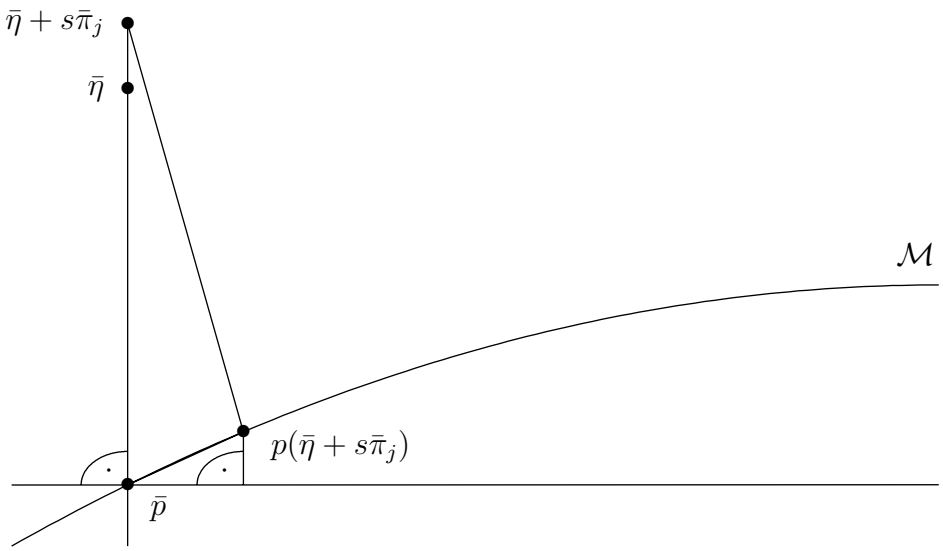


Figure 2.6: The triangle $\bar p, \bar\eta + s\bar\pi_j, p(\bar\eta + s\bar\pi_j)$

Now we express $\bar d$ in terms of s and $d(\bar\eta + s\bar\pi_j)$, for $j = 1, \dots, k$. Compare figure 2.6 concerning the following considerations.
The triangle $\bar p, \zeta_j(s) := \bar\eta + s\bar\pi_j, p(\zeta_j(s))$ with the parameter $s \in \mathbb{R}$ leads to the identity

$$\sum_{l=1, l\neq j}^{k} \bar d_l \bar\pi_l + (\bar d_j + s)\bar\pi_j \; = \tag{2.40}$$

$$= \sum_{l=1}^{n-k} x_l(\zeta_j(s))\hat\pi_l + \sum_{l=1}^{k} \varphi_l\Big(x(\zeta_j(s))\Big)\bar\pi_l + \sum_{l=1}^{k} d_l(\zeta_j(s))\pi_l(\zeta_j(s)).$$

- Application of the scalarproduct with $\hat\pi_m$, $m = 1, \dots, n-k$ to (2.40) yields

$$0 \;=\; x_m(\zeta_j(s)) + \sum_{l=1}^{k} d_l(\zeta_j(s))\langle \pi_l(\zeta_j(s)), \hat\pi_m\rangle \;\Rightarrow$$

$$x_m(\zeta_j(s)) \;=\; -\sum_{l=1}^{k} d_l(\zeta_j(s))\langle \pi_l(\zeta_j(s)), \hat\pi_m\rangle.$$

  We define the $(n-k)\times k$-matrices $\hat R_j(s)$ entrywise $\hat r_{jml}(s) := \langle \pi_l(\zeta_j(s)), \hat\pi_m\rangle$, such that there holds

$$x(\zeta_j(s)) \;=\; -\hat R_j(s) d(\zeta_j(s)), \quad j = 1, \dots, k. \tag{2.41}$$

- Application of the scalarproduct with $\bar\pi_m$, $m = 1, \dots, k$ to (2.40) yields

$$\bar d_m \;=\; \varphi_m\Big(x(\zeta_j(s))\Big) + \sum_{l=1}^{k} d_l(\zeta_j(s))\langle \pi_l(\zeta_j(s)), \bar\pi_m\rangle, \quad m \neq j,$$

$$\bar d_j + s \;=\; \varphi_j\Big(x(\zeta_j(s))\Big) + \sum_{l=1}^{k} d_l(\zeta_j(s))\langle \pi_l(\zeta_j(s)), \bar\pi_j\rangle, \quad m = j.$$

First we note that $\varphi(0) = 0$ implies $\varphi(x) = \int_0^1 J\varphi(\sigma x)d\sigma x$ for all $x \in \mathcal{U}$. This motivates the definition of the matrices

$$S_j(s) := \int_0^1 J\varphi\Big(\sigma x(\zeta_j(s))\Big)\, d\sigma$$

such that together with (2.41) there holds

$$\varphi(x(\zeta_j(s))) = S_j(s)x(\zeta_j(s)) = -S_j(s)\hat{R}_j(s)d(\zeta_j(s)), \quad j = 1, \dots, k.$$

Second we define the matrices $\bar{R}_j(s)$ elementewise by $\bar{r}_{jml}(s) := \langle \pi_l(\zeta_j(s)), \bar{\pi}_m \rangle$. Alltogether this leads to

$$\bar{d} + se_j = (-S_j(s)\hat{R}_j(s) + \bar{R}_j(s))d(\zeta_j(s)), \quad j = 1, \dots, k. \tag{2.42}$$

Insertion of (2.42) into (2.39) yields

$$\frac{d(\zeta_j(s)) - \bar{d}}{s} = e_j + \frac{1}{s}\Big((I - \bar{R}_j(s)) + S_j(s)\hat{R}_j(s)\Big)d(\zeta_j(s)).$$

For $s \to 0$ we obtain

$$\lim_{s\to 0} \frac{d(\zeta_j(s)) - \bar{d}}{s} = e_j + B_j(\bar{\eta})\bar{d},$$

where

$$\begin{aligned} B_j(\bar{\eta}) &:= \lim_{s\to 0} \frac{1}{s}\Big((I - \bar{R}_j(s)) + S_j(s)\hat{R}_j(s)\Big) = \\ &= -\Big(\bar{\pi}_m^\top J\pi_l(\bar{\eta})\bar{\pi}_j\Big)_{ml} + J\varphi(0) \cdot \Big(\hat{\pi}_m^\top J\pi_l(\bar{\eta})\bar{\pi}_l\Big)_{ml}. \end{aligned}$$

There follows

$$L(\bar{\eta}) = I + \Big([B_j(\bar{\eta})\bar{d}]_i\Big)_{ij} = I + B(\bar{\eta})\bar{d},$$

where the linear matrix valued mapping $B(\bar{\eta})\colon \mathbb{R}^k \to \mathbb{R}^{k\times k}$ is defined via $d \to B(\bar{\eta})d := ([B_j(\bar{\eta})d]_i)_{ij}$. The smoothness assumptions (2.19) on $\pi_i(\cdot)$ and (2.29) on $\varphi(\cdot)$ imply that the 2-norms of the matrices $B_j(\bar{\eta})$ are uniformly bounded for $\bar{\eta} \in \mathcal{G}$ with a moderate bound. This implies that there exists a moderate real constant $M_B \geq 0$ such that in addition $\|B(\bar{\eta})\|_2 \leq M_B$ for all $\bar{\eta} \in \mathcal{G}$. Note that M_B depends on L_φ, L_π and on k. □

The identity (2.37) from proposition 2.3.2 allows us to rewrite the differential equation (2.35) for $d(y(t))$ in the following form

$$\frac{d}{dt}d(y(t)) = \Big(K(y(t)) + \Big(I + B(y(t))d(y(t))\Big)\Lambda(y(t))\Big)d(y(t)). \tag{2.43}$$

We insert $d(y(t))$ into the Lyapunov norm

$$\|d(y(t))\|_{V(y(t))} = \langle V(y(t))d(y(t)), d(y(t))\rangle^{\frac{1}{2}}$$

and differentiate the square of the result

$$\begin{aligned} &\langle V(y(t))d(y(t)), d(y(t))\rangle' = \\ &\quad = \langle V(y(t))d(y(t))', d(y(t))\rangle + \langle V(y(t))d(y(t)), d(y(t))'\rangle = \\ &\qquad + \langle V(y(t))'d(y(t)), d(y(t))\rangle + \\ &\quad = \langle V(y(t))'d(y(t)), d(y(t))\rangle + 2\langle V(y(t))d(y(t))', d(y(t))\rangle. \end{aligned} \tag{2.44}$$

We consider the first term of (2.44)

$$\begin{aligned} V(y(t))' &= \Big(\nabla v_{ij}(y(t))^\top y'(t)\Big)_{ij} = \Big(\nabla v_{ij}(y(t))^\top f(y(t))\Big)_{ij} = \\ &= \Big(\nabla v_{ij}(y(t))^\top f\Big(p(y(t))\Big)\Big)_{ij} + \\ &\quad + \sum_{l=1}^{k}\Big(\nabla v_{ij}(y(t))^\top \pi_l(y(t))[\Lambda(y(t))d(y(t))]_l\Big)_{ij} = \\ &= \Big(\nabla v_{ij}(y(t))^\top f\Big(p(y(t))\Big)\Big)_{ij} + \\ &\quad + \frac{1}{\varepsilon}\sum_{l=1}^{k}\Big(\nabla v_{ij}(y(t))^\top \pi_l(y(t))[G(y(t))d(y(t))]_l\Big)_{ij}. \end{aligned}$$

Now we define the $k\times k$-matrix $W(y)$ elementewise by $w_{ij}(y) := \nabla v_{ij}(y)^\top f(p(y))$ as well as the linear operator family

$$\begin{aligned} C(y)\colon \mathbb{R}^k &\to \mathbb{R}^{k\times k} \\ d &\to C(y)d := \sum_{l=1}^{k}(\nabla v_{ij}(y)^\top \pi_l(y)[G(y)d]_l)_{ij}, \end{aligned}$$

where $y \in \mathcal{G}$. This enables us to formulate

$$V(y(t))' = W(y(t)) + \frac{1}{\varepsilon}C(y(t))d(y(t)).$$

Note that due to the smoothness assumptions (2.12), (2.21), (2.26) there exist moderate real constants $M_W, M_C \geq 0$ such that there holds

$$\|W(y)\|_2 \leq M_W, \quad \|C(y)\|_2 \leq M_C, \quad \forall y \in \mathcal{G}.$$

Note that M_W depends on $M_{f|_{\mathcal{M}}}, L_V, k$ and M_C on L_V, M_G, k. These bounds imply the estimate

$$\langle V(y(t))'d(y(t)), d(y(t))\rangle \leq \Big(M_W + \frac{1}{\varepsilon}M_C\|d(y(t))\|_2\Big)\|d(y(t))\|_2^2. \tag{2.45}$$

Now we consider the second term of (2.44). Insertion of (2.43), (2.23) and application of (2.21), (2.36), (2.38) yields

$$\begin{aligned} &2\langle V(y(t))d(y(t))', d(y(t))\rangle = \\ &\quad = 2\langle V(y(t))\Big(K(y(t)) + \Big(I + B(y(t))d(y(t))\Big)\Lambda(y(t))\Big)d(y(t)), d(y(t))\rangle = \\ &\quad = 2\langle V(y(t))\Big(K(y(t)) + \frac{1}{\varepsilon}B(y(t))d(y(t))G(y(t))\Big)d(y(t)), d(y(t))\rangle + \\ &\qquad + \langle\Big(\Lambda(y(t))^\top V(y(t)) + V(y(t))\Lambda(y(t))\Big)d(y(t)), d(y(t))\rangle \leq \\ &\quad \leq 2M_V\Big(M_K + \frac{1}{\varepsilon}M_B M_G\|d(y(t))\|_2\Big)\|d(y(t))\|_2^2 - \frac{1}{\varepsilon}\|d(y(t))\|_2^2. \end{aligned} \tag{2.46}$$

The bounds (2.45), (2.46) enable us to estimate (2.44), i.e.

$$\langle V(y(t))d(y(t)), d(y(t))\rangle' \leq \big(\mathcal{A}_0 + \frac{1}{\varepsilon}\mathcal{A}_1\|d(y(t))\|_2 - \frac{1}{\varepsilon}\big)\|d(y(t))\|_2^2,$$

where $\mathcal{A}_0 := 2M_V M_K + M_W$, $\mathcal{A}_1 := 2M_V M_B M_G + M_C$. For ε sufficiently small, i.e. for

$$\varepsilon \leq \frac{1}{2\mathcal{A}_0} \tag{2.47}$$

there follows

$$\langle V(y(t))d(y(t)), d(y(t))\rangle' \leq -\frac{1}{2\varepsilon}\big(\frac{1}{2} - \mathcal{A}_1\|d(y(t))\|_2\big)\|d(y(t))\|_2^2.$$

Now for $\|d(y(t))\|_2$ sufficiently small, i.e. for

$$\|d(y(t))\|_2 \leq \frac{1}{4\mathcal{A}_1} \tag{2.48}$$

there holds

$$\langle V(y(t))d(y(t)), d(y(t))\rangle' \leq -\frac{1}{2\varepsilon}\|d(y(t))\|_2^2.$$

This leads to

$$\langle V(y(t))d(y(t)), d(y(t))\rangle' \leq -\frac{1}{2\varepsilon M_V}\langle V(y(t))d(y(t)), d(y(t))\rangle. \tag{2.49}$$

We conclude

$$\langle V(y(t))d(y(t)), d(y(t))\rangle \leq \exp(-\frac{1}{2\varepsilon M_V}\, t)\langle V(y_0)d(y_0), d(y_0)\rangle.$$

In the 2-norm there holds

$$\|d(y(t))\|_2 \leq (M_V M_{V^{-1}})^{\frac{1}{2}} \exp(-\frac{1}{4\varepsilon M_V}\, t)\|d(y_0)\|_2. \tag{2.50}$$

Now we set $\mathcal{K}_0 := (M_V M_{V^{-1}})^{\frac{1}{2}}$, $\mathcal{K}_1 := 4M_V$ and prove the following theorem.

Theorem 2.3.1 *For ε, $\|d(y_0)\|_2$ sufficiently small, i.e. (2.47) and*

$$\|d(y_0)\|_2 \leq \frac{1}{4\mathcal{A}_1\mathcal{K}_0} \tag{2.51}$$

there holds

$$\|d(y(t))\|_2 \leq \mathcal{K}_0 \exp(-\frac{1}{\varepsilon\mathcal{K}_1}\, t)\|d(y_0)\|_2 \tag{2.52}$$

for all $t \geq 0$ with $y([0,t]) \subset \mathcal{G}$.

Proof: Without loss of generality let $\|d(y_0)\|_2 > 0$. Since $\mathcal{K}_0 \geq 1$ the restriction (2.51) ensures that (2.48) is fulfilled for t_0. Then (2.49) implies

$$\langle V(y_0)d(y_0), d(y_0)\rangle' < 0,$$

which by continuity yields

$$\langle V(y(t))d(y(t)), d(y(t))\rangle < \langle V(y_0)d(y_0), d(y_0)\rangle$$

at least on a short intervall. There follows $\|d(y(t))\|_2 < \mathcal{K}_0\|d(y_0)\|_2 \leq \frac{1}{4\mathcal{A}_1}$, i.e. (2.48) at least on a short intervall. Now assume that (2.48) is violated for some $t > 0$ then there exists a first $t^* > 0$ such that $\|d(y(t^*))\|_2 = \frac{1}{4\mathcal{A}_1}$, but by (2.50) there holds $\|d(y(t^*))\|_2 < \mathcal{K}_0\|d(y_0)\|_2 \leq \frac{1}{4\mathcal{A}_1}$ which is a contradiction. □

2.4 Some Estimates Arising from the Definition of the Problem Class

In this section we derive some basic estimates which arise from the assumptions on the differential equation. These estimates are further used in the existence and uniqueness proof for the algebraic equations as well as for the convergence proof of the Runge-Kutta methods.

Proposition 2.4.1 *Let $L_\phi := 1 + L_\varphi$, then there holds*

$$\|\phi(x) - \phi(\tilde{x})\|_2 \;\leq\; L_\phi \|x - \tilde{x}\|_2 \tag{2.53}$$

for all $x, \tilde{x} \in \mathcal{U}$ and all local parametrizations ϕ of $\mathcal{M}$.

Proof: The definition of ϕ and the smoothness assumption (2.29) on φ yields

$$\begin{aligned}
\|\phi(x) - \phi(\tilde{x})\|_2 &= \|\sum_{j=1}^{n-k}(x_j - \tilde{x}_j)\hat{\pi}_j + \sum_{j=1}^{k}(\varphi_j(x) - \varphi_j(\tilde{x}))\bar{\pi}_j\|_2 = \\
&= \|(x - \tilde{x}, \varphi(x) - \varphi(\tilde{x}))\|_2 \leq \\
&\leq \|x - \tilde{x}\|_2 + \|\varphi(x) - \varphi(\tilde{x})\|_2 \leq \\
&\leq \|x - \tilde{x}\|_2 + L_\varphi \|x - \tilde{x}\|_2,
\end{aligned}$$

for all $x, \tilde{x} \in \mathcal{U}$. □

Now we introduce the norm

$$\|(x, d)\|_* \;:=\; \max\{\|x\|_2, \|d\|_2\}$$

for all $(x, d) \in \mathbb{R}^{n-k} \times \mathbb{R}^k$.

Proposition 2.4.2 *Let $L_\eta := 2(L_\phi + 1)$, then there holds*

$$\|\eta(x,d) - \eta(\tilde{x},\tilde{d})\|_2 \;\leq\; L_\eta \|(x,d) - (\tilde{x},\tilde{d})\|_*, \tag{2.54}$$

$$\|\pi_i(\eta(x,d)) - \pi_i(\eta(\tilde{x},\tilde{d}))\|_2 \;\leq\; L_\pi L_\eta \|(x,d) - (\tilde{x},\tilde{d})\|_*, \quad i = 1, \ldots, k, \tag{2.55}$$

$$\|G(\eta(x,d)) - G(\eta(\tilde{x},\tilde{d}))\|_2 \;\leq\; k L_G L_\eta \|(x,d) - (\tilde{x},\tilde{d})\|_*, \tag{2.56}$$

$$\|\Lambda(\eta(x,d)) - \Lambda(\eta(\tilde{x},\tilde{d}))\|_2 \;\leq\; \frac{k L_G L_\eta}{\varepsilon} \|(x,d) - (\tilde{x},\tilde{d})\|_*, \tag{2.57}$$

$$\|V(\eta(x,d)) - V(\eta(\tilde{x},\tilde{d}))\|_2 \;\leq\; k L_V L_\eta \|(x,d) - (\tilde{x},\tilde{d})\|_* \tag{2.58}$$

for all local coordinates $(x, d), (\tilde{x}, \tilde{d}) \in \mathcal{W}$ with $\|\tilde{d}\|_2 \leq \frac{1}{2\sqrt{k} L_\pi}$.

Proof: From the definition of $\eta(x, d)$ there follows

$$\eta(x,d) \;=\; \phi(x) + \sum_{j=1}^{k} d_j \pi_j(\eta(x,d)),$$

for all $(x,d) \in \mathcal{W}$. This identity, the smoothness assumption (2.19) and (2.53) from proposition 2.4.1 can be used to derive

$$\begin{aligned}
&\|\eta(x,d) - \eta(\tilde{x},\tilde{d})\|_2 \; = \\
&\quad = \; \|\phi(x) - \phi(\tilde{x}) + \sum_{j=1}^{k} \Big(d_j \pi_j(\eta(x,d)) - \tilde{d}_j \pi_j(\eta(\tilde{x},\tilde{d})) \Big) \|_2 \; \leq \\
&\quad \leq \; L_\phi \|x - \tilde{x}\|_2 + \| \sum_{j=1}^{k} (d_j - \tilde{d}_j) \pi_j(\eta(x,d)) \|_2 + \\
&\qquad + \; \| \sum_{j=1}^{k} \tilde{d}_j \Big(\pi_j(\eta(x,d)) - \pi_j(\eta(\tilde{x},\tilde{d})) \Big) \|_2 \; \leq \\
&\quad \leq \; L_\phi \|x - \tilde{x}\|_2 + \|d - \tilde{d}\|_2 + \sqrt{k} L_\pi \|\eta(x,d) - \eta(\tilde{x},\tilde{d})\|_2 \|\tilde{d}\|_2
\end{aligned}$$

for all $(x,d), (\tilde{x},\tilde{d}) \in \mathcal{W}$. Under the restriction $\|\tilde{d}\|_2 \leq \frac{1}{2\sqrt{k} L_\pi}$ there follows

$$\|\eta(x,d) - \eta(\tilde{x},\tilde{d})\|_2 \; \leq \; 2(L_\phi \|x - \tilde{x}\|_2 + \|d - \tilde{d}\|_2)$$

for $(x,d), (\tilde{x},\tilde{d}) \in \mathcal{W}$. This proves (2.54). The estimates (2.55)-(2.58) are immediate consequences of (2.54) and the smoothness assumptions (2.19), (2.22), (2.26). □

Notation: Now we denote the coordinates of $f \circ \phi$ corresponding to the ONS B. To each local parametrization ϕ of $\mathcal{M}$ we introduce the function $\hat{\psi} : \mathcal{U} \to \mathbb{R}^{n-k}$ where $\hat{\psi} := (\hat{\psi}_1, \ldots, \hat{\psi}_{n-k})^\top$ is defined via

$$\hat{\psi}_j(x) \; := \; \langle f(\phi(x)), \hat{\pi}_j \rangle, \quad j = 1, \ldots, n-k \tag{2.59}$$

for $x \in \mathcal{U}$. In the same way we define the funcion $\psi : \mathcal{U} \to \mathbb{R}^k$ via the components

$$\psi_j(x) \; := \; \langle f(\phi(x)), \bar{\pi}_j \rangle, \quad j = 1, \ldots, k \tag{2.60}$$

for $x \in \mathcal{U}$. These coordinates first appear in the next section, where the algebraic equation for the implicit Euler method is parametrized via the local coordinates (x,d).

Proposition 2.4.3 *Let $L_\psi := L_{f|\mathcal{M}} L_\phi$, then there holds*

$$\left. \begin{array}{l} \|\hat{\psi}(x)\|_2 \\ \|\psi(x)\|_2 \end{array} \right\} \; \leq \; \|f(\phi(x))\|_2 \; \leq \; M_{f|\mathcal{M}}, \tag{2.61}$$

and

$$\left. \begin{array}{l} \|\hat{\psi}(x) - \hat{\psi}(\tilde{x})\|_2 \\ \|\psi(x) - \psi(\tilde{x})\|_2 \end{array} \right\} \; \leq \; \|f(\phi(x)) - f(\phi(\tilde{x}))\|_2 \; \leq \; L_\psi \|x - \tilde{x}\|_2, \tag{2.62}$$

for all $x, \tilde{x} \in \mathcal{U}$ and all local parametrizations ϕ

Proof: The inequalities are immediate consequences of the identity

$$f(\phi(x)) \; = \; \sum_{j=1}^{n-k} \hat{\psi}_j(x) \hat{\pi}_j + \sum_{j=1}^{k} \psi_j(x) \bar{\pi}_j$$

and the smoothness assumptions (2.12), (2.13) and (2.53) from proposition 2.4.1. □

Notation: Now we introduce the orthogonal projections from the space $\mathcal{E}(\eta(x,d))$ into $\mathcal{E}(\bar{\eta})^{\perp}$ and into $\mathcal{E}(\bar{\eta})$. The projections are parametrized via the local coordinates (x,d). The first projection $\hat{\Theta}(x,d)$ is defined entrywise as

$$\hat{\theta}_{ij}(x,d) \;:=\; \langle \pi_j(\eta(x,d)), \hat{\pi}_i \rangle, \quad i=1,\dots,n-k,\; j=1,\dots,k \tag{2.63}$$

for all $(x,d) \in \mathcal{W}$. The second projection $\Theta(x,d)$ is defined entrywise as

$$\theta_{ij}(x,d) \;:=\; \langle \pi_j(\eta(x,d)), \bar{\pi}_i \rangle, \quad i,j=1,\dots,k \tag{2.64}$$

for all $(x,d) \in \mathcal{W}$. These projectios first appear in the context of the parametrization of the algebraic equation of the implicit Euler method. Note that $\hat{\Theta}(0,\bar{d}) = 0$ and $\Theta(0,\bar{d}) = I$.

Proposition 2.4.4 *Let $L_\Theta := \sqrt{k} L_\pi L_\eta$, then there holds*

$$\|\hat{\Theta}(x,d)\|_2, \|\Theta(x,d)\|_2 \;\leq\; 1, \tag{2.65}$$

$$\left.\begin{aligned} \|\hat{\Theta}(x,d) - \hat{\Theta}(\tilde{x},\tilde{d})\|_2 \\ \|\Theta(x,d) - \Theta(\tilde{x},\tilde{d})\|_2 \end{aligned}\right\} \;\leq\; L_\Theta \|(x,d) - (\tilde{x},\tilde{d})\|_*, \tag{2.66}$$

$$\left.\begin{aligned} \|\hat{\Theta}(\tilde{x},\tilde{d})\|_2 \\ \|I - \Theta(\tilde{x},\tilde{d})\|_2 \end{aligned}\right\} \;\leq\; 2L_\Theta \max\{\|(\tilde{x},\tilde{d})\|_*, \|\bar{d}\|_2\} \tag{2.67}$$

for all $(x,d),(\tilde{x},\tilde{d}) \in \mathcal{W}$ with $\|\tilde{d}\|_2 \leq \frac{1}{2\sqrt{k}L_\pi}$.

Proof: The estimates in (2.65) follow from the definition of $\hat{\Theta}(x,d)$, $\Theta(x,d)$. The first part of (2.66) follows from

$$\begin{aligned} \|\hat{\Theta}(x,d) - \hat{\Theta}(\tilde{x},\tilde{d})\|_2 \;&=\; \|\Big(\langle \hat{\pi}_i, \pi_j(\eta(x,d)) - \pi_j(\eta(\tilde{x},\tilde{d}))\rangle\Big)_{ij}\|_2 \;\leq \\ &\leq\; \sqrt{k} \max_j \|\Big(\langle \hat{\pi}_i, \pi_j(\eta(x,d)) - \pi_j(\eta(\tilde{x},\tilde{d}))\rangle\Big)_{i}\|_2 \;\leq \\ &\leq\; \sqrt{k} \max_j \|\pi_j(\eta(x,d)) - \pi_j(\eta(\tilde{x},\tilde{d}))\|_2 \;\leq \\ &\leq\; \sqrt{k} L_\pi L_\eta \|(x,d) - (\tilde{x},\tilde{d})\|_*. \end{aligned}$$

The second part of (2.66) can be derived in the same way. Now (2.67) is a consequence of (2.66) and $\hat{\Theta}(0,\bar{d}) = 0$, $\Theta(0,\bar{d}) = I$. □

The following proposition describes the change of the Lyapunov norm.

Proposition 2.4.5 *Let $L_{V^{\frac{1}{2}}} := k L_V L_{\surd}$, where $L_{\surd}$ is defined as in proposition A.2.1, then there holds*

$$\Big| \|\cdot\|_{V(\eta(x,d))} - \|\cdot\|_{V(\eta(\tilde{x},\tilde{d}))} \Big| \;\leq\; L_{V^{\frac{1}{2}}} L_\eta \|(x,d) - (\tilde{x},\tilde{d})\|_* \|\cdot\|_2 \tag{2.68}$$

for all $(x,d),(\tilde{x},\tilde{d}) \in \mathcal{W}$ with $\|\tilde{d}\|_2 \leq \frac{1}{2\sqrt{k}L_\pi}$.

Proof: We use (A.13) of proposition A.2.1 from the appendix and (2.58) to derive

$$\begin{aligned}
&\left| \|\cdot\|_{V(\eta(x,d))} - \|\cdot\|_{V(\eta(\tilde{x},\tilde{d}))} \right| = \\
&\quad = \left| \|V(\eta(x,d))^{\frac{1}{2}} \cdot \|_2 - \|V(\eta(\tilde{x},\tilde{d}))^{\frac{1}{2}} \cdot \|_2 \right| \leq \\
&\quad \leq \|V(\eta(x,d))^{\frac{1}{2}} - V(\eta(\tilde{x},\tilde{d}))^{\frac{1}{2}}\|_2 \| \cdot \|_2 \leq \\
&\quad \leq L_{\sqrt{}} k L_V L_\eta \|(x,d) - (\tilde{x},\tilde{d})\|_* \| \cdot \|_2
\end{aligned}$$

for all $(x,d),(\tilde{x},\tilde{d}) \in \mathcal{W}$ with $\|\tilde{d}\|_2 \leq \frac{1}{2\sqrt{k}L_\pi}$. □

Chapter 3

The Implicit Euler Method

Let $u(t) \in \mathcal{M}$, $t \in [0,T]$ be a smooth solution of the differential equation (2.11) with initial value $u(0) = u_0 \in \mathcal{M}$. The solution $u(t)$ is approximated by the implicit Euler method. A discretization of the intervall $[0,T]$ is given by $t_\nu := \nu h$, $\nu \in \mathbb{N}_0$, where h is the constant step size. Starting with an initial approximation $\eta_0 \in \mathcal{G}$ for u_0 further approximations η_ν for $u(t_\nu)$ are defined via the equation

$$\eta_\nu \;=\; \eta_{\nu-1} + hf(\eta_\nu). \tag{3.1}$$

Figure 3.1 shows one step of the implicit Euler method.

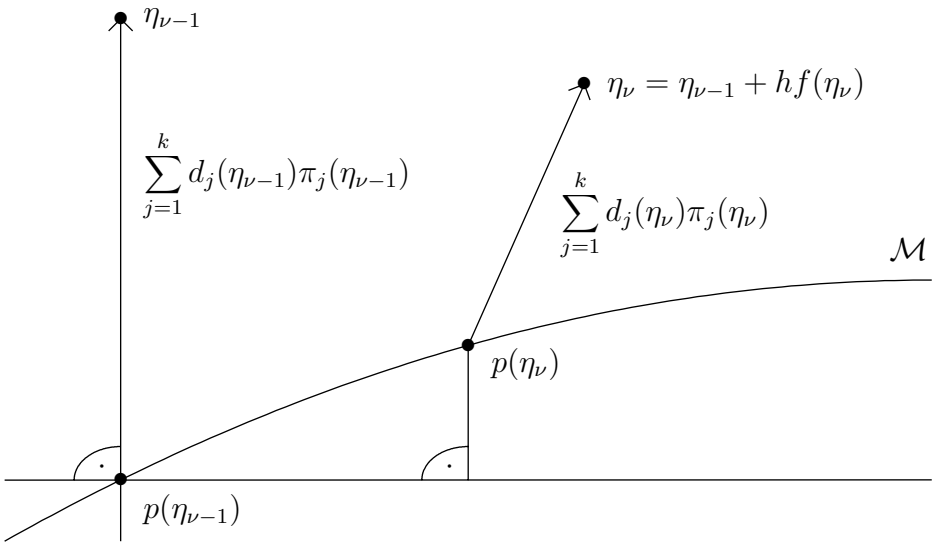


Figure 3.1: The implicit Euler method

To be shure that a step from $\eta_{\nu-1}$ to η_ν is well defined it has to be shown that the algebraic equation

$$\eta \;=\; \bar{\eta} + hf(\eta) \tag{3.2}$$

posesses a locally unique solution $\eta \in \mathcal{G}$, where $\bar{\eta} := \eta_{\nu-1} \in \mathcal{G}$ is given. Therefore the equation (3.2) is parametrized via $\eta(x,d)$ in a neighbourhood of $\bar{\eta}$. Then a transformation into a suiteable fixed point form for (x,d) and application of the fixed point theorem A.1.1 leads to existence and local uniqueness of a solution (x,d) of the fixed point equation. Then $\eta(x,d)$ is the locally unique solution of the algebraic equation (3.2).

3.1 Parametrization of the Algebraic Equation

In the following considerations we make use of the short notation

$$\bar{p} := p(\bar{\eta}), \quad \bar{d} := d(\bar{\eta}), \quad \bar{\pi}_j := \pi_j(\bar{\eta}),$$
$$p := p(\eta), \quad d := d(\eta), \quad \pi_j := \pi_j(\eta)$$

and $\Lambda := \Lambda(\eta)$, where $\bar{\eta}, \eta \in \mathcal{G}$. The identities

$$\bar{\eta} \;=\; \bar{p} + \sum_{j=1}^{k} \bar{d}_j \bar{\pi}_j, \quad \eta \;=\; p + \sum_{j=1}^{k} d_j \pi_j, \quad f(\eta) \;=\; f(p) + \sum_{j=1}^{k} [\Lambda d]_j \pi_j$$

for $\bar{\eta}, \eta \in \mathcal{G}$, which hold due to the assumptions in section 2.2 (compare page 11), lead to the following formulation of the algebraic equation (3.2):

$$p + \sum_{j=1}^{k} d_j \pi_j \;=\; \bar{p} + \sum_{j=1}^{k} \bar{d}_j \bar{\pi}_j + h\Big(f(p) + \sum_{j=1}^{k} [\Lambda d]_j \pi_j\Big). \tag{3.3}$$

The formulations (3.3) [1)] and (3.2) are equivalent in the sense that the solutions coincide.
Now we use the local parametrization $\phi(x)$ of $\mathcal{M}$ and the local coordinate transformation $\eta(x,d)$ corresponding to $\bar{\eta}$. Since $\eta(x,d)$ is a bijection from $\mathcal{W}$ to $\mathcal{V}$ we can replace the functions $p = p(\eta)$, $d = d(\eta)$, $\pi_j = \pi_j(\eta)$ and $\Lambda = \Lambda(\eta)$ for $\eta \in \mathcal{V}$ by $p = p(\eta(x,d))$, $d = d(\eta(x,d))$, $\pi_j = \pi_j(\eta(x,d))$ and $\Lambda = \Lambda(\eta(x,d))$, where $(x,d) \in \mathcal{W}$. Because of $p(\eta(x,d)) = \phi(x)$ the short notation $\varphi := \varphi(x)$ allows us to replace p on the left hand side of (3.3) by

$$p \;=\; \phi(x) \;=\; \bar{p} + \sum_{j=1}^{n-k} x_j \hat{\pi}_j + \sum_{j=1}^{k} \varphi_j \bar{\pi}_j.$$

This eliminates $\bar{p}$ in (3.3). Furthermore we relace p in $f(p)$ by $\phi(x)$. For the coordinate vectors of $f(p) = f(\phi(x))$ we use the short notation

$$\hat{\psi} := \hat{\psi}(x), \quad \psi := \psi(x).$$

This allows us to replace $f(p)$ on the right hand side of (3.3) by

$$f(p) \;=\; \sum_{j=1}^{n-k} \hat{\psi}_j \hat{\pi}_j + \sum_{j=1}^{k} \psi_j \bar{\pi}_j.$$

The resulting equations for $(x,d) \in \mathcal{W}$ can be formulated in the ONS B:

$$x_i + \sum_{j=1}^{k} d_j \langle \pi_j, \hat{\pi}_i \rangle \;=\; h\Big(\hat{\psi}_i + \sum_{j=1}^{k} [\Lambda d]_j \langle \pi_j, \hat{\pi}_i \rangle\Big), \tag{3.4}$$

$$\varphi_i + \sum_{j=1}^{k} d_j \langle \pi_j, \bar{\pi}_i \rangle \;=\; \bar{d}_i + h\Big(\psi_i + \sum_{j=1}^{k} [\Lambda d]_j \langle \pi_j, \bar{\pi}_i \rangle\Big). \tag{3.5}$$

Now let for short

$$\hat{\Theta} \;:=\; \hat{\Theta}(x,d), \quad \Theta \;:=\; \Theta(x,d),$$

[1)] Note that in this formulation the unknown is $\eta \in \mathcal{G}$, which is "hidden" due to the short notation.

such that $\hat{\Theta} = (\langle \pi_j, \hat{\pi}_i \rangle)_{ij}$ and $\Theta = (\langle \pi_j, \bar{\pi}_i \rangle)_{ij}$ in our short notation. Then the matrix-vector notation of the parametrized euler equations (3.4), (3.5) reads as follows

$$x + \hat{\Theta} d = h(\hat{\psi} + \hat{\Theta}\Lambda d), \tag{3.6}$$

$$\varphi + \Theta d = \bar{d} + h(\psi + \Theta\Lambda d). \tag{3.7}$$

In this formulation the unknowns are the coordinates (x, d).

Proposition 3.1.1 *Let (x, d) be the unique solution of (3.6), (3.7) in a neighbourhood $\mathcal{B} \subseteq \mathcal{W}$ of $0 \in I\!R^{n-k} \times I\!R^k$, then $\eta := \eta(x, d)$ is the unique solution of (3.2) in $\eta(\mathcal{B})$.*

Proof: $\eta(\cdot, \cdot)$ is a bijection from $\mathcal{B}$ to $\eta(\mathcal{B})$, which implies the proposition. □

The next step is to find an appropriate fixed point form for the parametrized equations. Therefore we consider the equations (3.6), (3.7) in a closed ball $\overline{\mathcal{B}}_r \subset \mathcal{W}$, where

$$\mathcal{B}_r := \{(x, d) \colon \|(x, d)\|_* < r\}$$

and $r > 0$. The parameter r will be fixed in section 3.3. Here we derive some restrictions on r such that the fixed point theorem can be applied. In order to enable a reformulation of equation (3.7) we rewrite (3.7) as

$$\Theta(I - h\Lambda)d = \bar{d} - \varphi + h\psi \tag{3.8}$$

and prove the regularity of the matrix Θ.

Proposition 3.1.2 *Let $\overline{\mathcal{B}}_r \subset \mathcal{W}$, where r is restricted by*

$$r \leq \min\left\{\frac{1}{2\sqrt{k}L_\pi}, \frac{1}{4L_\Theta}\right\}. \tag{3.9}$$

In addition let $\|\bar{d}\|_2$ be restricted by

$$\|\bar{d}\|_2 \leq \frac{1}{4L_\Theta}, \tag{3.10}$$

then $\Theta(x, d)$ is regular for $(x, d) \in \overline{\mathcal{B}}_r$ and the norm of the inverse satisfies the bound

$$\|\Theta(x, d)^{-1}\|_2 \leq 2, \quad \forall (x, d) \in \overline{\mathcal{B}}_r. \tag{3.11}$$

Proof: Since $\Theta(0, \bar{d}) = I$ and by (2.67) of proposition 2.4.4 there holds

$$\begin{aligned} \|\Theta(x, d)\tilde{d}\|_2 &\geq (1 - \|\Theta(x, d) - \Theta(0, \bar{d})\|_2)\|\tilde{d}\|_2 \geq \\ &\geq (1 - 2L_\Theta \max\{\|(x, d)\|_*, \|\bar{d}\|_2\})\|\tilde{d}\|_2 \geq \frac{1}{2}\|\tilde{d}\|_2 \end{aligned}$$

for all $(x, d) \in \overline{\mathcal{B}}_r$, $\tilde{d} \in \mathbb{R}^k$. □

The following fixed point form arises

$$x = h\hat{\psi} - \hat{\Theta}(I - h\Lambda)d, \tag{3.12}$$

$$d = (I - h\Lambda)^{-1}\Theta^{-1}(\bar{d} - \varphi + h\psi) \tag{3.13}$$

where the regularity of Θ follows from proposition 3.1.2 and the regularity of $I - h\Lambda$ follows from the location of the spectrum of Λ. In order to obtain contractivity further manipulations are useful. Insertion of (3.13) into (3.12) yields

$$x \;=\; h\hat{\psi} - \hat{\Theta}\Theta^{-1}(\bar{d} - \varphi + h\psi). \tag{3.14}$$

Then we replace the argument x of $\varphi = \varphi(x)$ in (3.13) by the right hand side of (3.14) which results in the desired fixed point form. The function $F := (F_1, F_2)$ of the fixed point form

$$(x, d) \;=\; F(x, d) \tag{3.15}$$

is defined as

$$\begin{aligned} F_1 \;&:=\; h\hat{\psi} - H_1(\bar{d} - \varphi + h\psi), \\ F_2 \;&:=\; H_2(\bar{d} - \varphi \circ F_1 + h\psi), \end{aligned}$$

where

$$H_1 \;:=\; \hat{\Theta}\Theta^{-1}, \quad H_2 \;:=\; (I - h\Lambda)^{-1}\Theta^{-1}.$$

In the next section it is shown that the fixed point theorem A.1.1 can be applied to obtain existence and uniqueness for (3.15).[2)]

3.2 Solvability of the Algebraic Equation

Our next goal is to apply the fixed point theorem A.1.1 to solve the algebraic equation. Therefore we have to show that the function $F(x, d)$ fulfills the assumptions of the fixed point theorem. First we consider the F-differences to obtain conditions for contractivity. The following propositions are required.

Proposition 3.2.1 *Let $M_0 := (M_V M_{V^{-1}})^{\frac{1}{2}}$ and $M_1 := 2M_V$, then there holds*

$$\left\| \Big(I - h\Lambda(\eta(x, d)) \Big)^{-1} \right\|_2 \;\leq\; \frac{M_0}{1 + \frac{h}{\varepsilon M_1}} \tag{3.16}$$

for all $(x, d) \in \mathcal{W}$.

Proof: From (2.23) and (2.25) there follows

$$\|e^{h\Lambda t}\|_2 \;\leq\; M_0 \exp(-\frac{h}{\varepsilon M_1} t), \quad \forall t \geq 0,$$

where we use $\Lambda := \Lambda(\eta(x, d))$ for short.[3)] The inverse Laplace-transformation leads to

$$\|(I - h\Lambda)^{-1}\|_2 \;=\; \| \int_0^\infty e^{-t} e^{h\Lambda t} dt \|_2 \;\leq\; \frac{M_0}{1 + \frac{h}{\varepsilon M_1}}.$$

□

[2)] Note that the manipulations which led to (3.15) leave the solution set invariant (compare section A.1.1 in the appendix).

[3)] Compare [6] for more details.

Proposition 3.2.2 *Let $\overline{\mathcal{B}}_r \subset \mathcal{W}$, $0 < r \leq \frac{1}{2\sqrt{k}L_\pi}$ and $L_{(I-h\Lambda)^{-1}} := \frac{k}{4}M_1 L_G L_\eta M_0^2$, then there holds*

$$\begin{aligned} &\Big\|\Big(I - h\Lambda(\eta(x,d))\Big)^{-1} - \Big(I - h\Lambda(\eta(\tilde{x},\tilde{d}))\Big)^{-1}\Big\|_2 \leq \\ &\leq L_{(I-h\Lambda)^{-1}}\|(x,d) - (\tilde{x},\tilde{d})\|_* \end{aligned} \tag{3.17}$$

for all $(x,d),(\tilde{x},\tilde{d}) \in \overline{\mathcal{B}}_r$.

Proof: Under the short notation $\Lambda := \Lambda(\eta(x,d))$, $\tilde{\Lambda} := \Lambda(\eta(\tilde{x},\tilde{d}))$ and with the help of (3.16), (2.57) there holds

$$\begin{aligned} \|(I - h\Lambda)^{-1} - (I - h\tilde{\Lambda})^{-1}\|_2 &\leq \|(I - h\Lambda)^{-1}\|_2\|(I - h\tilde{\Lambda})^{-1}\|_2 h\|\Lambda - \tilde{\Lambda}\|_2 \leq \\ &\leq \Big(\frac{M_0}{1 + \frac{h}{\varepsilon M_1}}\Big)^2 \frac{h}{\varepsilon} k L_G L_\eta \|(x,d) - (\tilde{x},\tilde{d})\|_* \end{aligned}$$

for all $(x,d),(\tilde{x},\tilde{d}) \in \overline{\mathcal{B}}_r$. We define

$$L_{(I-h\Lambda)^{-1}} := \sup_{\frac{h}{\varepsilon}>0} \frac{\frac{h}{\varepsilon} k L_G L_\eta M_0^2}{(1 + \frac{h}{\varepsilon M_1})^2} = \frac{k}{4} M_1 L_G L_\eta M_0^2,$$

which completes the proof. □

Proposition 3.2.3 *Let $L_{\Theta^{-1}} := 4L_\Theta$, then under the assumptions of proposition 3.1.2 there holds*

$$\|\Theta(x,d)^{-1} - \Theta(\tilde{x},\tilde{d})^{-1}\|_2 \leq L_{\Theta^{-1}}\|(x,d) - (\tilde{x},\tilde{d})\|_* \tag{3.18}$$

for all $(x,d),(\tilde{x},\tilde{d}) \in \overline{\mathcal{B}}_r$.

Proof: Under the short notation $\Theta := \Theta(x,d)$, $\tilde{\Theta} := \Theta(\tilde{x},\tilde{d})$ and with the help of (3.11), (2.66) there holds

$$\begin{aligned} \|\Theta^{-1} - \tilde{\Theta}^{-1}\|_2 &\leq \|\Theta^{-1}\|_2\|\tilde{\Theta}^{-1}\|_2\|\Theta - \tilde{\Theta}\|_2 \leq \\ &\leq 4L_\Theta\|(x,d) - (\tilde{x},\tilde{d})\|_* \end{aligned}$$

for all $(x,d),(\tilde{x},\tilde{d}) \in \overline{\mathcal{B}}_r$. □

Proposition 3.2.4 *Let $L_{H_1} := 2L_\Theta + L_{\Theta^{-1}}$ and $L_{H_2} := 2L_{(I-h\Lambda)^{-1}} + M_0 L_{\Theta^{-1}}$ and $M_{H_2} := 2M_0$. Under the assumptions of proposition 3.1.2 there holds*

$$\|H_1(x,d) - H_1(\tilde{x},\tilde{d})\|_2 \leq L_{H_1}\|(x,d) - (\tilde{x},\tilde{d})\|_*, \tag{3.19}$$

$$\|H_2(x,d) - H_2(\tilde{x},\tilde{d})\|_2 \leq L_{H_2}\|(x,d) - (\tilde{x},\tilde{d})\|_*, \tag{3.20}$$

$$\|H_1(x,d)\|_2 \leq \begin{cases} 2L_{H_1}\max\{r, \|\bar{d}\|_2\}, \\ 2, \end{cases} \tag{3.21}$$

$$\|H_2(x,d)\|_2 \leq \frac{2M_0}{1 + \frac{h}{\varepsilon M_1}} \leq M_{H_2} \tag{3.22}$$

for all $(x,d),(\tilde{x},\tilde{d}) \in \overline{\mathcal{B}}_r$.

Proof: We use the short notation $H_1 := H_1(x,d), \tilde{H}_1 := H_1(\tilde{x},\tilde{d}), \ldots$, where the tilde-symbol indicates the argument $(\tilde{x},\tilde{d})$. In this notation there holds

$$\begin{aligned} H_1 - \tilde{H}_1 &= \hat{\Theta}\Theta^{-1} - \tilde{\hat{\Theta}}\tilde{\Theta}^{-1} = \\ &= (\hat{\Theta} - \tilde{\hat{\Theta}})\Theta^{-1} + \tilde{\hat{\Theta}}(\Theta^{-1} - \tilde{\Theta}^{-1}). \end{aligned}$$

Application of (2.66), (3.11), (2.65) and (3.18) yields

$$\|H_1 - \tilde{H}_1\|_2 \leq (2L_\Theta + L_{\Theta^{-1}})\|(x,d) - (\tilde{x},\tilde{d})\|_*$$

for all $(x,d),(\tilde{x},\tilde{d}) \in \overline{\mathcal{B}}_r$. In the same way

$$\begin{aligned} H_2 - \tilde{H}_2 &= (I - h\Lambda)^{-1}\Theta^{-1} - (I - h\tilde{\Lambda})^{-1}\tilde{\Theta}^{-1} = \\ &= \Big((I - h\Lambda)^{-1} - (I - h\tilde{\Lambda})^{-1}\Big)\Theta^{-1} + (I - h\tilde{\Lambda})^{-1}\Big(\Theta^{-1} - \tilde{\Theta}^{-1}\Big) \end{aligned}$$

and (3.17), (3.16) leads to

$$\|H_2 - \tilde{H}_2\|_2 \leq \Big(2L_{(I-h\Lambda)^{-1}} + \frac{M_0}{1 + \frac{1}{\varepsilon M_1}} L_{\Theta^{-1}}\Big)\|(x,d) - (\tilde{x},\tilde{d})\|_*$$

for all $(x,d),(\tilde{x},\tilde{d}) \in \overline{\mathcal{B}}_r$. Setting $(\tilde{x},\tilde{d}) := (0,\bar{d})$ the first part of (3.21) follows from (3.19). The second part of (3.21) can be obtained from (3.11) and (2.65). The bound (3.22) is a consequence of (3.16), (3.11). $\square$.

Proposition 3.2.5 *There exist moderate real constants $B_{F_1}, C_{F_1}, B_{F_2} \geq 0$, such that under the assumptions of proposition 3.1.2 there holds*

$$\|F_1(x,d) - F_1(\tilde{x},\tilde{d})\|_2 \leq B_{F_1} \max\{h, \|\bar{d}\|_2, r\}\|(x,d) - (\tilde{x},\tilde{d})\|_*, \tag{3.23}$$

$$\|F_1(x,d)\|_2 \leq C_{F_1} \max\{h, \|\bar{d}\|_2, r\}, \tag{3.24}$$

$$\|F_2(x,d) - F_2(\tilde{x},\tilde{d})\|_2 \leq B_{F_2} \max\{h, \|\bar{d}\|_2, r\}\|(x,d) - (\tilde{x},\tilde{d})\|_*, \tag{3.25}$$

for all $(x,d),(\tilde{x},\tilde{d}) \in \overline{\mathcal{B}}_r$. The constants $B_{F_1}, C_{F_1}, B_{F_2}$ can be chosen as stated in the proof.

Proof: Under the short notation $F_1 := F_1(x,d), \tilde{F}_1 := F_1(\tilde{x},\tilde{d}), \ldots$ as usual, there holds

$$\begin{aligned} F_1 - \tilde{F}_1 &= h(\hat{\psi} - \tilde{\hat{\psi}}) - (H_1 - \tilde{H}_1)(\bar{d} - \varphi + h\psi) - \\ &\quad - \tilde{H}_1\Big(-(\varphi - \tilde{\varphi}) + h(\psi - \tilde{\psi})\Big). \end{aligned}$$

Application of (2.62), (3.19), (3.21) yields

$$\begin{aligned} \|F_1 - \tilde{F}_1\| \leq\ & hL_\psi\|x - \tilde{x}\| + \\ & + L_{H_1}(\|\bar{d}\|_2 + L_\varphi\|x\| + hM_{f|_{\mathcal{M}}})\|(x,d) - (\tilde{x},\tilde{d})\|_* + \\ & + 2L_{H_1} \max\{r, \|\bar{d}\|_2\}(L_\varphi + hL_\psi)\|x - \tilde{x}\|_2. \end{aligned}$$

We choose $B_{F_1} := L_\psi + L_{H_1}(1 + 3L_\varphi + M_{f|\mathcal{M}} + h2L_\psi)$, which proves (3.23). Furthermore the bound (3.21) yields

$$\begin{aligned} \|F_1\|_2 &= \|h\hat{\psi} - H_1(\bar{d} - \varphi + h\psi)\|_2 \leq \\ &\leq hM_{f|\mathcal{M}} + 2(\|\bar{d}\|_2 + L_\varphi r + hM_{f|\mathcal{M}}), \end{aligned}$$

such that $C_{F_1} := 3M_{f|\mathcal{M}} + 2(1 + L_\varphi)$. Now we consider the F_2-difference:

$$\begin{aligned} F_2 - \tilde{F}_2 &= (H_2 - \tilde{H}_2)(\bar{d} - \varphi \circ F_1 + h\psi) + \\ &\quad + \tilde{H}_2\Big(-(\varphi \circ F_1 - \varphi \circ \tilde{F}_1) + h(\psi - \tilde{\psi})\Big). \end{aligned}$$

Application of (3.20), (3.24), (3.22) and (3.23) yields

$$\begin{aligned} \|F_2 - \tilde{F}_2\| \leq\ & L_{H_2}(\|\bar{d}\|_2 + L_\varphi C_{F_1} \max\{h, \|\bar{d}\|_2, r\} + hM_{f|\mathcal{M}})\|(x,d) - (\tilde{x},\tilde{d})\|_* + \\ & + M_{H_2}(L_\varphi B_{F_1} \max\{h, \|\bar{d}\|_2, r\}\|(x,d) - (\tilde{x},\tilde{d})\|_* + hL_\psi\|x - \tilde{x}\|). \end{aligned}$$

Now $B_{F_2} := L_{H_2}(1 + L_\varphi C_{F_1} + M_{f|\mathcal{M}}) + M_{H_2}(L_\varphi B_{F_1} + L_\psi)$ completes the proof. □

Proposition 3.2.5 ensures that under mild restrictions on $h, \|\bar{d}\|_2, r$ we have contractivity for the function F. Next we consider the first iteration step:

Proposition 3.2.6 *Under the restriction* $\|\bar{d}\|_2 \leq \frac{1}{2\sqrt{k}L_\pi}$ *there holds*

$$\|F_1(0,0)\|_2 \leq C_{F_1(0)} \max\{h, \|\bar{d}\|_2\}, \tag{3.26}$$

$$\|F_2(0,0)\|_2 \leq C_{F_2(0)} \max\{h, \|\bar{d}\|_2\}, \tag{3.27}$$

where $C_{F_1(0)} := 3M_{f|\mathcal{M}} + 2$ *and* $C_{F_2(0)} := M_{H_2}(1 + L_\varphi C_{F_1(0)} + M_{f|\mathcal{M}})$.

Proof: The bounds (3.21), (3.22) yield

$$\begin{aligned} \|F_1(0,0)\|_2 &= \|h\hat{\psi}(0) - H_1(0,0)(\bar{d} + h\psi(0))\|_2 \leq \\ &\leq hM_{f|\mathcal{M}} + 2(\|\bar{d}\|_2 + hM_{f|\mathcal{M}}), \\ \|F_2(0,0)\|_2 &= \|H_2(0,0)(\bar{d} - \varphi \circ F_1(0,0) + h\psi(0))\|_2 \leq \\ &\leq M_{H_2}(\|\bar{d}\|_2 + L_\varphi C_{F_1(0)} \max\{h, \|\bar{d}\|_2\} + hM_{f|\mathcal{M}}), \end{aligned}$$

wher in the second estimate additionally (3.26) is used. □

Now the fixed point theorem A.1.1 can be applied:

Theorem 3.2.1 (Algebraic equation) *Let* $\overline{\mathcal{B}}_r \subset \mathcal{W}$ *and* $C_* := \max\{2C_{F_1(0)}, 2C_{F_2(0)}, 1\}$ *and* $C^* := 2\max\{B_{F_1}, B_{F_2}, \sqrt{k}L_\pi, 2L_\Theta\}$. *If the restriction*

$$C_* \max\{h, \|\bar{d}\|_2\} \leq r \leq \frac{1}{C^*} \tag{3.28}$$

is fulfilled, then the fixed point form (3.15) for the parametrized euler equations (3.6), (3.7) in $\overline{\mathcal{B}}_r$ *posesses a unique solution* (x,d). *Furthermore* $\eta := \eta(x,d)$ *is the unique solution of the algebraic equation (3.2) in* $\eta(\overline{\mathcal{B}}_r)$.

Proof: The restriction (3.28) and proposition 3.2.5 ensures that we have

$$\|F(x,d) - F(\tilde{x},\tilde{d})\|_* \;\leq\; K\|(x,d) - (\tilde{x},\tilde{d})\|_*,$$

where $K := \max\{h, \|\bar{d}\|_2, r\}\max\{B_{F_1}, B_{F_2}\} \leq \frac{1}{2} < 1$, i.e. contractivity. Furthermore we can use (3.28) and proposition 3.2.6 to obtain

$$\|F(0,0)\|_* \;\leq\; \max\{C_{F_1(0)}, C_{F_2(0)}\}\max\{h, \|\bar{d}\|_2\} \;\leq\; \frac{1}{2}r \;\leq (1-K)r.$$

Application of theorem A.1.1 yields the desired result. □

Remarks: Theorem 3.2.1 can be applied in two different ways:

- First let $r := \frac{1}{C^*}$. Then $\overline{\mathcal{B}}_r$ is a ball of $O(1)$ in which the solution (x,d) is unique.
- Second let $r := C_* \max\{h, \|\bar{d}\|_2\}$, which leads to an estimate for the solution (x,d):

$$\|(x,d)\|_* \;\leq\; C_* \max\{h, \|\bar{d}\|_2\} \tag{3.29}$$

3.3 Convergence Estimates for the Implicit Euler Method

Let $u(t) \in \mathcal{M}$, $t \in [0,T]$ be a smooth solution of the differential equation (2.11) with initial value $u(0) = u_0 \in \mathcal{M}$. The corresponding approximations generated by the implicit Euler method are denoted by η_ν, $\nu \in \mathbb{N}_0$, where η_0 is an approximation for u_0. In order to derive error bounds we split the error

$$e_\nu \;:=\; \eta_\nu - u(t_\nu) \;=\; (\eta_\nu - p(\eta_\nu)) + (p(\eta_\nu) - u(t_\nu)) \tag{3.30}$$

up into the stiff component

$$\eta_\nu - p(\eta_\nu) \;=\; \sum_{j=1}^{k} d_j(\eta_\nu)\pi_j(\eta_\nu)$$

and into the smooth component $p(\eta_\nu) - u(t_\nu)$. Furthermore the smooth component is split up into

$$p(\eta_\nu) - u(t_\nu) \;=\; (p(\eta_\nu) - \tilde{u}_{\nu-1}(t_\nu)) + (\tilde{u}_{\nu-1}(t_\nu) - u(t_\nu)), \tag{3.31}$$

where $p(\eta_\nu) - \tilde{u}_{\nu-1}(t_\nu)$ is the local smooth error, i.e. $\tilde{u}_{\nu-1}(t)$ is the solution of the differential equation (2.11) with initial value $\tilde{u}_{\nu-1}(t_{\nu-1}) = p(\eta_{\nu-1})$. For an illustration compare figure 3.2.
In the following sections we first estimate the stiff error coponent, which together with an estimate for the smooth error component leads to convergence of the implicit Euler method.

Notation: We introduce the convention that $\mathcal{C}$'s denote generic constants, i.e. $\mathcal{C}$ is a well defined constant which is moderate, independent of the stiffness parameter ε.

3.3.1 Recursion for the Stiff Error Component

We consider one step $\eta_{\nu-1} \to \eta_\nu$ of the implicit Euler method. We set $\bar{\eta} := \eta_{\nu-1}$ and deal with the corresponding local coordinates (x,d) in a neighbourhood of $\bar{\eta}$. First we set $\tilde{u}(t) := \tilde{u}_{\nu-1}(t_{\nu-1}+t)$ such that $\tilde{u}(0) = \bar{p} = p(\bar{\eta})$ and $\tilde{u}(h) = \tilde{u}_{\nu-1}(t_\nu)$. The solution $\tilde{u}(t)$ lies in $\mathcal{M}$ and for h sufficiently small there is $\tilde{u}(t) \in \phi(\mathcal{U})$ for $t \in [0,h]$. This means that there exists $\tilde{x}(t) \in \mathcal{U}$ such that $\phi(\tilde{x}(t)) = \tilde{u}(t)$ for

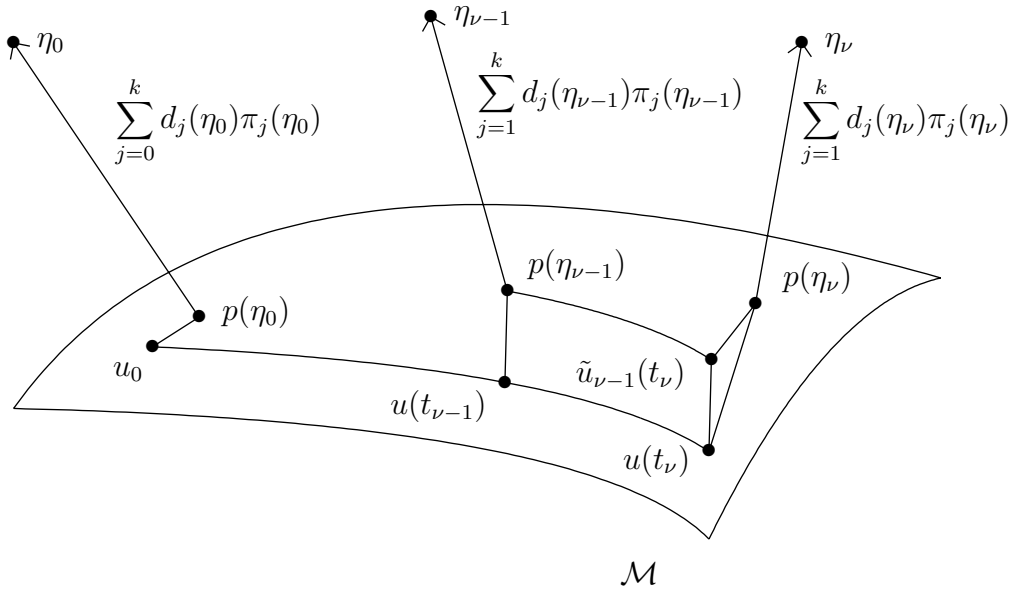


Figure 3.2: The error for the implicit Euler method

$t \in [0,h]$. Then $\tilde{x}(t)$ is a solution of the differential euqation $\tilde{x}'(t) = \hat{\psi}(\tilde{x}(t))$ with $\tilde{x}(0) = 0$. Setting $\bar{d} = d(\bar{\eta})$ we assume that for

$$r \;:=\; C_* \max\{h, \|\bar{d}\|_2\} \tag{3.32}$$

the assumptions of theorem 3.2.1 are fulfilled, such that there exists a locally unique solution $\eta = \eta(x,d)$ of the algebraic equation (3.2), where (x,d) are the local coordinates of the solution.[4)] Now the new approximation $\eta_\nu := \eta$ is well defined. Furthermore we can use the estimate (3.29). In the following considerations we make use of the short notation of the previous sections. The formulation (3.3) of the algebraic equation (3.2) leads us to the identity

$$\sum_{j=1}^{k} \delta_j \pi_j \;=\; \sum_{j=1}^{k} \bar{d}_j \pi_j - \chi, \tag{3.33}$$

where we define the auxiliary quantities

$$\delta \;:=\; (I - h\Lambda)d, \quad \chi \;:=\; p - \bar{p} - hf(p).$$

(δ_j, $j = 1,\dots,k$ denote the components of δ.) Application of the scalarproduct with π_i, $i = 1,\dots,k$ to (3.33) yields

$$\delta \;=\; \Theta^\top \bar{d} - (\langle \chi, \pi_i \rangle)_{i=1}^k. \tag{3.34}$$

Now we introduce the short notation $V := V(\eta)$ such that there holds

$$\|d\|_V \;\;\leq\;\; \|(I - h\Lambda)^{-1}\|_V \|\delta\|_V, \tag{3.35}$$

$$\|\delta\|_V \;\;\leq\;\; \|\Theta^\top\|_V \|\bar{d}\|_V + M_V^{\frac{1}{2}} \|\chi\|_2. \tag{3.36}$$

The bound (3.36) is a consequence of (3.34) and

$$\|(\langle \chi, \pi_i \rangle)_{i=1}^k\|_V \;\leq\; M_V^{\frac{1}{2}} \|(\langle \chi, \pi_i \rangle)_{i=1}^k\|_2 \;=\; M_V^{\frac{1}{2}} \|\chi\|_2.$$

[4)] Later we will show that under a suiteable induction asssumption on $\bar{d}$ together with a moderate step size restriction, these assumptions are fulfilled in each step.

Our goal is to derive a bound for $\|d\|_V$ under assumptions on $\|\bar{d}\|_{\bar{V}}$, where $\bar{V} := V(\bar{\eta})$, which will lead us to an inductive estimate of the stiff error component. Therefore we first note that by (2.68) from proposition 2.4.5 and (3.32) there holds

$$\begin{aligned}\|\bar{d}\|_V &\leq \|\bar{d}\|_{\bar{V}} + L_{V^{\frac{1}{2}}} L_\eta \|(x,d) - (0,\bar{d})\|_* \|\bar{d}\|_2 \leq \\ &\leq \|\bar{d}\|_{\bar{V}} + L_{V^{\frac{1}{2}}} L_\eta (r + \|\bar{d}\|_2)\|\bar{d}\|_2 \leq \\ &\leq \|\bar{d}\|_{\bar{V}} + \mathcal{C}\max\{h^2, \|\bar{d}\|_2^2\}, \end{aligned} \tag{3.37}$$

where in this case the generic constant $\mathcal{C}$ is $\mathcal{C} = L_{V^{\frac{1}{2}}} L_\eta 2C_*$. Now we estimate the remaining terms of (3.35), (3.36).

Proposition 3.3.1 *Under the assumptions of theorem 3.2.1 there holds*

$$\|(I - h\Lambda)^{-1}\|_V \leq \frac{1}{1 + \frac{h}{\varepsilon M_1}}. \tag{3.38}$$

Proof: The proof is similar to the proof of proposition 3.2.1 and therefore omitted. □

Proposition 3.3.2 *Under the assumptions of theorem 3.2.1 there exists a moderate positive constant $\mathcal{C}$ such that there holds*

$$\|\Theta^\top\|_V \leq 1 + \mathcal{C}\max\{h, \|\bar{d}\|_2\}. \tag{3.39}$$

The constant $\mathcal{C}$ can be chosen as $\mathcal{C} = 2M_0 L_\Theta C_*$.

Proof: Application of (2.67) from proposition 2.4.4 and (3.32) yields

$$\begin{aligned}\|\Theta^\top\|_V &\leq \|I\|_V + \|I - \Theta^\top\|_V \leq 1 + M_0\|I - \Theta\|_2 \leq \\ &\leq 1 + 2M_0 L_\Theta \max\{r, \|\bar{d}\|_2\} \leq \\ &\leq 1 + 2M_0 L_\Theta C_* \max\{h, \|\bar{d}\|_2\}. \end{aligned}$$

□

In order to estimate $\|\chi\|_2$ we introduce the truncation error

$$\tau := \tilde{u}(h) - \tilde{u}(0) - hf(\tilde{u}(h)) \tag{3.40}$$

corresponding to the solution $\tilde{u}(t)$ of the differential equation (2.11) with initial value $\tilde{u}(0) = \bar{p}$. Because of the smoothness assumption on $f \circ \phi$ there exists a moderate positive constant $\mathcal{C}$ such that

$$\|\tau\|_2 = \|\tilde{u}(h) + \tilde{u}'(h)\cdot(-h) - \tilde{u}(h-h)\|_2 \leq \mathcal{C}h^2. \tag{3.41}$$

With the help of the coordinate function $\tilde{x}(t)$ for $\tilde{u}(t)$, the bound for the truncation error (3.41) and (2.13), (2.53) there holds

$$\begin{aligned}\|\chi\|_2 &= \|p - \bar{p} - hf(p)\|_2 = \\ &= \|\Big(p - \tilde{u}(h)\Big) - h\Big(f(p) - f(\tilde{u}(h))\Big) + \tau\|_2 \leq \\ &\leq (1 + hL_{f|_{\mathcal{M}}})\|p - \tilde{u}(h)\|_2 + \mathcal{C}h^2 \leq \\ &\leq (1 + hL_{f|_{\mathcal{M}}})L_\phi\|x - \tilde{x}(h)\|_2 + \mathcal{C}h^2. \end{aligned} \tag{3.42}$$

Now we introduce the auxiliary quantities

$$\bar{\tau} \;:=\; \tilde{x}(h) - \tilde{x}(0) - h\hat{\psi}(\tilde{x}(h)), \quad \bar{\chi} \;:=\; \varphi - h\psi \tag{3.43}$$

and note that there holds $\|\bar{\tau}\|_2 \leq \|\tau\|_2$ and $\|\bar{\chi}\|_2 \leq \|\chi\|_2$. The next step is to bound

$$\begin{aligned} \|x - \tilde{x}(h)\|_2 &= \|h\hat{\psi} - H_1(\bar{d} - \varphi + h\psi) - \tilde{x}(h)\|_2 \;\leq \\ &\leq \|\bar{\tau}\|_2 + h\|\hat{\psi} - \hat{\psi}(\tilde{x}(h))\|_2 + \|H_1\|_2(\|\bar{d}\|_2 + \|\bar{\chi}\|_2) \;\leq \\ &\leq \mathcal{C}h^2 + hL_\psi\|x - \tilde{x}(h)\|_2 + \|H_1\|_2(\|\bar{d}\|_2 + \|\chi\|_2). \end{aligned} \tag{3.44}$$

Under the stepsize restriction

$$h \;\leq\; \frac{1}{2L_\psi} \tag{3.45}$$

and with the help of (3.21), (3.32) there follows

$$\begin{aligned} \|x - \tilde{x}(h)\|_2 &\leq 2\Big(\|H_1\|_2(\|\bar{d}\|_2 + \|\chi\|_2) + \mathcal{C}h^2\Big) \leq \\ &\leq \mathcal{C}\max\{h, \|\bar{d}\|_2\}\|\chi\|_2 + \mathcal{C}\max\{h^2, \|\bar{d}\|_2^2\}, \end{aligned} \tag{3.46}$$

where the constants $\mathcal{C}$ come from bounds for the expressions in the first line of (3.46). Insertion of (3.46) into (3.42) yields

$$\|\chi\|_2 \;\leq\; \mathcal{C}\max\{h, \|\bar{d}\|_2\}\|\chi\|_2 + \mathcal{C}\max\{h^2, \|\bar{d}\|_2^2\}, \tag{3.47}$$

with suiteable generic constants $\mathcal{C}$. Under the restrictions

$$h, \|\bar{d}\|_2 \;\leq\; \frac{1}{2\mathcal{C}}, \tag{3.48}$$

where $\mathcal{C}$ comes from the first term of (3.47) there follows

$$\|\chi\|_2 \;\leq\; 2\mathcal{C}\max\{h^2, \|\bar{d}\|_2^2\}, \tag{3.49}$$

where $\mathcal{C}$ comes from the second term of (3.47). A combination of the results (3.37), (3.38), (3.39) and (3.49) leads to

$$\begin{aligned} \|d\|_V &\leq \frac{1}{1 + \frac{h}{\varepsilon M_1}}\Big((1 + \mathcal{C}\max\{h, \|\bar{d}\|_2\})\|\bar{d}\|_V + 2M_V^{\frac{1}{2}}\mathcal{C}\max\{h^2, \|\bar{d}\|_2^2\}\Big) \;\leq \\ &\leq \frac{1}{1 + \frac{h}{\varepsilon M_1}}\Big(\|\bar{d}\|_{\bar{V}} + \mathcal{C}\max\{h^2, \|\bar{d}\|_{\bar{V}}^2\}\Big), \end{aligned} \tag{3.50}$$

where we used $\|\bar{d}\|_V \leq M_V^{\frac{1}{2}}\|\bar{d}\|_2$ and $\|\bar{d}\|_2 \leq M_{V-1}^{\frac{1}{2}}\|\bar{d}\|_{\bar{V}}$. Now we make the following induction assumption:

Induction assumption: For $K_0 \in [K_{\min}, \frac{1}{\varepsilon}]$ where $K_{\min} := M_1\mathcal{C}$ with $\mathcal{C}$ from (3.50) there holds

$$\|\bar{d}\|_{\bar{V}} \;\leq\; K_0\varepsilon h \tag{3.51}$$

for $h \leq h_{\max}$, where

$$h_{\max} \;:=\; \min\left\{\frac{1}{C_*C^*M_{V-1}^{\frac{1}{2}}}, \frac{1}{2L_\phi}, \frac{1}{2\mathcal{C}M_{V-1}^{\frac{1}{2}}}\right\} \tag{3.52}$$

with $\mathcal{C}$ from (3.48).

Since $\|\bar{d}\|_{\bar{V}} \leq K_0\varepsilon h \leq h$ and $\|\bar{d}\|_2 \leq M_{V^{-1}}^{\frac{1}{2}}\|\bar{d}\|_{\bar{V}}$ the assumptions (3.51), (3.52) imply (3.28), (3.45) and (3.48). Now (3.50) can be used to derive the induction conclusion:

Induction conclusion:

$$\begin{aligned}\|d\|_V &\leq \frac{1}{1+\frac{h}{\varepsilon M_1}}(K_0\varepsilon h + \mathcal{C}h^2) \leq \\ &\leq K_0\varepsilon h \frac{1+\frac{h}{M_1\varepsilon}\frac{M_1\mathcal{C}}{K_0}}{1+\frac{h}{\varepsilon M_1}} \leq K_0\varepsilon h \end{aligned} \tag{3.53}$$

for $h \leq h_{\max}$.

We can formulate the following proposition:

Proposition 3.3.3 (Stiff error component) *Let $K_0 \in [K_{\min}, \frac{1}{\varepsilon}]$, $K_{\min} \leq \frac{1}{\varepsilon}$ and*

$$\|d(\eta_0)\|_{V(\eta_0)} \leq K_0\varepsilon h \tag{3.54}$$

for $h \leq h_{\max}$. Then there holds

$$\|d(\eta_\nu)\|_{V(\eta_\nu)} \leq K_0\varepsilon h \tag{3.55}$$

for all $\nu \in \mathbb{N}_0$ with $t_\nu \in [0,T]$. The constants $K_{\min}, h_{\max}$ have to be chosen like in the considerations above.

Remark: Proposition 3.3.3 can be formulated in the 2-norm: Let

$$\|d(\eta_0)\|_2 \leq B_0\varepsilon h$$

for all $h \leq h_{\max}$, such that $K_0 := B_0 M_{V^{-1}}^{\frac{1}{2}} \in [K_{\min}, \frac{1}{\varepsilon}]$. Then there holds

$$\|d(\eta_\nu)\|_2 \leq M_0 B_0 \varepsilon h$$

for all $\nu \in \mathbb{N}_0$ with $t_\nu \in [0,T]$.

3.3.2 Estimates for the Smooth Error Component

We apply the bounds for the stiff error component to obtain an estimate for the smooth error component. The bounds (3.49) and (3.46) yield

$$\begin{aligned}\|\chi\|_2 &\leq \mathcal{C}h^2, \\ \|x - \tilde{x}(h)\|_2 &\leq \mathcal{C}h^2. \end{aligned} \tag{3.56}$$

The bound (3.56) can be applied to obtain

$$\|p - \tilde{u}(h)\|_2 \leq L_\phi \|x - \tilde{x}(h)\|_2 \leq \mathcal{C}h^2. \tag{3.57}$$

Now we use the index notation, i.e. in the step from $\bar{\eta} := \eta_{\nu-1}$ to $\eta_\nu := \eta$ the notation $p(\eta_\nu)$ for p and $\tilde{u}_{\nu-1}$ for $\tilde{u}$ is used. We derive the estimates

$$\begin{aligned}\|p(\eta_\nu) - u(t_\nu)\|_2 &\leq \|p(\eta_\nu) - \tilde{u}_{\nu-1}(t_\nu)\|_2 + \|\tilde{u}_{\nu-1}(t_{\nu-1}+h) - u(t_{\nu-1}+h)\|_2 \leq \\ &\leq \mathcal{C}h^2 + e^{L_{f|_{\mathcal{M}}}h}\|\tilde{u}_{\nu-1}(t_{\nu-1}) - u(t_{\nu-1})\|_2 \leq \\ &\leq e^{L_{f|_{\mathcal{M}}}t_\nu}\|p(\eta_0) - u_0\|_2 + \frac{e^{L_{f|_{\mathcal{M}}}t_\nu} - 1}{e^{L_{f|_{\mathcal{M}}}h} - 1}\mathcal{C}h^2 \leq \\ &\leq e^{L_{f|_{\mathcal{M}}}t_\nu}\|p(\eta_0) - u_0\|_2 + (e^{L_{f|_{\mathcal{M}}}t_\nu} - 1)\mathcal{C}h. \end{aligned} \tag{3.58}$$

Under the assumption that

$$\|p(\eta_0) - u_0\|_2 \leq C_0 h \tag{3.59}$$

for $h \leq h_{\max}$, where $C_0 \geq 0$ is a moderate constant, there follows

$$\begin{aligned} \|p(\eta_\nu) - u(t_\nu)\|_2 &\leq e^{L_{f|\mathcal{M}}{}^T} C_0 h + (e^{L_{f|\mathcal{M}}{}^T} - 1)\mathcal{C}h = \\ &= \mathcal{C}h \end{aligned} \tag{3.60}$$

for all $t_\nu \in [0, T]$.

3.3.3 Error Bounds for the Implicit Euler Method

Now the whole error as a combination of the stiff and the smooth error component can be estimated. From (3.55), (3.60) there follows

$$\begin{aligned} \|\eta_\nu - u(t_\nu)\|_2 &\leq \underbrace{\|\eta_\nu - p(\eta_\nu)\|_2}_{= \|d(\eta_\nu)\|_2} + \|p(\eta_\nu) - u(t_\nu)\|_2 \leq \\ &\leq M_V^{\frac{1}{2}} K_0 \varepsilon h + \mathcal{C}h = \mathcal{C}h \end{aligned} \tag{3.61}$$

for all $t_\nu \in [0, T]$. This leads to the following convergence theorem:

Theorem 3.3.1 (Convergence of the implicit Euler method) *Let $B_0, C_0 \geq 0$, such that $K_0 := B_0 M_{V^{-1}}^{\frac{1}{2}} \in [K_{\min}, \frac{1}{\varepsilon}]$ and*

$$\begin{aligned} \|\eta_0 - p(\eta_0)\|_2 &\leq B_0 \varepsilon h, \\ \|p(\eta_0) - u_0\|_2 &\leq C_0 h, \end{aligned}$$

for all $h \leq h_{\max}$. Then the implicit Euler method satisfies

$$\|\eta_\nu - p(\eta_\nu)\|_2 \leq M_0 B_0 \varepsilon h, \tag{3.62}$$

$$\|p(\eta_\nu) - u(t_\nu)\|_2 \leq \mathcal{C}h, \tag{3.63}$$

$$\|\eta_\nu - u(t_\nu)\|_2 \leq \mathcal{C}h, \tag{3.64}$$

for all $t_\nu \in [0, T], \nu \in \mathbb{N}$. The appearing constants have to be chosen like in the considerations above.

Remarks: concerning $K_0 \in [K_{\min}, \frac{1}{\varepsilon}]$:

- For $B_0 = K_{\min} M_{V^{-1}}^{-\frac{1}{2}}$ the stiff error component of the approximation for the initial value has to satisfy a rather restrictive assumption. The stiff components of the further approximations stay at the same level as the stiff component of the initial approximation, i.e. at the $O(\varepsilon h)$-level.

- For $B_0 = \frac{1}{\varepsilon} M_{V^{-1}}^{-\frac{1}{2}}$ there is a rather mild restriction for the stiff error component of the initial approximation, which results in a weaker bound for the stiff error components during the integration, i.e. the stiff error component is of $O(h)$. For this case the following considerations lead to an improved error result. The assumption

$$\|d(\eta_{\nu-1})\|_{V(\eta_{\nu-1})} \leq K_{\nu-1} \varepsilon h,$$

where $K_{\min} < K_{\nu-1} \leq \frac{1}{\varepsilon}$ leads to

$$\|d(\eta_\nu)\|_{V(\eta_\nu)} \;\leq\; K_{\nu-1}\varepsilon h \frac{1+\frac{h}{\varepsilon M_1}\frac{K_{\min}}{K_{\nu-1}}}{1+\frac{h}{\varepsilon M_1}} \;=\; K_\nu \varepsilon h,$$

where

$$K_\nu := \frac{1+\frac{h}{\varepsilon M_1}\frac{K_{\min}}{K_{\nu-1}}}{1+\frac{h}{\varepsilon M_1}} K_{\nu-1}.$$

There follows $\lim_{\nu\to\infty} K_\nu = K_{\min}$. During the integration this damping in the stiff error component leads to a stiff error component of $O(\varepsilon h)$.

Chapter 4

The Methods Radau Ia, IIa and Gauss

Let $u(t) \in \mathcal{M}$, $t \in [0,T]$ be a smooth solution of the differential equation (2.11) with initial value $u(0) = u_0 \in \mathcal{M}$. The solution $u(t)$ is approximated by an s-stage implicit Runge Kutta method. We restrict our considerations to the methods Radau Ia, IIa and Gauss.
A discretization of the intervall $[0,T]$ is given by $t_\nu := \nu h, \nu \in \mathbb{N}_0$, where h is the constant step size. Starting with the initial approximation $\eta_0 \in \mathcal{G}$ for u_0 further approximations η_ν for $u(t_\nu)$ are computed via the implicit Runge-Kutta method.
Let $s \in \mathbb{N}$ be the number of stages. The s-stage implicit Runge-Kutta methods Radau Ia, IIa and Gauss are defined via the Butcher-array

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

For more details compare section B.1 in the appendix. One step from $\eta_{\nu-1}$ to η_ν is described via the algebraic equations

$$Y_i \;=\; \eta_{\nu-1} + h \sum_{j=1}^{s} a_{ij} f(Y_j), \quad i = 1, \ldots, s \tag{4.1}$$

for the unknowns $Y_i \in \mathcal{G}$. The stages Y_i are approximations for $u(t_{\nu-1} + c_i h)$, where $i = 1, \ldots, s$. Insertion of Y_i into

$$\eta_\nu \;=\; \eta_{\nu-1} + h \sum_{i=1}^{s} b_i f(Y_i) \tag{4.2}$$

yields η_ν (compare figure 4.1).
First we have to ensure that a step from $\eta_{\nu-1}$ to η_ν is well defined, i.e. it has to be shown that the algebraic equations

$$Y_i \;=\; \bar{\eta} + h \sum_{j=1}^{s} a_{ij} f(Y_j), \quad i = 1, \ldots, s \tag{4.3}$$

for the unknowns $Y_i \in \mathcal{G}$, $i = 1, \ldots, s$ posess a locally unique solution, where $\bar{\eta} := \eta_{\nu-1} \in \mathcal{G}$ is given. Therefore the algebraic equations (4.3) are parametrized via $\eta(x,d)$ in a neighbourhood of $\bar{\eta}$. In this way we obtain equations in the unknown parameters (X_i, D_i) for $Y_i = \eta(X_i, D_i)$, $i = 1, \ldots, s$. A transformation into a suiteable fixed point form for (X_i, D_i) and application of the fixed point theorem A.1.1 leads to existence and local uniqueness of a solution (X_i, D_i), $i = 1, \ldots, s$ of the fixed

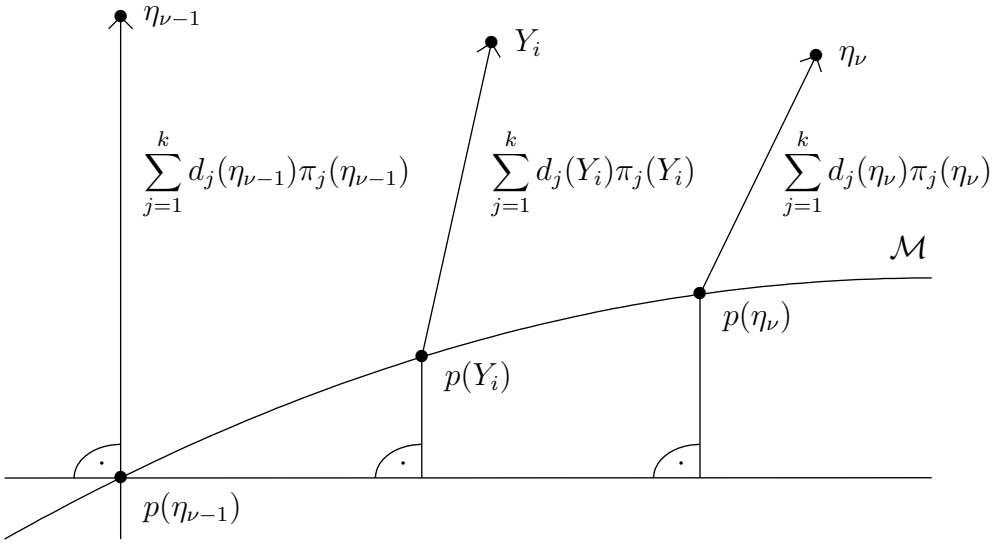


Figure 4.1: One step of the implicit Runge-Kutta method

point equation. Then $Y_i = \eta(X_i, D_i)$, $i = 1, \ldots, s$ is the locally unique solution of the algebraic equations (4.3).
Second we estimate the stiff and the smooth error components. This leads to a convergence result for the implicit Runge-Kutta methods Radau Ia, IIa and Gauss. In the case of ε very small compared to h nearly superconvergence concerning the smooth error components for the methods Radua IIa and Gauss is obtained.

4.1 Parametrization of the Algebraic Equations

The goal of this section is to formulate the algebraic equations (4.3) in an appropriate fixed point form, such that theorem A.1.1 can be applied to obtain existence and local uniqueness of a solution of the algebraic equations.
Formally the assumptions of section 2.2 lead to a separation of the stiff and smooth components in the algebraic equations. In order to simplify the notation we make use of the short notation

$$\bar{p} := p(\bar{\eta}), \quad \bar{d} := d(\bar{\eta}), \quad \bar{\pi}_j := \pi_j(\bar{\eta})$$

$$P_i := p(Y_i), \quad D_i := d(Y_i), \quad \Pi_{ij} := \pi_j(Y_i)$$

and $\Lambda_i := \Lambda(Y_i)$, where $i = 1, \ldots, s$ and $j = 1, \ldots, k$ and $\bar{\eta}, Y_i \in \mathcal{G}$. The identities (compare section 2.2) [1)]

$$\bar{\eta} = \bar{p} + \sum_{j=1}^{k} \bar{d}_j \bar{\pi}_j, \quad Y_i = P_i + \sum_{j=1}^{k} D_{ij}\Pi_{ij}, \quad f(Y_i) = f(P_i) + \sum_{j=1}^{k} [\Lambda_i D_i]_j \Pi_{ij},$$

for $\bar{\eta}, Y_i \in \mathcal{G}$, where $i = 1, \ldots, s$, lead to the following formulation of the algebraic equations (4.3)

$$P_i + \sum_{j=1}^{k} D_{ij}\Pi_{ij} = \tag{4.4}$$

[1)] The brackets $[\cdot]_j$ denote the j-th component of a vector.

$$= \bar{p} + \sum_{j=1}^{k} \bar{d}_j \bar{\pi}_j + h \sum_{j=1}^{s} a_{ij} \Big(f(P_j) + \sum_{l=1}^{k} [\Lambda_j D_j]_l \Pi_{jl} \Big), \quad i = 1, \ldots, s.$$

The formulations (4.4) and (4.3) are equivalent in the sense that the solutions coincide. [2)]
Now we use the local parametrization $\phi(x)$ of $\mathcal{M}$ and the local coordinate transformation $\eta(x,d)$ corresponding to $\bar{\eta}$. Since $\eta(x,d)$ is a bijection from $\mathcal{W}$ to $\mathcal{V}$ we can replace the quantities $P_i = p(Y_i)$, $D_i = d(Y_i)$, $\Pi_{ij} = \pi_j(Y_i)$ and $\Lambda_i = \Lambda(Y_i)$ for $Y_i \in \mathcal{V}$ by $P_i = p(\eta(X_i, D_i))$, $D_i = d(\eta(X_i, D_i))$, $\Pi_{ij} = \pi_j(\eta(X_i, D_i))$ and $\Lambda_i = \Lambda(\eta(X_i, D_i))$, where $(X_i, D_i) \in \mathcal{W}$. Because of $p(\eta(X_i, D_i)) = \phi(X_i)$ the short notation $\Phi_i := (\Phi_{i1}, \ldots, \Phi_{ik})^\top := \varphi(X_i)$ allows us to replace P_i on the left hand side of (4.4) by

$$P_i \;=\; \phi(X_i) \;=\; \bar{p} + \sum_{j=1}^{n-k} X_{ij} \hat{\pi}_j + \sum_{j=1}^{k} \Phi_{ij} \bar{\pi}_j.$$

This eliminates $\bar{p}$ in (4.4). Furthermore we replace P_i in $f(P_i)$ by $\phi(X_i)$ and obtain the equations

$$\sum_{j=1}^{n-k} X_{ij} \hat{\pi}_j + \sum_{j=1}^{k} \Phi_{ij} \bar{\pi}_j + \sum_{j=1}^{k} D_{ij} \Pi_{ij} \;= \tag{4.5}$$

$$= \sum_{j=1}^{k} \bar{d}_j \bar{\pi}_j + h \sum_{j=1}^{s} a_{ij} \Big(f(\phi(X_j)) + \sum_{l=1}^{k} [\Lambda_j D_j]_l \Pi_{jl} \Big), \quad i = 1, \ldots, s.$$

For the coordinate vectors of $f(\phi(X_i))$ we use the short notation

$$\hat{\Psi}_i := (\hat{\Psi}_{i1}, \ldots, \hat{\Psi}_{i\,n-k})^\top := \hat{\psi}(X_i), \quad \Psi_i := (\Psi_{i1}, \ldots, \Psi_{ik})^\top := \psi(X_i).$$

This allows us to replace $f(\phi(X_i))$ on the right hand side of (4.5) by

$$f(\phi(X_i)) = \sum_{j=1}^{n-k} \hat{\Psi}_{ij} \hat{\pi}_j + \sum_{j=1}^{k} \Psi_{ij} \bar{\pi}_j.$$

The resulting equations for $(X_i, D_i) \in \mathcal{W}$ can be formulated in the ONS B:

- Application of the scalarproduct $\langle \cdot, \hat{\pi}_m \rangle$, $m = 1, \ldots, n-k$ to (4.5) yields

$$X_{im} + \sum_{j=1}^{k} D_{ij} \langle \Pi_{ij}, \hat{\pi}_m \rangle \;= \tag{4.6}$$

$$= h \sum_{j=1}^{s} a_{ij} \Big(\underbrace{\langle f(\phi(X_j)), \hat{\pi}_m \rangle}_{= \hat{\Psi}_{jm}} + \sum_{l=1}^{k} [\Lambda_j D_j]_l \langle \Pi_{jl}, \hat{\pi}_m \rangle \Big),$$

  where $i = 1, \ldots, s$, $m = 1, \ldots, n-k$.

- Application of the scalarproduct $\langle \cdot, \bar{\pi}_m \rangle$, $m = 1, \ldots, k$ to (4.5) yields

$$\Phi_{im} + \sum_{j=1}^{k} D_{ij} \langle \Pi_{ij}, \bar{\pi}_m \rangle \;= \tag{4.7}$$

$$= \bar{d}_m + h \sum_{j=1}^{s} a_{ij} \Big(\underbrace{\langle f(\phi(X_j)), \bar{\pi}_m \rangle}_{= \Psi_{jm}} + \sum_{l=1}^{k} [\Lambda_j D_j]_l \langle \Pi_{jl}, \bar{\pi}_m \rangle \Big),$$

  where $i = 1, \ldots, s$, $m = 1, \ldots, k$.

[2)] Note that the unknowns of (4.4) are $Y_i, i = 1, \ldots, s$ which are hidden due to aur short notation.

Now let for short $\hat{\Theta}_i := \hat{\Theta}(X_i, D_i)$, $\Theta_i := \Theta(X_i, D_i)$ such that $\hat{\Theta}_i = (\langle \Pi_{ij}, \hat{\pi}_m \rangle)_{mj}$ and $\Theta_i = (\langle \Pi_{ij}, \bar{\pi}_m \rangle)_{mj}$ in our short notation. Then the matrix-vector notation of the parametrized algebraic equations (4.6), (4.7) reads as follows

$$X_i + \hat{\Theta}_i D_i = h \sum_{j=1}^{s} a_{ij} (\hat{\Psi}_j + \hat{\Theta}_j \Lambda_j D_j), \tag{4.8}$$

$$\Phi_i + \Theta_i D_i = \bar{d} + h \sum_{j=1}^{s} a_{ij} (\Psi_j + \Theta_j \Lambda_j D_j), \tag{4.9}$$

where $i = 1, \ldots, s$. In this formulation the unknowns are the coordinates $(X_i, D_i) \in \mathcal{W}$.

Proposition 4.1.1 *Let (X_i, D_i), $i = 1, \ldots, s$ be the unique solution of (4.8), (4.9) in $\mathcal{B}^s$, where $\mathcal{B}$ is a neighbourhood of $0 \in I\!R^{n-k} \times I\!R^k$, then $Y_i = \eta(X_i, D_i)$, $i = 1, \ldots, s$ is the unique solution of (4.3) in $\eta(\mathcal{B})^s$.*

Proof: The local coordinate transformation is a bijection, which implies the proposition. □

The next step is to find an appropriate fixed point form for the algebraic equations (4.8), (4.9). Therefore we consider the equations (4.8), (4.9) for $(X_i, D_i) \in \overline{\mathcal{B}}_r \subset \mathcal{W}$, where

$$\mathcal{B}_r := \{(x, d) \colon \|(x, d)\|_* < r\}$$

for $r > 0$. The parameter r will be fixed in section 4.3. Here we derive some restrictions on r such that the fixed point theorem can be applied. Now the following notation is introduced

$$\begin{aligned} X &:= (X_1, \ldots, X_s)^\top, & D &:= (D_1, \ldots, D_s)^\top \\ \hat{\Psi} &:= (\hat{\Psi}_1, \ldots, \hat{\Psi}_s)^\top, & \Psi &:= (\Psi_1, \ldots, \Psi_s)^\top, \\ \hat{\Omega} &:= \operatorname{diag}(\hat{\Theta}_1, \ldots, \hat{\Theta}_s), & \Phi &:= (\Phi_1, \ldots, \Phi_s)^\top \\ \Omega &:= \operatorname{diag}(\Theta_1, \ldots, \Theta_s), & \Gamma &:= \operatorname{diag}(\Lambda_1, \ldots, \Lambda_s) \end{aligned}$$

as well as $e := (1, \ldots, 1)^\top \in \mathbb{R}^s$. Sometimes we use $\hat{\Psi}(X)$, $\Psi(X)$, $\Phi(X)$, $\hat{\Omega}(X, D)$, $\Omega(X, D)$ $\Gamma(X, D)$ instead of $\hat{\Psi}$, Ψ, Φ, $\hat{\Omega}$, Ω, Γ to indicate the dependence on (X, D). In this notation the equations (4.8), (4.9) are of the form [3)]

$$X + \hat{\Omega} D = h(A \otimes I_{n-k})(\hat{\Psi} + \hat{\Omega} \Gamma D), \tag{4.10}$$

$$\Phi + \Omega D = e \otimes \bar{d} + h(A \otimes I_k)(\Psi + \Omega \Gamma D). \tag{4.11}$$

For further reformulations of the algebraic equations we need some propositions. First the following norms on $\mathbb{R}^{(n-k)\times s}$, $\mathbb{R}^{k \times s}$ and $\mathbb{R}^{(n-k)\times s} \times \mathbb{R}^{k \times s}$ are introduced

$$\begin{aligned} \|X\|_\circ &:= \|(\|X_1\|_2, \ldots, \|X_s\|_2)^\top\|_\infty, \\ \|D\|_\circ &:= \|(\|D_1\|_2, \ldots, \|D_s\|_2)^\top\|_\infty, \\ \|(X, D)\|_\diamond &:= \|(\|X\|_\circ, \|D\|_\circ)^\top\|_\infty. \end{aligned}$$

[3)] Compare section A.3 for more information about the direct product $\otimes$.

Proposition 4.1.2 *Let $\overline{\mathcal{B}}_r \subset \mathcal{W}$, where r satisfies the restriction*

$$r \;\leq\; \min\left\{\frac{1}{2\sqrt{k}L_\pi}, \frac{1}{4L_\Theta}\right\}. \tag{4.12}$$

In addition let $\|\bar{d}\|_2$ be restrited by

$$\|\bar{d}\|_2 \;\leq\; \frac{1}{4L_\Theta}\,, \tag{4.13}$$

then the matrix $\Omega(X,D)$ is regular with the bound

$$\|\Omega(X,D)^{-1}\|_\circ \;\leq\; 2, \tag{4.14}$$

for all (X,D) with $(X_i,D_i) \in \overline{\mathcal{B}}_r$ for $i = 1,\ldots,s$.

Proof: Application of proposition 3.1.2 yields the regularity of Θ_i with $\|\Theta_i^{-1}\|_2 \leq 2$ for $i = 1,\ldots,s$. Now the proposition follows from the definition of $\Omega(X,D)$. □

The following proposition requires the BSI-stability with the BSI-stability function $\Phi_I(\cdot)$. Compare section B.2.1 in the appendix.

Proposition 4.1.3 *Consider the methods Radau Ia, IIa and Gauss and let $\Upsilon(X,D) := I - h(A \otimes I_k)\Gamma(X,D)$. In addition let $\overline{\mathcal{B}}_r \subset \mathcal{W}$, where r is restricted by*

$$r \;\leq\; \min\left\{\frac{1}{2\sqrt{k}L_\pi}, \frac{1}{4kL_GL_\eta M_V}\right\} \tag{4.15}$$

and let $\|\bar{d}\|_2$ satisfy the restriction

$$\|\bar{d}\|_2 \;\leq\; \frac{1}{4kL_GL_\eta M_V}, \tag{4.16}$$

then the matrix $\Upsilon(X,D)$ is regular with the bound

$$\|\Upsilon(X,D)^{-1}\|_\circ \;\leq\; \sqrt{s}\mathcal{K}_0\Phi_I(-\frac{h}{\varepsilon\mathcal{K}_1}) \tag{4.17}$$

for all (X,D) with $(X_i,D_i) \in \overline{\mathcal{B}}_r$ for $i = 1,\ldots,s$, where $\mathcal{K}_0 := (M_V M_{V^{-1}})^{\frac{1}{2}}$, $\mathcal{K}_1 := 4M_V$.

Proof: Let us define the function $f_\bullet(t,d) := \Lambda(t)d$, $d \in \mathbb{R}^k$, where $\Lambda(t)$ is chosen such that there holds $\Lambda(t_{\nu-1} + c_i h) = \Lambda_i$ for $i = 1,\ldots,s$. Now we derive a one sided Lipschitz constant for the function $f_\bullet(t,d)$ which holds for $t \in \{t_{\nu-1} + c_i h \colon i = 1,\ldots,s\}$. Let us identify the inner poduct $\langle\cdot,\cdot\rangle_\bullet$ in section B.2.1 with the inner product $\langle\cdot,\cdot\rangle_{\bar{V}} := \langle\bar{V}\cdot,\cdot\rangle$, where we use the notation $\bar{V} := V(\bar{\eta})$. Now we estimate the inner products $\langle\Lambda_i d, d\rangle_{\bar{V}}$ for $i = 1,\ldots,s$. With the help of (2.23), (2.25), (2.57), (4.15), (4.16) and $\bar{\Lambda} := \Lambda(\bar{\eta})$ we derive

$$\begin{aligned}
\langle\Lambda_i d, d\rangle_{\bar{V}} \;&=\; \frac{1}{2}\langle(\bar{\Lambda}^\top\bar{V} + \bar{V}\bar{\Lambda})d, d\rangle + \langle(\Lambda_i - \bar{\Lambda})d, d\rangle_{\bar{V}} \;\leq \\
&\leq\; \Big(-\frac{1}{2\varepsilon} + \|\Lambda_i - \bar{\Lambda}\|_2\|\bar{V}\|_2\Big)\|d\|_2^2 \;\leq \\
&\leq\; -\frac{1}{\varepsilon}\Big(\frac{1}{2} - 2kL_GL_\eta M_V \max\{r, \|\bar{d}\|_2\}\Big)\|d\|_2^2, \leq \\
&\leq\; -\frac{1}{4\varepsilon}\|d\|_2^2 \;\leq\; -\frac{1}{\varepsilon 4M_V}\|d\|_{\bar{V}}^2.
\end{aligned}$$

Let us set $\mathcal{K}_1 := 4M_V$ such that $f_\bullet$ satisfies a one sided Lipschitz condition for $t \in \{t_{\nu-1} + c_i h \colon i = 1, \ldots, s\}$ with the one sided Lipschitz constant $m = -\frac{1}{\varepsilon \mathcal{K}_1}$ corresponding to the inner product $\langle \cdot, \cdot \rangle_{\bar{V}}$. Now we can apply the BSI-stability which holds for the methods Radau Ia, IIa and Gauss (compare appendix section B.2.1). Therefore we compare the identities

$$\begin{aligned} \Delta &= 0 + h(A \otimes I_k)\Gamma(X, D)\Delta + \Upsilon(X, D)\Delta, \\ 0 &= 0 + h(A \otimes I_k)\Gamma(X, D)0 \end{aligned}$$

for $\Delta \in \mathbb{R}^{k \times s}$. The BSI-stability of the methods Radau Ia, IIa and Gauss yields

$$|||\Delta|||_{\bar{V}} \leq \Phi_I(-\frac{h}{\varepsilon \mathcal{K}_1})|||\Upsilon(X, D)\Delta|||_{\bar{V}}$$

where $|||\Delta|||_{\bar{V}} = \|(\|\Delta_1\|_{\bar{V}}, \ldots, \|\Delta_s\|_{\bar{V}})^\top\|_2$ and Φ_I is the corresponding BSI-stability function. We change the norm and obtain

$$\|\Delta\|_\circ \leq \sqrt{s}\mathcal{K}_0 \Phi_I(-\frac{h}{\varepsilon \mathcal{K}_1})\|\Upsilon(X, D)\Delta\|_\circ,$$

where $\mathcal{K}_0 := (M_V M_{V^{-1}})^{\frac{1}{2}}$, which proves the desired result. □

Proposition 4.1.4 *Consider the methods Radau Ia, IIa and Gauss and let* $\Sigma(X, D) := \Omega(X, D) - h(A \otimes I_k)\Omega(X, D)\Gamma(X, D)$. *Let* $r, \|\bar{d}\|_2$ *fulfill the restrictions of proposition 4.1.3 and in addition*

$$r, \|\bar{d}\|_2 \ \leq \ \frac{1}{2R_\Sigma}, \tag{4.18}$$

where the constant R_Σ *is defined in the proof below, then* $\Sigma(X, D)$ *is regular with*

$$\|\Sigma(X, D)^{-1}\|_\circ \ \leq \ 2\sqrt{s}\mathcal{K}_0 \Phi_I(-\frac{h}{\varepsilon \mathcal{K}_1}) \ \leq \ M_{\Sigma^{-1}} \tag{4.19}$$

for all (X, D) *with* $(X_i, D_i) \in \overline{\mathcal{B}}_r \subset \mathcal{W}$ *for* $i = 1, \ldots, s$. *The constant* $M_{\Sigma^{-1}}$ *can be chosen as defined in the proof below.*

Proof: We extend our short notation by $\Sigma := \Sigma(X, D)$, $\Upsilon := \Upsilon(X, D)$ and consider the identity

$$\Sigma^{-1} \ = \ \Upsilon^{-1}\Big(I + (\Omega - I - h(A \otimes I_k)(\Omega - I)\Gamma)\Upsilon^{-1}\Big)^{-1}$$

which holds if we can prove the regularty of the second term on the right hand side. Therefore we apply (4.17), (2.67) and (2.21) to bound

$$\begin{aligned} &\|(\Omega - I - h(A \otimes I_k)(\Omega - I)\Gamma)\Upsilon^{-1}\|_\circ \leq \\ &\leq \ \sqrt{s}\mathcal{K}_0 \Phi_I(-\frac{h}{\varepsilon \mathcal{K}_1})(1 + \frac{h}{\varepsilon} M_G \|A\|_\infty) 2L_\Theta \max\{r, \|\bar{d}\|_2\} \ \leq \\ &\leq R_\Sigma \max\{r, \|\bar{d}\|_2\}, \end{aligned}$$

where R_Σ is defined as [4)]

$$R_\Sigma \ := \ \sup_{\frac{h}{\varepsilon} > 0} \sqrt{s}\mathcal{K}_0 \Phi_I(-\frac{h}{\varepsilon \mathcal{K}_1})(1 + \frac{h}{\varepsilon} M_G \|A\|_\infty) 2L_\Theta.$$

[4)] Due to the structure of the BSI-stability function Φ_I the constant R_Σ is moderate (compare section B.2.1).

Now the restriction (4.18) ensures the regularity as well as the bound

$$\Big\|\Big(I + (\Omega - I - h(A \otimes I_k)(\Omega - I)\Gamma)\Upsilon^{-1}\Big)^{-1}\Big\|_\circ \leq 2$$

Now the identity stated above yields

$$\|\Sigma^{-1}\|_\circ \leq 2\|\Upsilon^{-1}\|_\circ$$

which together with proposition 4.1.3 gives the desired result, where

$$M_{\Sigma^{-1}} := \sup_{\frac{h}{\varepsilon}>0} 2\sqrt{s}\mathcal{K}_0\Phi_I(-\frac{h}{\varepsilon\mathcal{K}_1})$$

is moderate due to the structure of the BSI-stability function (compare section B.2.1). □

Now we reformulate (4.10), (4.11) as a fixed point equation. Under the assumptions of proposition 4.1.2-4.1.4 the equations (4.10), (4.11) are equivalent to

$$X = h(A \otimes I_{n-k})\hat{\Psi} - \Big(\hat{\Omega} - h(A \otimes I_{n-k})\hat{\Omega}\Gamma\Big)D, \tag{4.20}$$

$$D = \Sigma^{-1}\Big(e \otimes \bar{d} - (\Phi - h(A \otimes I_k)\Psi)\Big). \tag{4.21}$$

The fixed point theorem requires a right hand side which is a contraction. In order to obtain this property we compute $h\Gamma D$ from (4.11)

$$h\Gamma D = \Omega^{-1}(A \otimes I_k)^{-1}\Big(\Omega D - (e \otimes \bar{d} - \Phi + h(A \otimes I_k)\Psi)\Big) \tag{4.22}$$

which leads to

$$\begin{aligned}\Big(\hat{\Omega} - h(A \otimes I_{n-k})\hat{\Omega}\Gamma\Big)D &= \\ = \Big(\hat{\Omega} - (A \otimes I_{n-k})\hat{\Omega}\Omega^{-1}(A \otimes I_k)^{-1}\Omega\Big)D &+ \\ + (A \otimes I_{n-k})\hat{\Omega}\Omega^{-1}(A \otimes I_k)^{-1}&\Big(e \otimes \bar{d} - \Phi + h(A \otimes I_k)\Psi\Big).\end{aligned}$$

We define

$$P := (\hat{\Omega} - Q\Omega)D, \tag{4.23}$$

$$Q := (A \otimes I_{n-k})\hat{\Omega}\Omega^{-1}(A \otimes I_k)^{-1} \tag{4.24}$$

such that we obtain

$$X = h(A \otimes I_k)\hat{\Psi} - P - Q(e \otimes \bar{d} - \Phi + h(A \otimes I_k)\Psi). \tag{4.25}$$

Now we replace the argument X of $\Phi = \Phi(X)$ in (4.21) by the right hand side of (4.25). This leads to the following fixed point equation

$$(X, D) = F(X, D) \tag{4.26}$$

where $F := (F_1, F_2)$ is defined via

$$F_1 := h(A \otimes I_{n-k})\hat{\Psi} - P - Q(e \otimes \bar{d} - \Phi + h(A \otimes I_k)\Psi), \tag{4.27}$$

$$F_2 := \Sigma^{-1}(e \otimes \bar{d} - \Phi \circ F_1 + h(A \otimes I_k)\Psi) \tag{4.28}$$

In the next section it is shown that the fixed point theorem A.1.1 can be applied to obtain existence and uniqueness for (4.26). [5)]

[5)] Note that the manipulations which led to (4.26) leave the solution set invariant (compare section A.1.1 in the appendix).

4.2 Solvability of the Algebraic Equations

Our next goal is to apply the fixed point theorem A.1.1 to solve the algebraic equations. Therefore we have to show that the function $F(X, D)$ fulfills the assumptions of the fixed point theorem. First we consider the F-differences to obtain conditions for contractivity. The following Propositions are required.

Proposition 4.2.1 *Let the assumptions of proposition 4.1.2-4.1.4 be fulfilled. Then there exists a moderate real constant $L_{\Sigma^{-1}} \geq 0$ such that*

$$\|\Sigma(X, D)^{-1} - \Sigma(\tilde{X}, \tilde{D})^{-1}\|_\circ \leq L_{\Sigma^{-1}} \|(X, D) - (\tilde{X}, \tilde{D})\|_\diamond \tag{4.29}$$

for all $(X, D), (\tilde{X}, \tilde{D})$ with $(X_i, D_i), (\tilde{X}_i, \tilde{D}_i) \in \overline{\mathcal{B}}_r \subset \mathcal{W}$ for $i = 1, \ldots, s$. The constant $L_{\Sigma^{-1}}$ can be chosen as stated in the proof.

Proof: We use $\tilde{\Sigma}$, $\tilde{\Omega}$, $\tilde{\Gamma}$ to denote the perturbed quantities $\Sigma(\tilde{X}, \tilde{D})$, $\Omega(\tilde{X}, \tilde{D})$, $\Gamma(\tilde{X}, \tilde{D})$. First we estimate

$$\begin{aligned}
\|\Sigma - \tilde{\Sigma}\|_\circ &= \|\Omega - \tilde{\Omega} - h(A \otimes I_k)(\Omega\Gamma - \tilde{\Omega}\tilde{\Gamma})\|_\circ \leq \\
&\leq \|\Omega - \tilde{\Omega}\|_\circ + h\|A\|_\infty(\|\Gamma - \tilde{\Gamma}\|_\circ + \|\Omega - \tilde{\Omega}\|_\circ\|\tilde{\Gamma}\|_\circ) \leq \\
&\leq \Big(L_\Theta + \frac{h}{\varepsilon}\|A\|_\infty(L_G L_\eta + L_\Theta M_G)\Big)\|(X, D) - (\tilde{X}, \tilde{D})\|_\diamond.
\end{aligned} \tag{4.30}$$

The identity $\Sigma^{-1} - \tilde{\Sigma}^{-1} = -\Sigma^{-1}(\Sigma - \tilde{\Sigma})\tilde{\Sigma}^{-1}$ and the estimates (4.19), (4.31) lead us to

$$\begin{aligned}
&\|\Sigma^{-1} - \tilde{\Sigma}^{-1}\|_\circ \leq \\
&\leq 4s\mathcal{K}_0^2\Phi_I(-\frac{h}{\varepsilon\mathcal{K}_1})^2\Big(L_\Theta + \frac{h}{\varepsilon}\|A\|_\infty(L_G L_\eta + M_G L_\Theta)\Big)\|(X, D) - (\tilde{X}, \tilde{D})\|_\diamond = \\
&= L_{\Sigma^{-1}}\|(X, D) - (\tilde{X}, \tilde{D})\|_\diamond
\end{aligned}$$

where the constant $L_{\Sigma^{-1}}$ is defined as

$$L_{\Sigma^{-1}} := \sup_{\frac{h}{\varepsilon} > 0} 4s\mathcal{K}_0^2\Phi_I(-\frac{h}{\varepsilon\mathcal{K}_1})^2\Big(L_\Theta + \frac{h}{\varepsilon}\|A\|_\infty(L_G L_\eta + M_G L_\Theta)\Big),$$

which is moderate (due to the special structure of the BSI-stability function; compare section B.2.1 in the appendix). □

Proposition 4.2.2 *Let the assumptions of proposition 4.1.2 be fulfilled. Then there holds*

$$\|P(X, D)\|_\circ \leq \begin{cases} C_P r, \\ M_P, \end{cases} \tag{4.31}$$

$$\|P(X, D) - P(\tilde{X}, \tilde{D})\|_\circ \leq B_P \max\{r, \|\bar{d}\|_2\}\|(X, D) - (\tilde{X}, \tilde{D})\|_\diamond, \tag{4.32}$$

$$\|Q(X, D)\|_\circ \leq \begin{cases} C_Q \max\{r, \|\bar{d}\|_2\}, \\ M_Q, \end{cases} \tag{4.33}$$

$$\|Q(X, D) - Q(\tilde{X}, \tilde{D})\|_\circ \leq L_Q\|(X, D) - (\tilde{X}, \tilde{D})\|_\diamond \tag{4.34}$$

for all $(X, D), (\tilde{X}, \tilde{D})$ with $(X_i, D_i), (\tilde{X}_i, \tilde{D}_i) \in \overline{\mathcal{B}}_r \subset \mathcal{W}$ for $i = 1, \ldots, s$. The constants C_P, M_P, B_P, C_Q, M_Q, L_Q can be chosen as stated in the proof.

Proof: In our short notation there holds

$$\begin{aligned}\|Q\|_\circ &= \|(A\otimes I_{n-k})\hat{\Omega}\Omega^{-1}(A\otimes I_k)^{-1}\|_\circ \leq \\ &\leq 4\|A\|_\infty\|A^{-1}\|_\infty L_\Theta \max\{r,\|\bar{d}\|_2\} = \\ &= C_Q\max\{r,\|\bar{d}\|_2\}\end{aligned}$$

where $C_Q := 4\|A\|_\infty\|A^{-1}\|_\infty L_\Theta$. On the other hand there holds

$$\|Q\|_\circ = \|(A\otimes I_{n-k})\hat{\Omega}\Omega^{-1}(A\otimes I_k)^{-1}\|_\circ \leq 2\|A\|_\infty\|A^{-1}\|_\infty =: M_Q.$$

Now we consider the Q-differences. We estimate

$$\begin{aligned}\|Q-\tilde{Q}\|_\circ &= \|(A\otimes I_{n-k})(\hat{\Omega}\Omega^{-1}-\tilde{\hat{\Omega}}\tilde{\Omega}^{-1})(A\otimes I_k)^{-1}\|_\circ \leq \\ &\leq 5\|A\|_\infty\|A^{-1}\|_\infty L_\Theta\|(X,D)-(\tilde{X},\tilde{D})\|_\diamond = \\ &= L_Q\|(X,D)-(\tilde{X},\tilde{D})\|_\diamond\end{aligned}$$

where $L_Q := 5\|A\|_\infty\|A^{-1}\|_\infty L_\Theta$. With the help of (4.33) there follows

$$\begin{aligned}\|P\|_\circ &= \|(\hat{\Omega}-Q\Omega)D\|_\circ \leq (1+M_Q)r = \\ &\leq C_Q r \leq C_Q\min\left\{\frac{1}{2\sqrt{k}L_\pi},\frac{1}{4L_\Theta}\right\} =: M_P,\end{aligned}$$

where $C_Q := (1+M_Q)$. For the P-differences we have

$$\begin{aligned}\|P-\tilde{P}\|_\circ &= \|(\hat{\Omega}-\tilde{\hat{\Omega}})D-(Q-\tilde{Q})\Omega D-\tilde{Q}(\Omega-\tilde{\Omega})D+ \\ &\quad +(\tilde{\hat{\Omega}}-\tilde{Q}\tilde{\Omega})(D-\tilde{D})\|_\circ \leq \\ &\leq r(L_\Theta+L_Q+M_QL_\Theta)\|(X,D)-(\tilde{X},\tilde{D})\|_\diamond + \\ &\quad +(2L_\Theta+C_Q)\max\{r,\|\bar{d}\|_2\}\|D-\tilde{D}\|_\circ\end{aligned}$$

such that the constant B_P in (4.32) can be defined as $B_P := (3+M_Q)L_\Theta + L_Q + C_Q$. □

Proposition 4.2.3 *Let the assumptions of proposition 4.1.2-4.1.4 be fulfilled. Then there holds*

$$\|F_1(X,D)-F_1(\tilde{X},\tilde{D})\|_\circ \leq B_{F_1}\max\{h,\|\bar{d}\|_2,r\}\|(X,D)-(\tilde{X},\tilde{D})\|_\diamond, \tag{4.35}$$

$$\|F_1(X,D)\|_\circ \leq C_{F_1}\max\{h,\|\bar{d}\|_2,r\}, \tag{4.36}$$

$$\|F_2(X,D)-F_2(\tilde{X},\tilde{D})\|_\circ \leq B_{F_2}\max\{h,\|\bar{d}\|_2,r\}\|(X,D)-(\tilde{X},\tilde{D})\|_\diamond \tag{4.37}$$

for all (X,D), $(\tilde{X},\tilde{D})$ with $(X_i,D_i),(\tilde{X}_i,\tilde{D}_i)\in\overline{\mathcal{B}}_r\subset\mathcal{W}$ for $i=1,\ldots,s$. The constants B_{F_1}, C_{F_1} and B_{F_2} can be chosen as stated in the proof.

Proof: In our short notation there holds

$$\|F_1-\tilde{F}_1\|_\circ =$$

$$
\begin{aligned}
&= \|h(A\otimes I_{n-k})(\hat{\Psi}-\tilde{\hat{\Psi}})-(P-\tilde{P})-(Q-\tilde{Q})(e\otimes\bar{d}-\Phi+h(A\otimes I_k)\Psi)+\\
&\quad+\tilde{Q}(\Phi-\tilde{\Phi}-h(A\otimes I_k)(\Psi-\tilde{\Psi}))\|_\circ \leq\\
&\leq h\|A\|_\infty L_\psi\|X-\tilde{X}\|_\circ + B_P\max\{r,\|\bar{d}\|_2\}\|(X,D)-(\tilde{X},\tilde{D})\|_\diamond +\\
&\quad+L_Q(\|\bar{d}\|_2+L_\phi r+h\|A\|_\infty M_{f|\mathcal{M}})\|(X,D)-(\tilde{X},\tilde{D})\|_\diamond +\\
&\quad+C_Q\max\{r,\|\bar{d}\|_2\}(L_\phi+h\|A\|_\infty L_\psi)\|X-\tilde{X}\|_\circ \leq\\
&\leq B_{F_1}\max\{h,\|\bar{d}\|_2,r\}\|(X,D)-(\tilde{X},\tilde{D})\|_\diamond
\end{aligned}
$$

where B_{F_1} is defined as

$$
\begin{aligned}
B_{F_1} &:= \|A\|_\infty L_\psi + B_P + L_Q(1+L_\phi+\|A\|_\infty M_{f|\mathcal{M}}) +\\
&\quad + C_Q(L_\phi+\|A\|_\infty L_\psi\frac{1}{4L_\Theta}).
\end{aligned}
$$

The estimate (4.36) follows from

$$
\begin{aligned}
\|F_1(X,D)\|_\circ &= \|h(A\otimes I_{n-k})\hat{\Psi}-P-Q(e\otimes\bar{d}-\Phi+h(A\otimes I_k)\Psi)\|_\circ \leq\\
&\leq h\|A\|_\infty M_{f|\mathcal{M}}+C_P r+M_Q(\|\bar{d}\|_2+L_\phi r+h\|A\|_\infty M_{f|\mathcal{M}}) \leq\\
&\leq C_{F_1}\max\{h,\|\bar{d}\|_2,r\}
\end{aligned}
$$

where C_{F_1} is defined as

$$
C_{F_1} := \|A\|_\infty M_{f|\mathcal{M}}+C_P+M_Q(1+L_\psi+\|A\|_\infty M_{f|\mathcal{M}}).
$$

Now we consider the F_2-differences

$$
\begin{aligned}
&\|F_2-\tilde{F}_2\|_\circ =\\
&\quad= \|(\Sigma^{-1}-\tilde{\Sigma}^{-1})(e\otimes\bar{d}-\Phi\circ F_1+h(A\otimes I_k)\Psi -\\
&\qquad-\tilde{\Sigma}^{-1}(\Phi\circ F_1-\Phi\circ\tilde{F}_1-h(A\otimes I_k)(\Psi-\tilde{\Psi}))\|_\circ \leq\\
&\quad\leq L_{\Sigma^{-1}}(\|\bar{d}\|_2+L_\phi C_{F_1}\max\{h,\|\bar{d}\|_2,r\}+h\|A\|_\infty M_{f|\mathcal{M}})\|(X,D)-(\tilde{X},\tilde{D})\|_\diamond +\\
&\qquad+M_{\Sigma^{-1}}(L_\phi C_{F_1}\max\{h,r,\|\bar{d}\|_2\}+h\|A\|_\infty L_\psi)\|(X,D)-(\tilde{X},\tilde{D})\|_\diamond \leq\\
&\quad\leq B_{F_2}\max\{h,\|\bar{d}\|_2,r\}\|(X,D)-(\tilde{X},\tilde{D})\|_\diamond
\end{aligned}
$$

where B_{F_2} is defined as

$$
B_{F_2} := L_{\Sigma^{-1}}(1+L_\phi C_{F_1}+\|A\|_\infty M_{f|\mathcal{M}})+M_{\Sigma^{-1}}(L_\phi C_{F_1}+\|A\|_\infty L_\psi).
$$

□

Proposition 4.2.3 ensures that under mild restrictions on $h, \|\bar{d}\|_2, r$ we have contractivity for F. Next we consider the first iteration step:

Proposition 4.2.4 *Let the assumptions of proposition 4.1.2-4.1.4 be fulfilled. Then there holds*

$$\|F_1(0,0)\|_\circ \leq C_{F_1(0)} \max\{h, \|\bar{d}\|_2\}, \tag{4.38}$$

$$\|F_2(0,0)\|_\circ \leq C_{F_2(0)} \max\{h, \|\bar{d}\|_2\}. \tag{4.39}$$

The constants $C_{F_1(0)}, C_{F_2(0)}$ can be chosen as stated in the proof.

Proof: We estimate

$$\begin{aligned}
\|F_1(0,0)\|_\circ &= \|h(A \otimes I_{n-k})\hat{\Psi}(0) - Q(0,0)(e \otimes \bar{d} + h(A \otimes I_k)\Psi(0))\|_\circ \leq \\
&\leq h\|A\|_\infty M_{f|\mathcal{M}} + M_Q(\|\bar{d}\|_2 + h\|A\|_\infty M_{f|\mathcal{M}}) \leq \\
&\leq C_{F_1(0)} \max\{h, \|\bar{d}\|_2\}
\end{aligned}$$

where $C_{F_1(0)} := \|A\|_\infty M_{f|\mathcal{M}}(1 + M_Q) + M_Q$. In the same way we derive

$$\begin{aligned}
\|F_2(0,0)\|_\circ &= \|\Sigma(0,0)^{-1}(e \otimes \bar{d} - \Phi \circ F_1(0,0) + h(A \otimes I_k)\Psi(0))\|_\circ \leq \\
&\leq M_{\Sigma^{-1}}(\|\bar{d}\|_2 + L_\phi C_{F_1(0)} \max\{h, \|\bar{d}\|_2\} + h\|A\|_\infty M_{f|\mathcal{M}}) \leq \\
&\leq C_{F_2(0)} \max\{h, \|\bar{d}\|_2\}
\end{aligned}$$

where $C_{F_2(0)} := M_{\Sigma^{-1}}(1 + L_\phi C_{F_1(0)} + \|A\|_\infty M_{f|\mathcal{M}})$. □

Theorem 4.2.1 (Algebraic equations) *Let $C_\diamond := \max\{2C_{F_1(0)}, 2C_{F_2(0)}, 1\}$ and $C^\diamond := 2\max\{B_{F_1}, B_{F_2}, R_\Sigma, \sqrt{k}L_\pi, 2L_\Theta, 2kL_G L_\eta M_V\}$ and $\overline{\mathcal{B}}_r \subset \mathcal{W}$. If the restriction*

$$C_\diamond \max\{h, \|\bar{d}\|_2\} \leq r \leq \frac{1}{C^\diamond} \tag{4.40}$$

is fulfilled, then the fixed point form (4.26) for the parametrized algebraic equations in $\overline{\mathcal{B}}_r^s$ posesses a unique solution (X_i, D_i), $i = 1, \ldots, s$. Furthermore $Y_i := \eta(X_i, D_i)$, $i = 1, \ldots, s$ is the unique solution of the algebraic equations (4.3) in $\eta(\overline{\mathcal{B}}_r)^s$.

Proof: Proposition 4.2.3 and the restriction (4.40) ensure the estimate

$$\|F(X,D) - F(\tilde{X}, \tilde{D})\|_\diamond \leq K\|(X,D) - (\tilde{X}, \tilde{D})\|_\diamond$$

where $K := \max\{h, \|\bar{d}\|_2, r\}\max\{B_{F_1}, B_{F_2}\} \leq \frac{1}{2} < 1$, i.e. contractivity. Furthermore we can use proposition 4.2.4 to obtain

$$\|F(0,0)\|_\diamond \leq \max\{C_{F_1(0)}, C_{F_2(0)}\}\max\{h, \|\bar{d}\|_2\} \leq \frac{1}{2}r \leq (1-K)r.$$

Application of theorem A.1.1 yields the desired result. □

Remarks: Theorem 4.2.1 can be applied in two different ways:

- First let $r := \frac{1}{C^\diamond}$. Then $\overline{\mathcal{B}}_r^s$ is a ball of $O(1)$ in which the solution (X_i, D_i), $i = 1, \ldots, s$ is unique.
- Second let $r := C_\diamond \max\{h, \|\bar{d}\|_2\}$, which leads to a first estimate for the solution (X, D):

$$\|(X,D)\|_\diamond \leq C_\diamond \max\{h, \|\bar{d}\|_2\} \tag{4.41}$$

4.3 Parametrization of the Implicit Runge-Kutta Scheme

First we apply theorem 4.2.1 with $r := C_\diamond \max\{h, \|\bar{d}\|_2\}$, which means that under a moderate stepsize restriction and a moderate restriction on $\|\bar{d}\|_2$ the stages Y_i, $i = 1, \ldots, s$ are well defined such that we can compute η_ν from $\bar{\eta} := \eta_{\nu-1}$. In this section we show that for $h, \|\bar{d}\|_2$ sufficiently small the new approximation η_ν is sufficiently close to $\bar{\eta} = \eta_{\nu-1}$ such that there exist coordinates (x, d) with $\eta_\nu = \eta(x, d)$. This allows us to formulate a Runge Kutta step in the parametrization $\eta(x, d)$. Furthermore an a-priori estimate for the coordinates (x, d) of η_ν is obtained.
In this context the following assumption is necessary. It can be obtained as a consequence of the assumptions stated in section 2.2 (which can be shown by an application of the fixed point theorem A.1.1 but is omitted here).

Assumption concerning the coordinates (x, d)

- There exist moderate real constants $C_1, C_2 > 0$ such that for all $\bar{\eta} \in \mathcal{G}$ with the corresponding parametrization $\eta(x, d)$ and for all $y \in \mathbb{R}^n$ with

$$\|\bar{\eta} - y\|_2, \|\bar{d}\|_2 \leq \frac{1}{C_1} \tag{4.42}$$

  there exists $(x, d) \in \mathcal{W}$ such that $y = \eta(x, d)$. Furthermore $(x, d) = (x, d)(y)$ satisfies the estimate

$$\|(x, d)\|_* \leq C_2 \max\{\|\bar{d}\|_2, \|\bar{\eta} - y\|_2\}. \tag{4.43}$$

The following propositions together with the previous assumption lead to the a-priory bound for (x, d).

Proposition 4.3.1 *Let $r := C_\diamond \max\{h, \|\bar{d}\|_2\}$ with $h, \|\bar{d}\|_2$ sufficiently small such that the assumptions of theorem 4.2.1 are fulfilled. Then the locally unique solution (X, D) of the parametrized algebraic equation satisfies*

$$\|D\|_\circ \leq C_D \Phi_I(-\frac{h}{\varepsilon \mathcal{K}_1}) \max\{h, \|\bar{d}\|_2\}, \tag{4.44}$$

$$\|h\Gamma D\|_\circ \leq C_{h\Gamma D} \max\{h, \|\bar{d}\|_2\}. \tag{4.45}$$

The constants $C_D, C_{h\Gamma D}$ can be chosen as in the proof below.

Proof: The equation (4.21) leads to

$$\begin{aligned}
\|D\|_\circ &= \|\Sigma^{-1}(e \otimes \bar{d} - \Phi + h(A \otimes I_k)\Psi)\|_\circ \leq \\
&\leq 2\sqrt{s}\mathcal{K}_0 \Phi_I(-\frac{h}{\varepsilon \mathcal{K}_1})(\|\bar{d}\|_2 + L_\phi C_\diamond \max\{h, \|\bar{d}\|_2\} + h\|A\|_\infty M_{f|\mathcal{M}}) \leq \\
&\leq C_D \Phi_I(-\frac{h}{\varepsilon \mathcal{K}_1}) \max\{h, \|\bar{d}\|_2\},
\end{aligned}$$

where the constant C_D is defined as

$$C_D := 2\sqrt{s}\mathcal{K}_0(1 + L_\phi C_\diamond + \|A\|_\infty M_{f|\mathcal{M}}).$$

Now we estimate (4.22)

$$\begin{aligned}\|h\Gamma D\|_\circ &= \|\Omega^{-1}(A\otimes I_k)^{-1}\Big(\Omega D-(e\otimes\bar d-\Phi+h(A\otimes I_k)\Psi)\Big)\|_\circ \le\\ &\le 2\|A^{-1}\|_\infty\Big(C_\diamond\max\{h,\|\bar d\|_2\}+\|\bar d\|_2+\\ &\quad +L_\phi C_\diamond\max\{h,\|\bar d\|_2\}+h\|A\|_\infty M_{f|\mathcal M}\Big)\le\\ &\le C_{h\Gamma D}\max\{h,\|\bar d\|_2\},\end{aligned}$$

where $C_{h\Gamma D}:=2\|A\|_\infty(C_\diamond(1+L_\phi)+1+\|A\|_\infty M_{f|\mathcal M})$. □

The next step is to bound the coordinates of

$$\eta_\nu-\bar\eta \;=\; h\sum_{i=1}^{s}b_i f(Y_i)$$

in the ONS B:

Proposition 4.3.2 *Under the assumptions of proposition 4.3.1 there holds*

$$\left.\begin{aligned}&\|h\sum_{i=1}^{s}b_i(\hat\Psi_i+\hat\Theta_i\Lambda_i D_i)\|_2\\ &\|h\sum_{i=1}^{s}b_i(\Psi_i+\Theta_i\Lambda_i D_i)\|_2\end{aligned}\right\}\le C_3\max\{h,\|\bar d\|_2\}. \tag{4.46}$$

The constant C_3 can be chosen as stated in the proof.

Proof: We estimate

$$\begin{aligned}\left.\begin{aligned}&\|h\sum_{i=1}^{s}b_i(\hat\Psi_i+\hat\Theta_i\Lambda_i D_i)\|_2\\ &\|h\sum_{i=1}^{s}b_i(\Psi_i+\Theta_i\Lambda_i D_i)\|_2\end{aligned}\right\} &\le \|b\|_1(hM_{f|\mathcal M}+C_{h\Gamma D}\max\{h,\|\bar d\|_2\})\le\\ &\le C_3\max\{h,\|\bar d\|_2\},\end{aligned}$$

where $C_3:=\|b\|_1(M_{f|\mathcal M}+C_{h\Gamma D})$. □

Application of proposition 4.3.2 yields

$$\|\eta_\nu-\bar\eta\|_2\le C_3\max\{h,\|\bar d\|_2\}. \tag{4.47}$$

This means that the restriction (4.42)of the assumption concerning the coordinates (x,d) is fulfilled if $h,\|\bar d\|_2$ satisfy

$$h,\|\bar d\|_2\le\frac{1}{C_1\max\{1,C_3\}}. \tag{4.48}$$

Proposition 4.3.3 *Let $r:=C_\diamond\max\{h,\|\bar d\|_2\}$ with $h,\|\bar d\|_2$ sufficiently small, such that the assumptions of proposition 4.3.2 are fulfilled. Then under the additional restriction (4.48) there exist coordinates $(x,d)\in\mathcal W$ with $\eta(x,d)=\eta_\nu$ and the a-priori estimate*

$$\|(x,d)\|_*\le C_*\max\{h,\|\bar d\|_2\} \tag{4.49}$$

holds, where C_* can be chosen as stated in the proof below.

Proof: We estimate

$$\begin{aligned}\|(x,d)\|_* &\leq C_2 \max\{\|\eta_\nu - \bar{\eta}\|_2, \|\bar{d}\|_2\} \leq \\ &\leq C_* \max\{h, \|\bar{d}\|_2\}\end{aligned}$$

where $C_* := C_2 \max\{C_3, 1\}$. □

Now we can reformulate

$$\eta(x,d) = \eta_\nu = \bar{\eta} + h \sum_{i=1}^{s} b_i f(Y_i) \tag{4.50}$$

in a representation for the coordinates (x,d). We use the short notation as usual and $p := p(\eta(x,d))$, $\pi_j := \pi_j(\eta(x,d))$ to formulate

$$p + \sum_{j=1}^{k} d_j \pi_j \;=\; \bar{p} + \sum_{j=1}^{k} \bar{d}_j \bar{\pi}_j + h \sum_{i=1}^{s} b_i (f(P_i) + \sum_{j=1}^{k} [\Lambda_i D_i]_j \Pi_{ij}). \tag{4.51}$$

It is due to $p(\eta(x,d)) = \phi(x)$ such that we can replace p by

$$p = \phi(x) = \bar{p} + \sum_{j=1}^{n-k} x_j \hat{\pi}_j + \sum_{j=1}^{k} \varphi_j \bar{\pi}_j,$$

where we use the short notation $\varphi := (\varphi_1, \ldots, \varphi_k)^\top := \varphi(x)$. Now with the help of the matrices

$$\hat{\Theta} := \hat{\Theta}(x,d), \quad \Theta := \Theta(x,d)$$

(4.51) can be written in the form

$$x + \hat{\Theta} d \;=\; h \sum_{i=1}^{s} b_i (\hat{\Psi}_i + \hat{\Theta}_i \Lambda_i D_i), \tag{4.52}$$

$$\varphi + \Theta d \;=\; \bar{d} + h \sum_{i=1}^{s} b_i (\Psi_i + \Theta_i \Lambda_i D_i). \tag{4.53}$$

Chapter 5

Convergence Estimates for the Methods Radau Ia, IIa and Gauss

In the following sections we make use of the convention that $\mathcal{C}$'s denote generic constants, i.e. $\mathcal{C}$ is a well defined constant which is moderate, independent of the stiffness parameter ε.
In order to derive convergence estimates for the methods Radau Ia, IIa and Gauss we require the following smoothness assumption:

Assumption concerning the smoothness of solutions in $\mathcal{M}$

- Solutions $u(t)$ of the differential equation (2.11) in $\mathcal{M}$ are assumed to be p_o+1 times continuously differentiable where

$$p_o := \begin{cases} 2s & \text{for Gauss,} \\ 2s-1 & \text{for Radau Ia, IIa.} \end{cases}$$

  Furthermore we assume that there exists a moderate real constant $\mathcal{C} > 0$ such that

$$\|u^{(l)}(t)\|_2 \leq \mathcal{C} \tag{5.1}$$

  for $l = 1, \ldots, p_o + 1$.

Now let $u(t) \in \mathcal{M}$, $t \in [0, T]$ be a smooth solution of the differential equation (2.11) with initial value $u(0) = u_0 \in \mathcal{M}$. The corresponding approximations generated by the implicit Runge Kutta methods Radau Ia, IIa and Gauss are denoted by η_ν, $\nu \in \mathbb{N}_0$, where η_0 is an approximation for u_0. In order to derive error bounds we split the error

$$e_\nu := \eta_\nu - u(t_\nu) = (\eta_\nu - p(\eta_\nu)) + (p(\eta_\nu) - u(t_\nu)) \tag{5.2}$$

up into the stiff error component

$$\eta_\nu - p(\eta_\nu) = \sum_{j=1}^{k} d_j(\eta_\nu)\pi_j(\eta_\nu)$$

and into the smooth error component $p(\eta_\nu) - u(t_\nu)$. Furthermore the smooth error component is split up into

$$p(\eta_\nu) - u(t_\nu) \;=\; (p(\eta_\nu) - \tilde{u}_{\nu-1}(t_\nu)) + (\tilde{u}_{\nu-1}(t_\nu) - u(t_\nu)) \tag{5.3}$$

where $p(\eta_\nu) - \tilde{u}_{\nu-1}(t_\nu)$ is the local smooth error, i.e. $\tilde{u}_{\nu-1}(t)$ is the solution of the differential equation (2.11) with initial value $\tilde{u}_{\nu-1}(t_{\nu-1}) = p(\eta_{\nu-1})$. For an illustration compare figure 5.1.

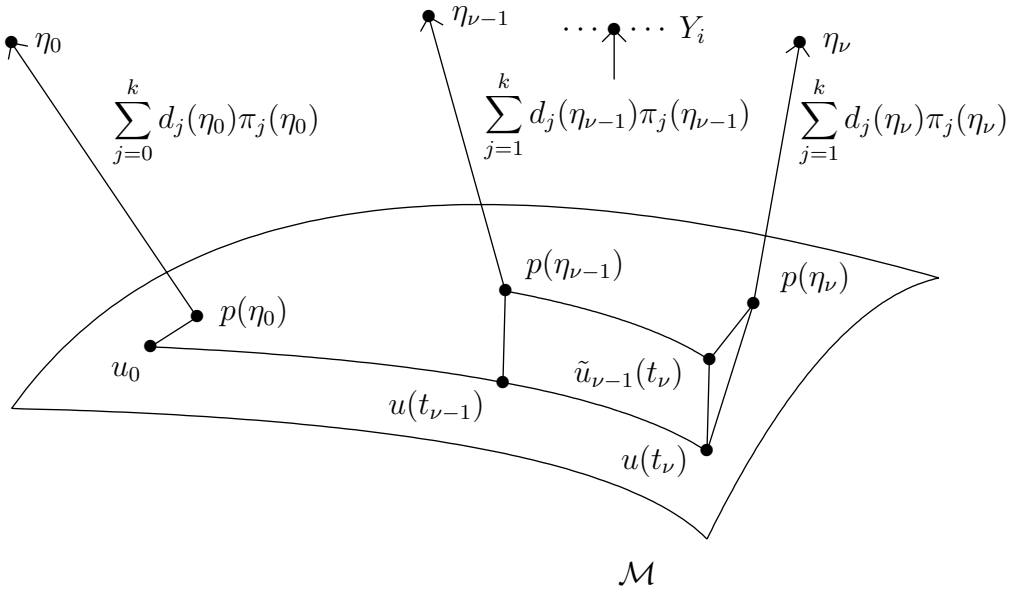


Figure 5.1: Decomposition of the error

5.1 Recursion for the Stiff Error Component

We consider one step $\eta_{\nu-1} \to \eta_\nu$ of the implicit Runge-Kutta methods Radau Ia, IIa and Gauss. In each step we let $\bar{\eta} := \eta_{\nu-1}$ and deal with the corresponding local coordinates (x, d) in a neighbourhood of $\bar{\eta}$. Furthermore we rename the auxiliary solutions $\tilde{u}(t) := \tilde{u}_{\nu-1}(t_{\nu-1}+t)$ such that $\tilde{u}(0) = \bar{p} = p(\bar{\eta})$ and $\tilde{u}(h) = \tilde{u}_{\nu-1}(t_\nu)$. The solution $\tilde{u}(t)$ lies in $\mathcal{M}$ and for h sufficiently small there is $\tilde{u}(t) \in \phi(\mathcal{U})$ for $t \in [0, h]$. This means that there exists $\tilde{x}(t) \in \mathcal{U}$ such that $\phi(\tilde{x}(t)) = \tilde{u}(t)$ for $t \in [0, h]$. Then $\tilde{x}(t)$ is a solution of the differential equation $\tilde{x}'(t) = \hat{\psi}(\tilde{x}(t))$ with $\tilde{x}(0) = 0$. Setting $\bar{d} = d(\bar{\eta})$ we assume that for

$$r := C_\diamond \max\{h, \|\bar{d}\|_2\} \tag{5.4}$$

the assumptions of theorem 4.2.1 as well as (4.48) are fulfilled. This means that a Runge-Kutta step is well defined and there exist local coordinates $(x, d) \in \mathcal{W}$ with $\eta_\nu = \eta(x, d)$ and local coordinates $(X_i, D_i) \in \mathcal{W}$ with $Y_i = \eta(X_i, D_i)$ for $i = 1, \ldots, s$. [1)] Furthermore we can use the estimates (4.41), (4.49). Under the short notation introduced in the previous sections we formulate (4.29), (4.51) as

$$\sum_{j=1}^{k} D_{ij}\Pi_{ij} = \sum_{j=1}^{k} \bar{d}_j\bar{\pi}_j + h\sum_{j=1}^{s} a_{ij}\sum_{l=1}^{k}[\Lambda_j D_j]_l\Pi_{jl} - \Xi_i, \tag{5.5}$$

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

$$\sum_{j=1}^{k} d_j\pi_j = \sum_{j=1}^{s} \bar{d}_j\bar{\pi}_j + h\sum_{j=1}^{s} b_j\sum_{l=1}^{k}[\Lambda_j D_j]_l\Pi_{jl} - \chi, \tag{5.6}$$

$$\chi := p - \bar{p} - h\sum_{j=1}^{s} b_j f(P_j).$$

[1)] Later we will show that under a suiteable induction assumption on $\bar{d}$ together with a moderate step size restriction, these assumptions are fulfilled in each step.

Application of the inner product $\langle \cdot, \pi_m \rangle$, $m = 1, \ldots, k$ to (5.5), (5.6) yields

$$\sum_{j=1}^{k} D_{ij} \langle \Pi_{ij}, \pi_m \rangle = \sum_{j=1}^{k} \bar{d}_j \langle \bar{\pi}_j, \pi_m \rangle + h \sum_{j=1}^{s} a_{ij} \sum_{l=1}^{k} [\Lambda_j D_j]_l \langle \Pi_{jl}, \pi_m \rangle - \langle \Xi_i, \pi_m \rangle, \tag{5.7}$$

$$d_m = \sum_{j=1}^{k} \bar{d}_j \langle \bar{\pi}_j, \pi_m \rangle + h \sum_{j=1}^{s} b_j \sum_{l=1}^{k} [\Lambda_j D_j]_l \langle \Pi_{jl}, \pi_m \rangle - \langle \chi, \pi_m \rangle \tag{5.8}$$

where $i = 1, \ldots, s$ and $m = 1, \ldots, k$. In order to reformulate (5.7), (5.8) in the matrix-vector notation we introduce the matrices $\mho_i := (\langle \Pi_{ij}, \pi_m \rangle)_{mj}$ for $i = 1, \ldots, s$ and obtain

$$\mho_i D_i = \Theta^\top \bar{d} + h \sum_{j=1}^{s} a_{ij} \mho_j \Lambda_j \mho_j^{-1} \mho_j D_j - (\langle \Xi_i, \pi_m \rangle)_{m=1}^{k}, \tag{5.9}$$

$$d = \Theta^\top \bar{d} + h \sum_{j=1}^{s} s b_j \mho_j \Lambda_j \mho_j^{-1} \mho_j D_j - (\langle \chi, \pi_m \rangle)_{m=1}^{k}, \tag{5.10}$$

where $i = 1, \ldots, s$. The next step is to apply the BS- and B- stability (compare sections B.2.2,B.2.3 in the appendix). Therefore we define the auxiliary function $f_\bullet(t, \tilde{d}) := \tilde{\Lambda}(t)\tilde{d}$, where $\tilde{\Lambda}(t)$ is chosen such that there holds $\tilde{\Lambda}(t_{\nu-1} + c_j h) = \mho_j \Lambda_j \mho_j^{-1}$ for $j = 1, \ldots, s$. In addition we introduce the corresponding auxiliary schemes

$$\tilde{D}_i = \Theta^\top \bar{d} + h \sum_{j=1}^{s} a_{ij} \mho_j \Lambda_j \mho_j^{-1} \tilde{D}_j, \quad i = 1, \ldots, s, \tag{5.11}$$

$$\tilde{d} = \Theta^\top \bar{d} + h \sum_{j=1}^{s} b_j \mho_j \Lambda_j \mho_j^{-1} \tilde{D}_j \tag{5.12}$$

and

$$0 = 0 + h \sum_{j=1}^{s} a_{ij} \mho_j \Lambda_j \mho_j^{-1} 0, \quad i = 1, \ldots, s, \tag{5.13}$$

$$0 = 0 + h \sum_{j=1}^{s} b_j \mho_j \Lambda_j \mho_j^{-1} 0. \tag{5.14}$$

To be shure that the function $f_\bullet$ as well as the schemes are well defined we have to prove the regularity of the matrices $\mho_i$, $i = 1, \ldots, s$ and $I - h(A \otimes I_k)\mho\Gamma\mho^{-1}$, where $\mho := \mathrm{diag}(\mho_1, \ldots, \mho_k)$.

Proposition 5.1.1 *Let the assumptions of proposition 4.3.3 be fulfilled. Then there exists a moderate real constant $R_\mho$ such that the restriction*

$$h, \|\bar{d}\|_2 \leq \frac{1}{2R_\mho} \tag{5.15}$$

implies the regularity of $\mho_i$, $i = 1, \ldots, s$ and there holds

$$\|\mho_i^{-1}\|_2 \leq 2, \quad i = 1, \ldots, s. \tag{5.16}$$

The constant $R_\mho$ can be chosen as stated in the proof below.

Proof: Similar to the proof of proposition 3.1.2 we use $\Theta(0,\bar{d}) = I$ to obtain

$$\begin{aligned}\|\Theta_i - I\|_2 &\leq 2L_\Theta \max\{\|(X_i, D_i)\|_*, \|\bar{d}\|_2\} \leq \\ &\leq 2L_\Theta \max\{C_\diamond, 1\} \max\{h, \|\bar{d}\|_2\}, \quad i = 1, \ldots, s,\end{aligned}$$

which holds due to (2.67), (4.41). Additionally we require

$$\begin{aligned}\| \underbrace{(\langle \Pi_{ij}, \bar{\pi}_m \rangle)_{mj}}_{=\Theta_i} - \underbrace{(\langle \Pi_{ij}, \pi_m \rangle)_{mj}}_{=\mho_i} \|_2 &\leq 2L_\Theta \max\{\|(x,d)\|_*, \|\bar{d}\|_2\} \leq \\ &\leq 2L_\Theta \max\{C_*, 1\} \max\{h, \|\bar{d}\|_2\}\end{aligned}$$

which can be proved similar to (2.67) and with the help of (4.49). A combination of these estimates yields

$$\begin{aligned}\|\mho_i \tilde{d}\|_2 &\geq (1 - \|\mho_i - I\|_2)\|\tilde{d}\|_2 \geq \\ &\geq (1 - \|\mho_i - \Theta_i\|_2 - \|\Theta_i - I\|_2)\|\tilde{d}\|_2 \geq \\ &\geq (1 - R_\mho \max\{h, \|\bar{d}\|_2\})\|\tilde{d}\|_2 \geq \frac{1}{2}\|\tilde{d}\|_2\end{aligned}$$

for $i = 1, \ldots, s$, $\tilde{d} \in \mathbb{R}^k$, where $R_\mho := 2L_\Theta(\max\{C_\diamond, 1\} + \max\{C_*, 1\})$. □

Proposition 5.1.2 *Let the assumptions of proposition 5.1.1 be fulfilled. Then there exists a moderate real constant $R_{\mho\Gamma\mho^{-1}}$ such that the restriction*

$$h, \|\bar{d}\|_2 \leq \frac{1}{4R_{\mho\Gamma\mho^{-1}}} \tag{5.17}$$

implies

$$\langle \mho_j \Lambda_j \mho_j^{-1} \tilde{d}, \tilde{d} \rangle_V \leq -\frac{1}{\varepsilon \mathcal{K}_1} \langle \tilde{d}, \tilde{d} \rangle_V, \quad j = 1, \ldots, s, \quad \tilde{d} \in I\!R^k, \tag{5.18}$$

where $V = V(\eta(x,d))$ and $\langle \cdot, \cdot \rangle_V := \langle V \cdot, \cdot \rangle$ is the corresponding inner product. Furthermore there holds

$$\|(I - h(A \otimes I_k)\mho\Gamma\mho^{-1})^{-1}\|_\circ \leq \sqrt{s}\mathcal{K}_0 \Phi_I(-\frac{h}{\varepsilon \mathcal{K}_1}). \tag{5.19}$$

The constant $R_{\mho\Gamma\mho^{-1}}$ can be chosen as stated in the proof below.

Proof: First we note that there holds

$$\begin{aligned}\|\mho_j - I\|_2 &\leq R_\mho \max\{h, \|\bar{d}\|_2\} \\ \|\mho_j^{-1} - I\|_2 &\leq \|\mho_j^{-1}\|_2 \|I - \mho_j\|_2 \leq 2R_\mho \max\{h, \|\bar{d}\|_2\}\end{aligned}$$

for $j = 1, \ldots, s$ (compare proof of proposition 5.1.1). This implies

$$\begin{aligned}&\|\mho_j \Lambda_j \mho_j^{-1} - \Lambda\|_2 \leq \\ &\leq \|(\mho_j - I)\Lambda_j \mho_j^{-1}\|_2 + \|(\Lambda_j - \Lambda)\mho_j^{-1}\|_2 + \|\Lambda(\mho_j^{-1} - I)\|_2 \leq \\ &\leq \frac{1}{\varepsilon} 2(2R_\mho M_G + kL_G L_\eta) \max\{h, \|\bar{d}\|_2\} = \frac{1}{\varepsilon}\mathcal{C} \max\{h, \|\bar{d}\|_2\},\end{aligned}$$

such (5.17) with $R_{\mathfrak{V}\Gamma\mathfrak{V}^{-1}} := \mathcal{C}M_V$, where $\mathcal{C}$ is defined via the previous equation, can be used to derive

$$\begin{aligned}
\langle \mathfrak{V}_j\Lambda_j\mathfrak{V}_j^{-1}\tilde{d},\tilde{d}\rangle_V &= \langle \Lambda\tilde{d},\tilde{d}\rangle_V + \langle (\mathfrak{V}_j\Lambda_j\mathfrak{V}_j^{-1}-\Lambda)\tilde{d},\tilde{d}\rangle_V \leq \\
&\leq \Big(-\frac{1}{2\varepsilon} + \|\mathfrak{V}_j\Lambda_j\mathfrak{V}_j^{-1}-\Lambda\|_2\|V\|_2\Big)\|\tilde{d}\|_2^2 \leq \\
&\leq -\frac{1}{\varepsilon}\Big(\frac{1}{2} - R_{\mathfrak{V}\Gamma\mathfrak{V}^{-1}}\max\{h,\|\bar{d}\|_2\}\Big)\|\tilde{d}\|_2^2 \leq \\
&\leq -\frac{1}{4\varepsilon}\|\tilde{d}\|_2^2 \leq -\frac{1}{\varepsilon 4M_V}\|\tilde{d}\|_V^2
\end{aligned}$$

for all $\tilde{d} \in \mathbb{R}^k$. Now the BSI-stability can be applied to the function $f_\bullet(t,\tilde{d}) = \tilde{\Lambda}(t)\tilde{d}$ where $\tilde{\Lambda}(t)$ is chosen such that there holds $\tilde{\Lambda}(t_{\nu-1}+c_jh) = \mathfrak{V}_j\Lambda_j\mathfrak{V}_j^{-1}$ for $j = 1,\ldots,s$. Since $f_\bullet(t,\tilde{d})$ satisfies a one sided Lipschitz condition with the one sided Lipschitz constant $m = -\frac{1}{\varepsilon\mathcal{K}_1}$ corresponding to the inner product $\langle\cdot,\cdot\rangle_V$ we obtain

$$|||\Delta|||_V \leq \Phi_I(-\frac{h}{\varepsilon\mathcal{K}_1})|||(I-h(A\otimes I_k)\mathfrak{V}\Gamma\mathfrak{V}^{-1})\Delta|||_V$$

for all $\Delta \in \mathbb{R}^{s\times k}$. We change the norm which yields

$$\|\Delta\|_\circ \leq \sqrt{s}\mathcal{K}_0\Phi_I(-\frac{h}{\varepsilon\mathcal{K}_1})\|(I-h(A\otimes I_k)\mathfrak{V}\Gamma\mathfrak{V}^{-1})\Delta\|_\circ$$

for all $\Delta \in \mathbb{R}^{s\times k}$. This proves the proposition. □

The estimate (5.18) of proposition 5.1.2 allows us to apply the B- and BS-stability inequalities to the schemes (5.9)-(5.14). Therefore we introduce

$$\Xi_i^* := (\langle \Xi_i,\pi_m\rangle)_{m=1}^k, \quad \chi^* := (\langle \chi,\pi_m\rangle)_{m=1}^k$$

and estimate

$$\begin{aligned}
\|d\|_V &\leq \|d-\tilde{d}\|_V + \|\tilde{d}-0\|_V \leq \\
&\leq \Phi_B(-\frac{h}{\varepsilon\mathcal{K}_1})\|\Theta^\top\bar{d}\|_V + \Phi_{BS}(-\frac{h}{\varepsilon\mathcal{K}_1})\Big(|||\Xi^*|||_V + \|\chi^*\|_V\Big).
\end{aligned}$$

The next step is to consider the quantity $\|\Theta^\top\bar{d}\|_V$. We apply proposition 2.4.5 and obtain

$$\begin{aligned}
\|\Theta^\top\bar{d}\|_V &\leq \|\bar{d}\|_V + \|(\Theta^\top - I)\bar{d}\|_V \leq \\
&\leq \|\bar{d}\|_{\bar{V}} + |\|\bar{d}\|_V - \|\bar{d}\|_{\bar{V}}| + M_V^{\frac{1}{2}}\|\Theta - I\|_2\|\bar{d}\|_2 \leq \\
&\leq \|\bar{d}\|_{\bar{V}} + L_{V^{\frac{1}{2}}}L_\eta(C_*+1)\max\{h,\|\bar{d}\|_2\}\|\bar{d}\|_2 + \\
&\quad + M_V^{\frac{1}{2}}2L_\Theta\max\{C_*,1\}\max\{h,\|\bar{d}\|_2\}\|\bar{d}\|_2 = \\
&= \|\bar{d}\|_{\bar{V}} + \mathcal{C}\max\{h,\|\bar{d}\|_2\}\|\bar{d}\|_2,
\end{aligned}$$

where $\mathcal{C}$ is defined via the previous equation. Insertion into the estimate for $\|d\|_V$ yields

$$\begin{aligned}
\|d\|_V &\leq \Phi_B(-\frac{h}{\varepsilon\mathcal{K}_1})\Big(\|\bar{d}\|_{\bar{V}} + \|\bar{d}\|_2\mathcal{C}\max\{h,\|\bar{d}\|_2\}\Big) + \\
&\quad + \sqrt{s}M_V^{\frac{1}{2}}\Phi_{BS}(-\frac{h}{\varepsilon\mathcal{K}_1})\Big(\|\Xi^*\|_\circ + \|\chi^*\|_2\Big). \qquad (5.20)
\end{aligned}$$

Remark: For the method Radau IIa the equations (5.10), (5.12), (5.14) are redundant such that the BS-stability of the Radau IIa method reduces to the BSI-stability. This means that in the case of the Radau IIa method the BS-stability function in (5.20) can be replaced by the BSI-stability function (and the quantity $\|\chi^*\|_2$ is redundant).

In order to derive bounds for $\|\Xi^*\|_\circ$, $\|\chi^*\|_2$ we introduce the truncation error

$$\mathcal{T}_i := \tilde{u}(c_i h) - \bar{p} - h\sum_{j=1}^{s} a_{ij} f(\tilde{u}(c_j h)), \quad i = 1, \ldots, s, \tag{5.21}$$

$$\tau := \tilde{u}(h) - \bar{p} - h\sum_{i=1}^{s} b_i f(\tilde{u}(c_i h)) \tag{5.22}$$

corresponding to the auxiliary solutions $\tilde{u}(t)$.

Proposition 5.1.3 *There exists a moderate real constant $\mathcal{C} \geq 0$ such that there holds*

$$\|\mathcal{T}_i\|_2 \leq \mathcal{C} h^{q_o+1}, \quad q_o = \begin{cases} s & \textit{Gauss, Radau IIa} \\ s-1 & \textit{Radau Ia} \end{cases} \tag{5.23}$$

$$\|\tau\|_2 \leq \mathcal{C} h^{p_o+1}, \quad p_o = \begin{cases} 2s & \textit{Gauss} \\ 2s-1 & \textit{Radau Ia, IIa} \end{cases} \tag{5.24}$$

for the auxiliary sulutions $\tilde{u}(t)$ with $\tilde{u}(0) = \bar{p} \in \mathcal{M}$.

Proof: A Taylor series expansion yields

$$\begin{aligned} \tilde{u}(h) &= \sum_{l=0}^{L} \frac{h^l}{l!} \tilde{u}^{(l)}(0) + O(h^{L+1}), \\ \tilde{u}'(c_i h) &= \sum_{l=0}^{L-1} \frac{(c_i h)^l}{l!} \tilde{u}^{(l+1)}(0) + O(h^L), \end{aligned}$$

where the O-constants are moderate. Insertion into the definition of τ yields

$$\begin{aligned} \tau &= \tilde{u}(h) - \bar{p} - h\sum_{i=1}^{s} b_i f(\tilde{u}(c_i h)) = \\ &= \tilde{u}(h) - \tilde{u}(0) - h\sum_{i=1}^{s} b_i \tilde{u}'(c_i h) = \\ &= \sum_{l=0}^{L} \frac{h^l}{l!} \tilde{u}^{(l)}(0) - \tilde{u}(0) - h\sum_{i=1}^{s} b_i \sum_{l=0}^{L-1} \frac{(c_i h)^l}{l!} \tilde{u}^{l-1}(0) + O(h^{L+1}) = \\ &= \sum_{l=0}^{L} \frac{h^l}{(l-1)!} \tilde{u}^{(l)}(0) \underbrace{\Big(\frac{1}{l} - \sum_{i=1}^{s} b_i c_i^{l-1}\Big)}_{=\ 0 \text{ for } B(l)} + O(h^{L+1}). \end{aligned}$$

Analoguously we obtain

$$\mathcal{T}_i = \tilde{u}(c_i h) - \bar{p} - h\sum_{j=1}^{s} a_{ij} f(\tilde{u}(c_j h)) =$$

$$
\begin{aligned}
&= \tilde{u}(c_i h) - \tilde{u}(0) - h\sum_{j=1}^{s} a_{ij}\tilde{u}'(c_j h) \;= \\
&= \sum_{l=0}^{L} \frac{(c_i h)^l}{l!}\tilde{u}^{(l)}(0) - \tilde{u}(0) - h\sum_{j=1}^{s} a_{ij}\sum_{l=0}^{L-1}\frac{(c_j h)^l}{l!}\tilde{u}^{(l-1)}(0) + O(h^{L+1}) \;= \\
&= \sum_{l=1}^{L} \frac{h^l}{(l-1)!}\tilde{u}^{(l)}(0)\underbrace{\Big(\frac{c_i^l}{l} - \sum_{j=1}^{s} a_{ij}c_j^{l-1}\Big)}_{=\,0 \text{ for } C(l)} + O(h^{L+1}).
\end{aligned}
$$

Now the simplifying conditions $B(\cdot)$, $C(\cdot)$ yield the desired proposition. Compare section B.1 in the appendix. □

We estimate with the help of (5.22)

$$
\begin{aligned}
\|\chi\|_2 &= \|p - \bar{p} - h\sum_{i=1}^{s} b_i f(P_i)\|_2 \;= \\
&= \Big\|p - \tilde{u}(h) + h\sum_{i=1}^{s} b_i\Big(f(\tilde{u}(c_i h)) - f(P_i)\Big) + \tau\Big\|_2 \;\leq \\
&\leq \|p - \tilde{u}(h)\|_2 + h\|b\|_1\Big\|\Big(f(\tilde{u}(c_i h)) - f(P_i)\Big)_{i=1}^{s}\Big\|_\circ + \|\tau\|_2 \;\leq \\
&\leq L_\phi\|x - \tilde{x}(h)\|_2 + h\underbrace{\|b\|_1 L_{f|_{\mathcal{M}}} L_\phi}_{=\mathcal{C}}\Big\|\Big(\tilde{x}(c_i h) - X_i\Big)_{i=1}^{s}\Big\|_\circ + \\
&\quad + \mathcal{C}h^{p_o+1}.
\end{aligned} \tag{5.25}
$$

Furthermore we estimate with the help of (5.23)

$$
\begin{aligned}
\|\Xi_i\|_2 &= \|P_i - \bar{p} - h\sum_{j=1}^{s} a_{ij} f(P_j)\|_2 \;= \\
&= \Big\|P_i - \tilde{u}(c_i h) + h\sum_{j=1}^{s} a_{ij}\Big(f(\tilde{u}(c_j h)) - f(P_j)\Big) + \mathcal{T}_i\Big\|_2 \;\leq \\
&\leq \underbrace{(1 + h\|A\|_\infty L_{f|_{\mathcal{M}}})L_\phi}_{\leq \mathcal{C}}\Big\|\Big(\tilde{x}(c_j h) - X_j\Big)_{j=1}^{s}\Big\|_\circ + \mathcal{C}h^{q_o+1}
\end{aligned} \tag{5.26}
$$

The next goal is to bound $\|x - \tilde{x}(h)\|_2$ and $\|\tilde{x}(c_i h) - X_i\|_2$. We require the following proposition.

Proposition 5.1.4 *Let the assumptions of proposition 4.3.1 be fulfilled. Then there exist mderate real constants $\mathcal{C} \geq 0$ such that there holds*

$$
\|D\|_\circ \;\leq\; \mathcal{C}\Phi_I(-\frac{h}{\varepsilon\mathcal{K}_1})\Big(\|\bar{d}\|_2 + \|\Xi\|_\circ\Big), \tag{5.27}
$$

$$
\|h\Gamma D\|_\circ \;\leq\; \mathcal{C}\Big(\|\bar{d}\|_2 + \|\Xi\|_\circ\Big), \tag{5.28}
$$

$$
\|\hat{\Omega}\|_\circ, \|\hat{\Theta}\|_2 \;\leq\; \mathcal{C}\max\{h, \|\bar{d}\|_2\}. \tag{5.29}
$$

The constants $\mathcal{C}$ can be chosen as stated in the proof below.

Proof: First we estimate

$$\|\Phi_i - h\sum_{j=1}^{s} a_{ij}\Psi_j\|_2 \;=\; \|(\langle \Xi_i, \bar{\pi}_m\rangle)_{m=1}^{k}\|_2 \;\leq\; \|\Xi_i\|_2$$

for $i = 1, \ldots, s$, which yields

$$\begin{aligned}\|D\|_\circ \;&=\; \Big\|\Sigma^{-1}\Big(e\otimes\bar{d} - (\Phi - h(A\otimes I_k)\Psi)\Big)\Big\|_\circ \;\leq\\ &\leq\; \underbrace{2\sqrt{s}\mathcal{K}_0}_{=\mathcal{C}}\,\Phi_I(-\frac{h}{\varepsilon\mathcal{K}_1})\Big(\|\bar{d}\|_2 + \|\Xi\|_\circ\Big).\end{aligned}$$

Analoguously there follows

$$\begin{aligned}\|h\Gamma D\|_\circ \;&=\; \Big\|\Omega^{-1}(A\otimes I_k)^{-1}\Big(\Omega D - e\otimes\bar{d} + \Phi - h(A\otimes I_k)\Psi\Big)\Big\|_\circ \leq\\ &\leq\; 2\|A^{-1}\|_\infty\Big(2\sqrt{s}\mathcal{K}_0\Phi_I(-\frac{h}{\varepsilon\mathcal{K}_1})(\|\bar{d}\|_2 + \|\Xi\|_\circ) +\\ &\quad + \|\bar{d}\|_2 + \|\Xi\|_\circ\Big) \;\leq\; \mathcal{C}(\|\bar{d}\|_2 + \|\Xi\|_\circ),\end{aligned}$$

where the generic constant

$$\mathcal{C} \;=\; \sup_{\frac{h}{\varepsilon}>0} 2\|A\|_\infty(2\sqrt{s}\mathcal{K}_0\Phi_I(-\frac{h}{\varepsilon\mathcal{K}_1}) + 1)$$

is moderate due to the special structure of the BSI-stability function. Since $\|(X,D)\|_\diamond \leq C_\diamond \max\{h, \|\bar{d}\|_2\}$ there holds

$$\|\hat{\Omega}\|_\circ, \|\hat{\Theta}\|_2 \;\leq\; \underbrace{2L_\Theta \max\{C_\diamond, 1\}}_{=\mathcal{C}} \max\{h, \|\bar{d}\|_2\},$$

which completes the proof. □

Now the estimate

$$\|\tilde{x}(c_i h) - h\sum_{j=1}^{s} a_{ij}\hat{\psi}(\tilde{x}(c_j h))\|_2 \;\leq\; \|\mathcal{T}_i\|_2$$

leads to

$$\begin{aligned}\|X_i - \tilde{x}(c_i h)\|_2 \;&=\; \Big\|h\sum_{j=1}^{s} a_{ij}\hat{\Psi}_j - \hat{\Theta}_i D_i + h\sum_{j=1}^{s} a_{ij}\hat{\Theta}_j\Lambda_j D_j - \tilde{x}(c_i h)\Big\|_2 =\\ &=\; \Big\|h\sum_{j=1}^{s} a_{ij}\Big(\hat{\Psi}_j - \hat{\psi}(\tilde{x}(c_j h))\Big) -\\ &\quad - \Big(\tilde{x}(c_i h) - h\sum_{j=1}^{s} a_{ij}\hat{\psi}(\tilde{x}(c_j h))\Big) - \hat{\Theta}_i D_i + h\sum_{j=1}^{s} a_{ij}\hat{\Theta}_j\Lambda_j D_j\Big\|_2 \leq\\ &\leq\; h\|A\|_\infty L_\psi\|(X_j - \tilde{x}(c_j h))_{j=1}^{s}\|_\circ +\\ &\quad + \|\mathcal{T}\|_\circ + \|\hat{\Omega}\|_\circ\Big(\|D\|_\circ + \|A\|_\infty\|h\Gamma D\|_\circ\Big),\end{aligned}$$

where $\mathcal{T} := (\mathcal{T}_1, \ldots, \mathcal{T}_s)^\top$. The stepsize restriction

$$h \leq \frac{1}{2\|A\|_\infty L_\psi} \tag{5.30}$$

together with proposition 5.1.4 implies

$$\begin{aligned}
\|X_i - \tilde{x}(c_i h)\|_2 &\leq 2\Big(\|\mathcal{T}\|_\circ + \|\hat{\Omega}\|_\circ(\|D\|_\circ + \|A\|_\infty \|h\Gamma D\|_\circ)\Big) \leq \\
&\leq 2\mathcal{C}h^{q_o+1} + 2\Big(\mathcal{C}\Phi_I(-\frac{h}{\varepsilon\mathcal{K}_1})(\|\bar{d}\|_2 + \|\Xi\|_\circ) + \\
&\quad + \|A\|_\infty \mathcal{C}(\|\bar{d}\|_2 + \|\Xi\|_\circ)\Big)\mathcal{C}\max\{h, \|\bar{d}\|_2\} \leq \\
&\leq \mathcal{C}h^{q_o+1} + \mathcal{C}(\|\bar{d}\|_2 + \|\Xi\|_\circ)\max\{h, \|\bar{d}\|_2\}.
\end{aligned} \tag{5.31}$$

The estimate

$$\|\tilde{x}(h) - h\sum_{i=1}^{s} b_i \hat{\psi}(\tilde{x}(c_i h))\|_2 \leq \|\tau\|_2$$

leads to

$$\begin{aligned}
\|x - \tilde{x}(h)\|_2 &= \|h\sum_{i=1}^{s} b_i \hat{\Psi}_i - \hat{\Theta}d + h\sum_{i=1}^{s} b_i \hat{\Theta}_i \Lambda_i D_i - \tilde{x}(h)\|_2 = \\
&= \Big\|h\sum_{i=1}^{s} b_i\Big(\hat{\Psi}_i - \hat{\psi}(\tilde{x}(c_i h))\Big) - \\
&\quad - \Big(\tilde{x}(h) - h\sum_{i=1}^{s} b_i \hat{\psi}(\tilde{x}(c_i h))\Big) - \hat{\Theta}d + h\sum_{i=1}^{s} b_i \hat{\Theta}_i \Lambda_i D_i\Big\|_2 \leq \\
&\leq h\|b\|_1 L_\psi \Big\|\Big(X_i - \tilde{x}(c_i h)\Big)_{i=1}^{s}\Big\|_\circ + \mathcal{C}h^{p_o+1} + \\
&\quad + \|\hat{\Theta}\|_2\|d\|_2 + \|b\|_1\|\hat{\Omega}\|_\circ\|h\Gamma D\|_\circ,
\end{aligned}$$

which together with proposition 5.1.4 and (5.31) implies

$$\begin{aligned}
\|x - \tilde{x}(h)\|_2 &\leq h\|b\|_1 L_\psi\Big(\mathcal{C}h^{q_o+1} + \mathcal{C}(\|\bar{d}\|_2 + \|\Xi\|_\circ)\max\{h, \|\bar{d}\|_2\}\Big) + \mathcal{C}h^{p_o+1} + \\
&\quad + \|d\|_2 \mathcal{C}\max\{h, \|\bar{d}\|_2\} + \|b\|_1 \mathcal{C}(\|\bar{d}\|_2 + \|\Xi\|_\circ)\max\{h, \|\bar{d}\|_2\} \leq \\
&\leq \mathcal{C}(\|d\|_2 + \|\bar{d}\|_2 + \|\Xi\|_\circ)\max\{h, \|\bar{d}\|_2\} + \\
&\quad + \mathcal{C}(h^{q_o+1} + h^{p_o+1}).
\end{aligned} \tag{5.32}$$

Insertion of (5.31) into (5.26) yields

$$\begin{aligned}
\|\Xi_i\|_2 &\leq \mathcal{C}\Big\|\Big(\tilde{x}(c_j h) - X_j\Big)_{j=1}^{s}\Big\|_\circ + \mathcal{C}h^{q_o+1} \leq \\
&\leq \mathcal{C}\Big(\mathcal{C}h^{q_o+1} + \mathcal{C}(\|\bar{d}\|_2 + \|\Xi\|_\circ)\max\{h, \|\bar{d}\|_2\}\Big) + \mathcal{C}h^{q_o+1} = \\
&= \mathcal{C}h^{q_o+1} + R_\Xi\Big(\|\bar{d}\|_2 + \|\Xi\|_\circ\Big)\max\{h, \|\bar{d}\|_2\},
\end{aligned} \tag{5.33}$$

where R_Ξ is defined via the previous equation. The restriction

$$h, \|\bar{d}\|_2 \leq \frac{1}{2R_\Xi} \tag{5.34}$$

implies

$$\begin{aligned} \|\Xi\|_\circ &\leq 2\Big(\mathcal{C}h^{q_o+1} + R_\Xi \|\bar{d}\|_2 \max\{h, \|\bar{d}\|_2\}\Big) \leq \\ &\leq \mathcal{C}\Big(h^{q_o+1} + \|\bar{d}\|_2 \max\{h, \|\bar{d}\|_2\}\Big). \end{aligned} \tag{5.35}$$

Analoguously we insert (5.31), (5.32) into (5.25) and obtain

$$\begin{aligned} \|\chi\|_2 &\leq L_\phi \|x - \tilde{x}(h)\|_2 + \mathcal{C}h \Big\| \Big(\tilde{x}(c_i h) - X_i\Big)_{i=1}^s \Big\|_\circ + \mathcal{C}h^{p_o+1} \leq \\ &\leq L_\phi\Big(\mathcal{C}(h^{q_o+2} + h^{p_o+1}) + \mathcal{C}(\|d\|_2 + \|\bar{d}\|_2 + \|\Xi\|_\circ) \max\{h, \|\bar{d}\|_2\}\Big) + \\ &\quad + h\mathcal{C}\Big(\mathcal{C}h^{q_o+1} + \mathcal{C}(\|\bar{d}\|_2 + \|\Xi\|_\circ) \max\{h, \|\bar{d}\|_2\}\Big) + \mathcal{C}h^{p_o+1} \leq \\ &\leq \mathcal{C}(\|d\|_2 + \|\bar{d}\|_2 + \|\Xi\|_\circ) \max\{h, \|\bar{d}\|_2\} + \mathcal{C}(h^{p_o+1} + h^{q_o+2}). \end{aligned}$$

Now we apply (5.35) which leads to

$$\begin{aligned} \|\chi\|_2 &\leq \mathcal{C}\Big(\|d\|_2 + \|\bar{d}\|_2 + \mathcal{C}(h^{q_o+1} + \|\bar{d}\|_2 \max\{h, \|\bar{d}\|_2\})\Big) \max\{h, \|\bar{d}\|_2\} + \\ &\quad + \mathcal{C}(h^{p_o+1} + h^{q_o+2}) \leq \\ &\leq \mathcal{C}\Big(h^{p_o+1} + h^{q_o+2} + (h^{q_o+1} + \|d\|_2 + \|\bar{d}\|_2) \max\{h, \|\bar{d}\|_2\}\Big). \end{aligned} \tag{5.36}$$

Insertion into (5.20) yields

$$\begin{aligned} \|d\|_V &\leq \Phi_B(-\frac{h}{\varepsilon\mathcal{K}_1})\Big(\|\bar{d}\|_{\bar{V}} + \|\bar{d}\|_2 \mathcal{C} \max\{h, \|\bar{d}\|_2\}\Big) + \\ &\quad + \mathcal{C}\Phi_{BS}(-\frac{h}{\varepsilon\mathcal{K}_1})\Big(\|\Xi^*\|_\circ + \|\chi^*\|_2\Big) \leq \\ &\leq \Phi_B(-\frac{h}{\varepsilon\mathcal{K}_1})\Big(\|\bar{d}\|_{\bar{V}} + \|\bar{d}\|_2 \mathcal{C} \max\{h, \|\bar{d}\|_2\}\Big) + \\ &\quad + \mathcal{C}\Phi_{BS}(-\frac{h}{\varepsilon\mathcal{K}_1})\Big(\mathcal{C}(h^{p_o+1} + h^{q_o+2} + (h^{q_o+1} + \|d\|_2 + \|\bar{d}\|_2) \max\{h, \|\bar{d}\|_2\}) + \\ &\quad + \mathcal{C}(h^{q_o+1} + \|\bar{d}\|_2 \max\{h, \|\bar{d}\|_2\})\Big) \leq \\ &\leq \Phi_B(-\frac{h}{\varepsilon\mathcal{K}_1})\Big(\|\bar{d}\|_{\bar{V}} + \|\bar{d}\|_2 \mathcal{C} \max\{h, \|\bar{d}\|_2\}\Big) + \\ &\quad + \Phi_{BS}(-\frac{h}{\varepsilon\mathcal{K}_1})\mathcal{C}\Big((\|d\|_V + \|\bar{d}\|_2) \max\{h, \|\bar{d}\|_2\} + h^{q_o+1}\Big), \end{aligned}$$

where we applied $\|d\|_2 \leq M_{V^{-1}}^{\frac{1}{2}} \|d\|_V$. If the restriction

$$h, \|\bar{d}\|_2 \leq (2\mathcal{C}\Phi_{BS}(-\frac{h}{\varepsilon\mathcal{K}_1}))^{-1} \tag{5.37}$$

is fulfilled then there holds

$$\frac{1}{1-\mathcal{C}\Phi_{BS}(-\frac{h}{\varepsilon\mathcal{K}_1})\max\{h,\|\bar{d}\|_2\}} \;\leq\; 1+2\mathcal{C}\Phi_{BS}(-\frac{h}{\varepsilon\mathcal{K}_1})\max\{h,\|\bar{d}\|_2\},$$

which implies

$$\begin{aligned}
\|d\|_V \;&\leq\; \Big(1+2\mathcal{C}\Phi_{BS}(-\frac{h}{\varepsilon\mathcal{K}_1})\max\{h,\|\bar{d}\|_2\}\Big)\Big(\Phi_B(-\frac{h}{\varepsilon\mathcal{K}_1})(\|\bar{d}\|_{\bar{V}}+\\
&\quad +\|\bar{d}\|_2\mathcal{C}\max\{h,\|\bar{d}\|_2\})+\Phi_{BS}(-\frac{h}{\varepsilon\mathcal{K}_1})\mathcal{C}(\|\bar{d}\|_2\max\{h,\|\bar{d}\|_2\}+h^{q_o+1})\Big)\;\leq\\
&\leq\; \Phi_B(-\frac{h}{\varepsilon\mathcal{K}_1})\Big(\|\bar{d}\|_{\bar{V}}+\|\bar{d}\|_2\mathcal{C}\max\{h,\|\bar{d}\|_2\}\Big)+\\
&\quad +\Phi_{BS}(-\frac{h}{\varepsilon\mathcal{K}_1})\Big(\mathcal{C}\|\bar{d}\|_2\max\{h,\|\bar{d}\|_2\}+\mathcal{C}h^{q_o+1}\Big). \qquad (5.38)
\end{aligned}$$

In (5.38) we made use of the estimate $\|\bar{d}\|_{\bar{V}} \leq M_V^{\frac{1}{2}}\|\bar{d}\|_2$.
For further approximations the concrete structure of the stability functions is essential.

5.1.1 The Stiff Error Component for the Method Radau IIa

In the case of the method Radau IIa the B-stability function is of the form

$$\Phi_B(-\frac{h}{\varepsilon\mathcal{K}_1}) \;=\; \frac{1}{\sqrt{1+\phi_B\frac{h}{\varepsilon\mathcal{K}_1}}} \qquad (5.39)$$

(compare section B.2.3). Due to the special structure of the method Radau IIa the BS-stability function can be replaced by the BSI-stability function, such that we have

$$\Phi_{BS}(-\frac{1}{\varepsilon\mathcal{K}_1}) \;=\; \Phi_I(-\frac{1}{\varepsilon\mathcal{K}_1}) \;=\; \frac{\mathcal{C}}{1+\phi_I\frac{h}{\varepsilon\mathcal{K}_1}} \qquad (5.40)$$

(compare section B.2.1,B.2.2).
In the following considerations we use the index notation, which means that we write d_ν and V_ν for $d(\eta_\nu)$ resp. $V(\eta_\nu)$. We want to derive error bounds by induction. First we assume that for $\bar{d} := d_{\nu-1}$ and h the assumptions of proposition 4.3.3 and (5.15), (5.17), (5.30), (5.34) and (5.37) are fulfilled. This allows us to utilize the estimate (5.38) with the special B- and BS-stability functions (5.39), (5.40). In order to derive estimates the square root in the B-stability function enforces us to apply (5.38) two times at once, i.e. we have to consider two steps of the method Radau IIa at once. Therefore we have to ensure that the second step of the method is well defined. This means that now d_ν has to satisfy the same assumptions as required for $\bar{d} = d_{\nu-1}$. Since we know that there holds

$$\|d_\nu\|_2 \;\leq\; C_*\max\{h,\|d_{\nu-1}\|_2\}$$

these assumptions can be reformulated as restrictions for h, $\|d_{\nu-1}\|_2$. Therefore we set

$$R_d \;:=\; \max\left\{C_\diamond C^\diamond, C_1\max\{1,C_3\}, 2R_{\mho}, 2R_{\mho\Gamma\mho^{-1}}, 2R_\Xi, 2\mathcal{C}\Phi_{BS}(-\frac{h}{\varepsilon\mathcal{K}_1})^{-1}\right\}$$

and assume

$$h, \|d_{\nu-1}\|_2 \leq \frac{1}{C_* R_d}. \tag{5.41}$$

Now the second step is well defined and we can apply (5.38) two times, which implies

$$\begin{aligned}
\|d_{\nu+1}\|_{V_\nu} &\leq \frac{1}{\sqrt{1+\phi_B \frac{h}{\varepsilon \mathcal{K}_1}}}\Big(\|d_\nu\|_{V_\nu} + \|d_\nu\|_{V_\nu} M_{V^{-1}}^{\frac{1}{2}} \mathcal{C} \max\{h, \|d_\nu\|_2\}\Big) + \\
&\quad + \frac{\mathcal{C}}{1+\phi_I \frac{h}{\varepsilon \mathcal{K}_1}}\Big(\|d_\nu\|_2 \max\{h, \|d_\nu\|_2\} + \mathcal{C} h^{q_o+1}\Big) \leq \\
&\leq \frac{1}{1+\phi_{\min} \frac{h}{\varepsilon \mathcal{K}_1}}\Big(\|d_{\nu-1}\|_{V_{\nu-1}} + \|d_{\nu-1}\|_2 \mathcal{C} \max\{h, \|d_{\nu-1}\|_2\} + \mathcal{C} h^{q_o+1}\Big) \leq \\
&\leq \frac{1}{1+\phi_{\min} \frac{h}{\varepsilon \mathcal{K}_1}}\Big(\|d_{\nu-1}\|_{V_{\nu-1}} + \\
&\quad + \mathcal{C}(\|d_{\nu-1}\|_{V_{\nu-1}} \max\{h, \|d_{\nu-1}\|_{V_{\nu-1}}\} + h^{q_o+1})\Big),
\end{aligned} \tag{5.42}$$

where $\phi_{\min} := \min\{\phi_I, \phi_B\}$. This is the error recursion suiteable for the error induction. We make the induction over all even idices, i.e. $\nu = 0, 2, 4, \ldots$. Then the bounds for the odd indices will follow directly from the bounds for the even indices.

Induction assumption: For $K_0 \in [K_{\min}, \frac{1}{\varepsilon}]$ where $K_{\min} := 2\phi_{\min}^{-1} \mathcal{K}_1 \mathcal{C}$ with $\mathcal{C}$ from (5.42) there holds

$$\|d_{\nu-1}\|_{V_{\nu-1}} \leq K_0 \varepsilon h^s \tag{5.43}$$

for $\nu - 1$ even and $h \leq h_{\max}$, where

$$h_{\max} := \min\left\{\frac{1}{C_* R_d M_{V^{-1}}^{\frac{1}{2}}}, \frac{1}{2\|A\|_\infty L_\psi}\right\}. \tag{5.44}$$

Since $\|d_{\nu-1}\|_{V_{\nu-1}} \leq K_0 \varepsilon h^s \leq h$ and $\|d_{\nu-1}\|_2 \leq M_{V^{-1}}^{\frac{1}{2}} \|d_{\nu-1}\|_{V_{\nu-1}}$ the assumptions (5.43), (5.44) imply that (5.41) is fulfilled. Now (5.42) can be used to derive

Induction conclusion:

$$\begin{aligned}
\|d_{\nu+1}\|_{V_{\nu+1}} &\leq \frac{K_0 \varepsilon h^s}{1+\phi_{\min} \frac{h}{\varepsilon \mathcal{K}_1}}\Big(1 + \mathcal{C}(\max\{h, \varepsilon K_0 h^{q_o}\} + \frac{1}{K_0}\frac{h}{\varepsilon})\Big) \leq \\
&\leq \frac{K_0 \varepsilon h^s}{1+\phi_{\min} \frac{h}{\varepsilon \mathcal{K}_1}}\Big(1 + \phi_{\min} \frac{h}{\varepsilon \mathcal{K}_1} \phi_{\min}^{-1} \mathcal{K}_1 \mathcal{C}(\varepsilon + \frac{1}{K_0})\Big) \leq \\
&\leq K_0 \varepsilon h^s
\end{aligned} \tag{5.45}$$

for $h \leq h_{\max}$.

In order to derive bounds for ν odd we use (5.38) together with (5.40) and obtain

$$\|d_\nu\|_{V_\nu} \leq \Phi_B(-\frac{h}{\varepsilon \mathcal{K}_1})\Big(\|d_{\nu-1}\|_{V_{\nu-1}} + \|d_{\nu-1}\|_2 \mathcal{C} \max\{h, \|d_{\nu-1}\|_2\}\Big) +$$

$$+\frac{\mathcal{C}}{1+\phi_I\frac{h}{\varepsilon\mathcal{K}_1}}\Big(\mathcal{C}\|d_{\nu-1}\|_2\max\{h,\|d_{\nu-1}\|_2\}+\mathcal{C}h^{s+1}\Big)\ \leq$$

$$\leq\ \mathcal{C}\|d_{\nu-1}\|_{V_{\nu-1}}+\frac{\mathcal{C}}{1+\phi_I\frac{h}{\varepsilon\mathcal{K}_1}}h^{s+1}\ \leq$$

$$\leq\ \mathcal{C}K_0\varepsilon h^s+\frac{\mathcal{C}}{1+\phi_I\frac{h}{\varepsilon\mathcal{K}_1}}\frac{h}{\varepsilon K_0}K_0\varepsilon h^s\ \leq$$

$$\leq\ (\mathcal{C}+\frac{\mathcal{C}}{K_0})K_0\varepsilon h^s\ =:\ K_1\varepsilon h^s \tag{5.46}$$

for ν odd and $h\leq h_{\max}$. Now we can formulate the following proposition:

Proposition 5.1.5 (Stiff error component for Radau IIa) *Let $K_0\in[K_{\min},\frac{1}{\varepsilon}]$ and*

$$\|d(\eta_0)\|_{V(\eta_0)}\ \leq\ K_0\varepsilon h^s \tag{5.47}$$

for $h\leq h_{\max}$. Then there holds

$$\|d(\eta_\nu)\|_{V(\eta_\nu)}\ \leq\ \begin{cases} K_0\varepsilon h^s & \text{for } \nu \text{ even}\\ K_1\varepsilon h^s & \text{for } \nu \text{ odd}\end{cases} \tag{5.48}$$

for all $\nu\in\mathbb{N}_0$ with $t_\nu\in[0,T]$. The constants $K_{\min},h_{\max},K_1$ have to be chosen like in the considerations above.

Remark: Prposition 5.1.5 can be formulated in the 2-norm: Let

$$\|d_0\|_2\ \leq\ B_0\varepsilon h^s,$$

for all $h\leq h_{\max}$, such that $K_0:=B_0M_{V^{-1}}^{\frac{1}{2}}\in[K_{\min},\frac{1}{\varepsilon}]$. Then there holds

$$\|d_\nu\|_2\ \leq\ \begin{cases}\mathcal{K}_0B_0\varepsilon h^s & \text{for } \nu \text{ even}\\ \mathcal{K}_0B_1\varepsilon h^s & \text{for } \nu \text{ odd}\end{cases}$$

for all $\nu\in\mathbb{N}_0$ with $t_\nu\in[0,T]$, where the constant B_1 is defined as $B_1:=(\mathcal{C}+\frac{\mathcal{C}}{K_0})B_0$ with $\mathcal{C}$ from (5.46).

5.1.2 The Stiff Error Component for the Methods Radau Ia and Gauss

Now let $q_o\geq 2$. Under the assumption

$$\|d_{\nu-1}\|_{V_{\nu-1}}\leq K_{\nu-1}h^{q_o} \tag{5.49}$$

for all

$$h\leq\frac{1}{K_{\nu-1}}$$

the estimate (5.38) together with $\Phi_B(-\frac{h}{\varepsilon\mathcal{K}_1})\leq 1$ and $\Phi_{BS}(-\frac{h}{\varepsilon\mathcal{K}_1})\leq\mathcal{C}$ implies

$$\begin{aligned}\|d_\nu\|_{V_\nu}\ &\leq\ (1+\mathcal{C}\max\{h,\|d_{\nu-1}\|_{V_{\nu-1}}\})\|d_{\nu-1}\|_{V_{\nu-1}}+\mathcal{C}h^{q_o+1}\leq\\ &\leq\ (1+\mathcal{C}h\max\{1,K_{\nu-1}h^{q_o-1}\})K_{\nu-1}h^{q_o}+\mathcal{C}h^{q_o}\ \leq\\ &\leq\ \Big((1+\mathcal{C}h)K_{\nu-1}+\mathcal{C}h\Big)h^{q_o}\leq K_\nu h^{q_o},\end{aligned} \tag{5.50}$$

where

$$K_\nu := (1+\mathcal{C}h)K_{\nu-1} + \mathcal{C}h.$$

The recursive definition of K_ν implies

$$\begin{aligned} K_\nu &\leq (1+\mathcal{C}h)^\nu K_0 + \frac{(1+\mathcal{C}h)^\nu - 1}{1+\mathcal{C}h-1}\mathcal{C}h \leq \\ &\leq e^{\mathcal{C}T}K_0 + e^{\mathcal{C}T} - 1 =: K_\infty. \end{aligned} \tag{5.51}$$

Now if we assume

$$h \leq \frac{1}{K_\infty} \tag{5.52}$$

then there holds

$$h \leq \frac{1}{K_\nu}$$

for all $\nu \in \mathbb{N}_0$, which means that we can conclude from (5.49) to (5.50) for all $\nu \in \mathbb{N}$. Since for $h \leq h_{\max}$ with

$$h_{\max} := \min\left\{\frac{1}{C_* R_d M_{V-1}^{\frac{1}{2}}}, \frac{1}{2\|A\|_\infty L_\psi}, \frac{1}{K_\infty M_{V-1}^{\frac{1}{2}}}\right\}$$

all necessary restrictions on $h, \|d_{\nu-1}\|_2$ are fulfilled we obtain

$$\|d_\nu\|_{V_\eta} \leq K_\infty h^{q_o}$$

for all $\nu \in \mathbb{N}$ and $h \leq h_{\max}$.

Proposition 5.1.6 (Stiff error component for Radau Ia and Gauss) *Let $q_o \geq 2$ and $K_0 > 0$ with*

$$\|d(\eta_0)\|_{V(\eta_0)} \leq K_0 h^{q_o} \tag{5.53}$$

for $h \leq h_{\max}$. Then for the methods Radau Ia and Gauss there holds

$$\|d(\eta_\nu)\|_{V(\eta_\nu)} \leq K_\infty h^{q_o} \tag{5.54}$$

for all $\nu \in \mathbb{N}_0$ with $t_\nu \in [0,T]$. The constants K_∞, $h_{\max}$ have to be chosen as stated in the considerations above.

5.2 Estimates for the Smooth Error Component

Let $q_o \geq 2$.[2)] In order to derive bounds for the smooth error components we use proposition 5.1.5 in the case of the method Radau IIa and proposition 5.1.6 in the case of the methods Radau Ia and Gauss. This means that we can use the estimates

$$\|\bar{d}\|_{\bar{V}}, \|d\|_V \leq \begin{cases} \mathcal{C}\varepsilon h^{q_o} & \text{Radau IIa} \\ \mathcal{C}h^{q_o} & \text{Radau Ia and Gauss} \end{cases} \tag{5.55}$$

[2)] For $q_o = 1$ the method Radau IIa reduces to the implicit euler method, compare chapter 3.

to bound

$$\begin{aligned}
\|\Xi\|_\circ &\leq \mathcal{C}(h^{q_o+1} + \|\bar{d}\|_2 \max\{h, \|\bar{d}\|_2\}) \leq \\
&\leq \mathcal{C}(h^{q_o+1} + \left\{ \begin{array}{c} \mathcal{C}\varepsilon h^{q_o} \\ \mathcal{C}h^{q_o} \end{array} \right\} \max\{h, \left\{ \begin{array}{c} \mathcal{C}\varepsilon h^{q_o} \\ \mathcal{C}h^{q_o} \end{array} \right\}\}) \leq \\
&\leq \mathcal{C}h^{q_o+1} \quad \left\{ \begin{array}{l} \text{for Radau IIa} \\ \text{for Radau Ia and Gauss.} \end{array} \right. \qquad (5.56)
\end{aligned}$$

Now we estimate

$$\begin{aligned}
\|x - \tilde{x}(h)\|_2 &\leq \mathcal{C}(\|d\|_2 + \|\bar{d}\|_2 + \|\Xi\|_\circ) \max\{h, \|\bar{d}\|_2\} + \\
&\quad + \mathcal{C}(h^{q_o+2} + h^{p_o+1}) \leq \\
&\leq \mathcal{C}(\left\{ \begin{array}{c} \mathcal{C}\varepsilon h^{q_o} \\ \mathcal{C}h^{q_o} \end{array} \right\} + \left\{ \begin{array}{c} \mathcal{C}\varepsilon h^{q_o} \\ \mathcal{C}h^{q_o} \end{array} \right\} + \mathcal{C}h^{q_o+1}) \max\{h, \left\{ \begin{array}{c} \mathcal{C}\varepsilon h^{q_o} \\ \mathcal{C}h^{q_o} \end{array} \right\}\} + \\
&\quad + \mathcal{C}(h^{q_o+2} + h^{p_o+1}) \leq \\
&\leq \left\{ \begin{array}{ll} \mathcal{C}(h+\varepsilon)h^{q_o+1} & \text{for Radau IIa} \\ \mathcal{C}h^{q_o+1} & \text{for Radau Ia and Gauss} \end{array} \right. \qquad (5.57)
\end{aligned}$$

where we used $q_o + 2 \leq p_o + 1$ for $q_o \geq 2$. This implies

$$\begin{aligned}
\|p - \tilde{u}(h)\|_2 &\leq L_\phi \|x - \tilde{x}(h)\|_2 \leq L_\phi \left\{ \begin{array}{c} \mathcal{C}(h+\varepsilon)h^{q_o+1} \\ \mathcal{C}h^{q_o+1} \end{array} \right\} \\
&\leq \left\{ \begin{array}{ll} \mathcal{C}(h+\varepsilon)h^{q_o+1} & \text{for Radau IIa} \\ \mathcal{C}h^{q_o+1} & \text{for Radau Ia and Gauss.} \end{array} \right. \qquad (5.58)
\end{aligned}$$

In the followning estimates we use the index notation, i.e. in the step from $\bar{\eta} := \eta_{\nu-1}$ to η_ν the notation $p(\eta_\nu)$ for p and $\tilde{u}_{\nu-1}$ for $\tilde{u}$ is used:

$$\begin{aligned}
&\|p(\eta_\nu) - \tilde{u}(t_\nu)\|_2 \leq \\
&\leq \|p(\eta_\nu) - \tilde{u}_{\nu-1}(t_\nu)\|_2 + \|\tilde{u}_{\nu-1}(t_{\nu-1}+h) - u(t_{\nu-1}+h)\|_2 \leq \\
&\leq \left\{ \begin{array}{c} \mathcal{C}(h+\varepsilon)h^{q_o+1} \\ \mathcal{C}h^{q_o+1} \end{array} \right\} + e^{L_{f|\mathcal{M}}h} \|p(\eta_{\nu-1}) - u(t_{\nu-1})\|_2 \leq \\
&\leq e^{L_{f|\mathcal{M}}t_\nu} \|p(\eta_0) - u_0\|_2 + \frac{e^{L_{f|\mathcal{M}}t_\nu} - 1}{e^{L_{f|\mathcal{M}}h} - 1} \left\{ \begin{array}{c} \mathcal{C}(h+\varepsilon)h^{q_o+1} \\ \mathcal{C}h^{q_o+1} \end{array} \right\} \leq \\
&\leq e^{L_{f|\mathcal{M}}t_\nu} \|p(\eta_0) - u_0\|_2 + (e^{L_{f|\mathcal{M}}t_\nu} - 1) \left\{ \begin{array}{ll} \mathcal{C}(h+\varepsilon)h^{q_o} & \text{Radau IIa} \\ \mathcal{C}h^{q_o} & \text{Radau Ia, Gauss.} \end{array} \right. \qquad (5.59)
\end{aligned}$$

Now the assumption

$$\|p(\eta_0) - u_0\|_2 \leq \left\{ \begin{array}{ll} C_0(h+\varepsilon)h^{q_o} & \text{Radau IIa} \\ C_0 h^{q_o} & \text{Radau Ia, Gauss} \end{array} \right. \qquad (5.60)$$

for $h \leq h_{\max}$ implies

$$\begin{aligned} \|p(\eta_\nu) - u(t_\nu)\|_2 &\leq e^{L_{f|\mathcal{M}}T} \left\{ \begin{array}{c} C_0(h+\varepsilon)h^{q_o} \\ C_0 h^{q_o} \end{array} \right\} + (e^{L_{f|\mathcal{M}}T} - 1) \left\{ \begin{array}{c} \mathcal{C}(h+\varepsilon)h^{q_o} \\ \mathcal{C}h^{q_o} \end{array} \right\} \leq \\ &\leq \left\{ \begin{array}{ll} \mathcal{C}(h+\varepsilon)h^{q_o} & \text{Radau IIa} \\ \mathcal{C}h^{q_o} & \text{Radau Ia, Gauss.} \end{array} \right. \end{aligned} \tag{5.61}$$

In the next section we combine the stiff and the smooth error components to obtain convergence estimates for the methods Radau Ia, IIa and Gauss.

5.3 Error Bounds for the Methods Radau Ia, IIa and Gauss

The whole error as a combination of the stiff and the smooth error component can be estimated by

$$\begin{aligned} \|\eta_\nu - u(t_\nu)\|_2 &\leq \|\eta_\nu - p(\eta_\nu)\|_2 + \|p(\eta_\nu) - u(t_\nu)\|_2 \leq \\ &\leq \left\{ \begin{array}{c} \mathcal{C}\varepsilon h^{q_o} \\ \mathcal{C}h^{q_o} \end{array} \right\} + \left\{ \begin{array}{c} \mathcal{C}(h+\varepsilon)h^{q_o} \\ \mathcal{C}h^{q_o} \end{array} \right\} \leq \\ &\leq \left\{ \begin{array}{ll} \mathcal{C}(h+\varepsilon)h^{q_o} & \text{Radau IIa} \\ \mathcal{C}h^{q_o} & \text{Radau Ia, Gauss} \end{array} \right. \end{aligned} \tag{5.62}$$

for all $\nu \in \mathbb{N}_0$ where $t_\nu \in [0, T]$. This leads to the following convergence theorem:

Theorem 5.3.1 (Convergence of the method Radau IIa) *Let B_0, $C_0 \geq 0$, such that $K_0 := B_0 M_{V^{-1}}^{\frac{1}{2}} \in [K_{\min}, \frac{1}{\varepsilon}]$ and*

$$\begin{aligned} \|p(\eta_0) - u_0\|_2 &\leq C_0(h+\varepsilon)h^s, \\ \|\eta_0 - p(\eta_0)\|_2 &\leq B_0 \varepsilon h^s \end{aligned}$$

for all $h \leq h_{\max}$. Then the Radau IIa methods satisfies

$$\|\eta_\nu - u(t_\nu)\|_2 \leq \mathcal{C}(h+\varepsilon)h^s, \tag{5.63}$$

$$\|p(\eta_\nu) - u(t_\nu)\|_2 \leq \mathcal{C}(h+\varepsilon)h^s, \tag{5.64}$$

$$\|\eta_\nu - p(\eta_\nu)\|_2 \leq \left\{ \begin{array}{ll} \mathcal{K}_0 B_0 \varepsilon h^s & \text{for } \nu \text{ even} \\ \mathcal{K}_0 B_1 \varepsilon h^s & \text{for } \nu \text{ odd} \end{array} \right. \tag{5.65}$$

for all $\nu \in \mathbb{N}$ with $t_\nu \in [0, T]$. The appearing constants have to be chosen like in the considerations above.

Remarks: concerning $K_0 \in [K_{\min}, \frac{1}{\varepsilon}]$:

- For $B_0 = K_{\min} M_{V^{-1}}^{-\frac{1}{2}}$ the stiff errorr component of the approximation for the initial value has to satisfy a rather restrictive assumption. The stiff components of the further approximations stay at the same level as the stiff component of the initial approximation, i.e. at the $O(\varepsilon h^s)$-level.

- For $B_0 = \frac{1}{\varepsilon} M_{V^{-1}}^{-\frac{1}{2}}$ there is a rather mild restriction for the stiff component of the initial approximation, which results in a weaker bound for the stiff error components during the integration, i.e. the stiff error component is of $O(h^s)$. For this case the following considerations lead to an improved error result. The assumption

$$\|d_{\nu-1}\|_{V_{\nu-1}} \leq K_{\nu-1}\varepsilon h,$$

where $K_{\min} < K_{\nu-1} \leq \frac{1}{\varepsilon}$ leads to

$$\|d_{\nu+1}\|_{V_{\nu+1}} \leq K_{\nu-1}\varepsilon h^s \frac{1 + \phi_{\min}\frac{h}{\varepsilon\mathcal{K}_1}\frac{K_{\min}}{K_{\nu-1}}}{1 + \phi_{\min}\frac{h}{\varepsilon\mathcal{K}_1}} =: K_{\nu+1}\varepsilon h,$$

where

$$K_{\nu+1} := \frac{1 + \phi_{\min}\frac{h}{\varepsilon\mathcal{K}_1}\frac{K_{\min}}{K_{\nu-1}}}{1 + \phi_{\min}\frac{h}{\varepsilon\mathcal{K}_1}} K_{\nu-1}.$$

There follows $\lim_{\nu\to\infty} K_\nu = K_{\min}$ for $\nu - 1$ even . For odd ν's we obtain $K_\nu = (\mathcal{C} + \frac{\mathcal{C}}{K_{\nu-1}})K_{\nu-1}$ with

$$\|d_\nu\|_{V_\nu} \leq K_\nu \varepsilon h^s$$

and $\lim_{\nu\to\infty} K_\nu = (\mathcal{C} + \frac{\mathcal{C}}{K_{\min}})K_{\min}$ for ν odd. During the integration this damping in the stiff error component leads to a stiff error of $O(\varepsilon h)$.

Theorem 5.3.2 (Convergence of the methods Radau Ia and Gauss) *Let $B_0, C_0 \geq 0$ and $K_0 := B_0 M_{V^{-1}}^{\frac{1}{2}}$ and*

$$\begin{aligned} \|p(\eta_0) - u_0\|_2 &\leq C_0 h^{q_o} \\ \|\eta_0 - p(\eta_0)\|_2 &\leq B_0 h^{q_o} \end{aligned}$$

for all $h \leq h_{\max}$. Then the methods Radau Ia and Gauss satisfy

$$\|\eta_\nu - u(t_\nu)\|_2 \leq \mathcal{C} h^{q_o} \tag{5.66}$$

$$\|p(\eta_\nu) - u(t_\nu)\|_2 \leq \mathcal{C} h^{q_o} \tag{5.67}$$

$$\|\eta_\nu - p(\eta_\nu)\|_2 \leq M_V^{\frac{1}{2}} K_\infty h^{q_o} \tag{5.68}$$

for all $\nu \in \mathbb{N}_0$ with $t_\nu \in [0, T]$. The appearing constants have to be chosen like in the considerations above.

5.4 The Strongly Stiff Case

We consider the methods Radau Ia, IIa and Gauss for $q_o \geq 2$ and stipulate the following smoothness assumptions.

Smoothness assumptions concerning the functions p, d, π_i and $\hat{\psi}$

- Let the functions $p(y), d(y), \pi_i(y)$ be (p_o+1)-times continuously differentiable with

$$\Big\|\frac{\partial^{|l|}p(y)}{\partial y_1^{l_1}\cdots\partial y_n^{l_n}}\Big\|_2,\ \Big\|\frac{\partial^{|l|}d(y)}{\partial y_1^{l_1}\cdots\partial y_1^{l_n}}\Big\|_2,\ \Big\|\frac{\partial^{|l|}\pi_i(y)}{\partial y_1^{l_1}\cdots\partial y_n^{l_n}}\Big\|_2 \le \mathcal{C} \tag{5.69}$$

  for all $y \in \mathcal{G}$, where $l := (l_1, \ldots, l_n) \in \mathbb{N}_0^n$ with

$$1 \le |l| := l_1 + \cdots + l_n \le p_o + 1$$

  and $i = 1, \ldots, k$.

- Let $\hat{\psi}(x)$ be (p_o+1)-times continuously differentiable with

$$\Big\|\frac{\partial^{|l|}\hat{\psi}(x)}{\partial x_1^{l_1}\cdots\partial x_{n-k}^{l_{n-k}}}\Big\|_2 \le \mathcal{C} \tag{5.70}$$

  for all $x \in \mathcal{U}$ where $l = (l_1, \ldots, l_{n-k}) \in \mathbb{N}_0^{n-k}$ and

$$|l| := l_1 + \cdots + l_{n-k} \le p_o + 1.$$

In order to derive estimates the folowing propositions are required.

Proposition 5.4.1 (Angle condition) *Let $\bar{e}$ be a direction-vector in the tangential space $\mathcal{T}(p(\bar{\eta}))$ of the manifold $\mathcal{M}$ corresponding to the point $p(\bar{\eta}), \bar{\eta} \in \mathcal{G}$ with $\|\bar{e}\|_2 = 1$. Then there holds*

$$\|(\langle \bar{e}, \pi_i(\bar{\eta})\rangle)_{i=1}^k\|_2 \le \frac{L_\varphi}{\sqrt{1+L_\varphi^2}} < 1. \tag{5.71}$$

This means that the angle between $\mathcal{L}\{\pi_1(\bar{\eta}), \ldots, \pi_k(\bar{\eta})\}$ and the manifold $\mathcal{M}$ is greater equal α with

$$\cos\alpha = \frac{L_\varphi}{\sqrt{1+L_\varphi^2}}\,.$$

Proof: We estimate

$$L_\varphi \ge \frac{\|(\langle \bar{e}, \bar{\pi}_i\rangle)_{i=1}^k\|_2}{\|(\langle \bar{e}, \hat{\pi}_i)_{i=1}^{n-k}\|_2} = \frac{\|(\langle \bar{e}, \bar{\pi}_i\rangle)_{i=1}^k\|_2}{\sqrt{1-\|(\langle \bar{e}, \bar{\pi}_i\rangle)_{i=1}^k\|_2^2}}$$

where the first inequality follows from

$$\bar{e} \in \mathcal{L}\{\hat{\pi}_1 + \sum_{j=1}^k \frac{\partial}{\partial x_1}\varphi_j(0)\bar{\pi}_j, \ldots, \hat{\pi}_{n-k} + \sum_{j=1}^k \frac{\partial}{\partial x_{n-k}}\varphi_j(0)\bar{\pi}_j\}$$

and the definition of L_φ. A reformulation of this inequality leads to

$$\|(\langle \bar{e}, \pi_i(\bar{\eta})\rangle)_{i=1}^k\|_2 \le \frac{L_\varphi}{\sqrt{1+L_\varphi^2}} < 1$$

which completes the proof. □

Proposition 5.4.2 *Let $\check{\eta}(t)$ be a function which satisfies the differential equation (2.11) at $\check{t}$, then for $\check{p}(t) := p(\check{\eta}(t))$, $\check{d}(t) := d(\check{\eta}(t))$ and $\check{\pi}_i(t) := \pi_i(\check{\eta}(t))$ there holds*

$$\|\check{p}'(\check{t}) - f(\check{p}(\check{t}))\|_2 \leq \mathcal{C}\|\check{d}(\check{t})\|_2 \max_i \|\check{\pi}_i'(\check{t})\|_2. \tag{5.72}$$

The constant $\mathcal{C}$ has to be chosen as defined in (5.73).

Proof: Let $\check{\Lambda}(t) := \Lambda(\check{\eta}(t))$, then the identities

$$\begin{aligned} \check{\eta}(t) &= \check{p}(t) + \sum_{i=1}^{k} \check{d}_i(t)\check{\pi}_i(t), \\ f(\check{\eta}(t)) &= f(\check{p}(t)) + \sum_{i=1}^{k} [\check{\Lambda}(t)\check{d}(t)]_i \check{\pi}_i(t) \end{aligned}$$

and $\check{\eta}'(\check{t}) = f(\check{\eta}(\check{t}))$ imply

$$\underbrace{\check{p}'(\check{t}) - f(\check{p}(\check{t}))}_{\in \mathcal{T}(\check{p}(\check{t}))} = -\sum_{i=1}^{s} \check{d}_i(\check{t})\check{\pi}_i'(\check{t}) - \underbrace{\sum_{i=1}^{s} [(I_k - \check{\Lambda}(\check{t})\check{d}(\check{t})]_i \check{\pi}_i(\check{t}_i)}_{\in \mathcal{L}\{\check{\pi}_1(\check{t})),\ldots,\check{\pi}_k(\check{t}))\}}.$$

We extend $\check{\pi}_1(\check{t}), \ldots, \check{\pi}_k(\check{t})$ to an ONS $\check{\pi}_1(\check{t}), \ldots, \check{\pi}_k(\check{t})$, $\hat{\pi}_i(\check{t}), \ldots, \hat{\pi}_{n-k}(\check{t})$ such that there holds

$$\check{\pi}_i'(\check{t}) = \sum_{j=1}^{k} \langle \check{\pi}_i'(\check{t}), \check{\pi}_j(\check{t})\rangle \check{\pi}_j(\check{t}) + \sum_{j=1}^{n-k} \langle \check{\pi}_i'(\check{t}), \hat{\pi}_j(\check{t})\rangle \hat{\pi}_j(\check{t}).$$

This implies

$$\underbrace{\check{p}'(\check{t}) - f(\check{p}(\check{t}))}_{\in \mathcal{T}(\check{p}(\check{t}))} = \underbrace{-\sum_{j=1}^{n-k} [(\langle \check{\pi}_i'(\check{t}), \hat{\pi}_j(\check{t})\rangle)_{ji} \check{d}(\check{t})]_j \hat{\pi}_j(\check{t})}_{\in \mathcal{L}\{\check{\pi}_1(\check{t})),\ldots,\check{\pi}_k(\check{t}))\}^{\perp}} + \underbrace{\sum_{j=1}^{k} \langle \check{p}'(\check{t}) - f(\check{p}(\check{t})), \check{\pi}_j(\check{t})\rangle \check{\pi}_j(\check{t})}_{\in \mathcal{L}\{\check{\pi}_1(\check{t})),\ldots,\check{\pi}_k(\check{t}))\}}.$$

Now proposition 5.4.1 can be applied to obtain

$$\begin{aligned} \|\check{p}'(\check{t}) - f(\check{p}(\check{t}))\|_2^2 &= \|(\langle \check{\pi}_i'(\check{t}), \hat{\pi}_j(\check{t})\rangle)_{ji} \check{d}(\check{t})\|_2^2 + \|(\langle \check{p}'(\check{t}) - f(\check{p}(\check{t})), \check{\pi}_j(\check{t})\rangle)_{j=1}^{k}\|_2^2 \leq \\ &\leq (\sqrt{k} \max_i \|\check{\pi}_i'(\check{t})\|_2 \|\check{d}(\check{t})\|_2)^2 + (\cos\alpha)^2 \|\check{p}'(\check{t}) - f(\check{p}(\check{t}))\|_2^2. \end{aligned}$$

There follows

$$\|\check{p}'(\check{t}) - f(\check{p}(\check{t}))\|_2 \leq \underbrace{\sqrt{\frac{k}{1 - (\cos\alpha)^2}}}_{=\mathcal{C}} \max_i \|\check{\pi}_i'(\check{t})\|_2 \|\check{d}(\check{t})\|_2. \tag{5.73}$$

□

Notation: The collocation polynomial of the methods Radau Ia, IIa and Gauss is denoted by $\check{\eta}(t)$, which means that there holds

$$\check{\eta}'(c_i h) = f(\check{\eta}(c_i h)), \quad i = 1, \ldots, s.$$

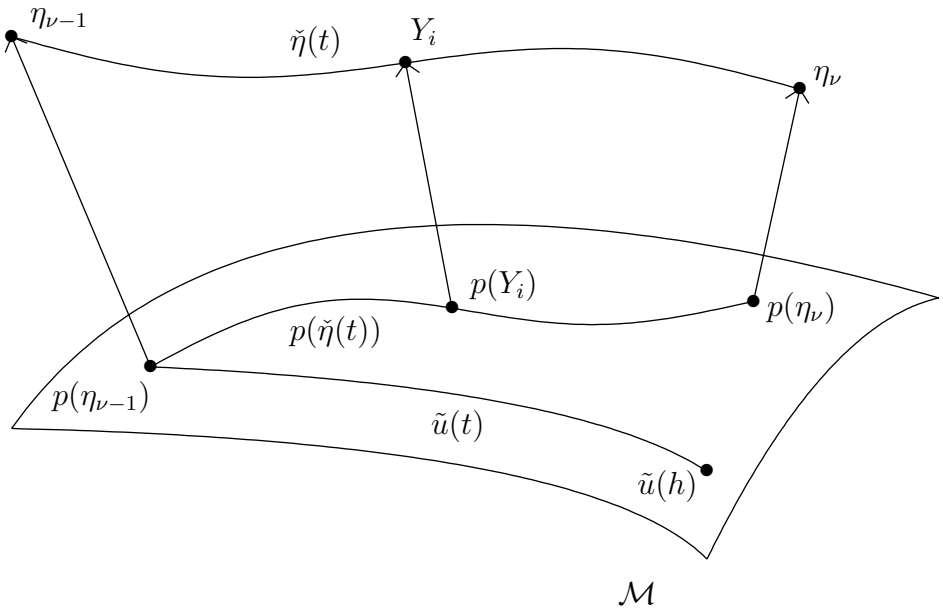


Figure 5.2: The collocation polynomial $\check{\eta}(t)$

Corresponding to the collocation polynomial we define $\check{p}(t) := p(\check{\eta}(t))$ for $\check{\eta}(t) \in \mathcal{G}$. Compare figure 5.2.

Now we estimate the quantities $\|D\|_\circ$, $\|h\Gamma D\|_\circ$, which are required for the next proposition. The estimates (5.27), (5.55), (5.56) lead to

$$\begin{aligned}
\|D\|_\circ &\leq \mathcal{C}\Phi_I(-\frac{h}{\varepsilon\mathcal{K}_1})\big(\|\bar{d}\|_2 + \|\Xi\|_\circ\big) \leq \\
&\leq \mathcal{C}\Phi_I(-\frac{h}{\varepsilon\mathcal{K}_1})\Big(\left\{ \begin{array}{c} \mathcal{C}\varepsilon h^{q_o} \\ \mathcal{C}\frac{h}{\varepsilon}\varepsilon h^{q_o-1} \end{array} \right\} + \mathcal{C}\frac{h}{\varepsilon}\varepsilon h^{q_o}\Big) \leq \\
&\leq \left\{ \begin{array}{ll} \mathcal{C}\varepsilon h^{q_o} & \text{Radau Ia} \\ \mathcal{C}\min\{\varepsilon, h\}h^{q_o-1} & \text{Radau IIa, Gauss.} \end{array} \right. \qquad (5.74)
\end{aligned}$$

The estimates (5.28), (5.55), (5.56) lead to

$$\begin{aligned}
\|h\Gamma D\|_\circ &\leq \mathcal{C}(\|\bar{d}\|_2 + \|\Xi\|_\circ) \leq \mathcal{C}(\left\{ \begin{array}{c} \mathcal{C}\varepsilon h^{q_o} \\ \mathcal{C}h^{q_o} \end{array} \right\} + \mathcal{C}h^{q_o+1}) \leq \\
&\leq \left\{ \begin{array}{ll} \mathcal{C}(\varepsilon + h)h^{q_o} & \text{Radau IIa} \\ \mathcal{C}h^{q_o} & \text{Radau Ia, Gauss.} \end{array} \right. \qquad (5.75)
\end{aligned}$$

Now the following proposition can be proofed.

Proposition 5.4.3 *The collocation polynomial $\check{\eta}(t)$ of the methods Radau Ia, IIa and Gauss satisfies*

$$\|\check{p}'(c_i h) - f(\check{p}(c_i h))\|_2 \leq \left\{ \begin{array}{ll} \mathcal{C}\varepsilon h^s & \textit{Radau IIa} \\ \mathcal{C}\min\{h, \varepsilon\}h^{q_o-1} & \textit{Radau Ia, Gauss} \end{array} \right. \qquad (5.76)$$

for $h \in (0, h_{\max}]$ and $i = 1, \ldots, s$, where $\check{p}(t) = p(\check{\eta}(t))$. The constants have to be chosen as stated in the proof.

Proof: Proposition 5.4.2 together with (5.74) yields

$$\begin{aligned}\|\check{p}'(c_ih) - f(\check{p}(c_ih))\|_2 &\leq \mathcal{C}\|D_i\|_2 \max_j \|\check{\pi}_j'(c_ih)\|_2 \leq \\ &\leq \max_j \|\check{\pi}_j'(c_ih)\|_2 \begin{cases} \mathcal{C}\varepsilon h^{q_o} & \text{Radau IIa} \\ \mathcal{C}\min\{\varepsilon, h\}h^{q_o-1} & \text{Radau Ia, Gauss.} \end{cases}\end{aligned}$$

The estimate (5.75) can be used to bound

$$\begin{aligned}\|\check{\pi}_j'(c_ih)\|_2 &= \|J\pi_j(\check{\eta}(c_ih))\check{\eta}'(c_ih)\|_2 = \|J\pi_j(Y_i)f(Y_i)\|_2 = \\ &= \|J\pi_j(Y_i)(f(P_i) + \frac{1}{h}\sum_{m=1}^{k}[h\Lambda_iD_i]_m\Pi_{im})\|_2 \leq \\ &\leq \mathcal{C}\Big(M_{f|_{\mathcal{M}}} + \frac{1}{h}\begin{cases} \mathcal{C}(\varepsilon+h)h^{q_o} \\ \mathcal{C}h^{q_o} \end{cases}\Big) \leq \mathcal{C}\end{aligned}$$

which completes the proof. □

Proposition 5.4.4 *Let $h \in (0, \bar{h}_{\max}]$, where $\bar{h}_{\max}$ is defined in the proof below. Furthermore let $\check{\eta}(t)$ be the collocation polynomial and $\tilde{u}(t)$ the solution of the differential equation (2.11) with initial value $\bar{p}$, then there holds*

$$\max_{t\in[0,h]} \|\tilde{u}(t) - \check{\eta}(t)\|_2 \leq \begin{cases} \mathcal{C}(h+\varepsilon)h^{q_o} & \textit{Radau IIa} \\ \mathcal{C}h^{q_o} & \textit{Radau Ia, Gauss} \end{cases} \tag{5.77}$$

$$\max_{t\in[0,h]} \|\tilde{u}^{(\kappa)}(t) - \check{\eta}^{(\kappa)}(t)\|_2 \leq \begin{cases} \mathcal{C}(h+\varepsilon)h^{q_o-\kappa} & \textit{Radau IIa} \\ \mathcal{C}h^{q_o-\kappa} & \textit{Radau Ia, Gauss} \end{cases} \tag{5.78}$$

where $\kappa = 1, \ldots, s$. The constants $\mathcal{C}$ have to be chosen as stated in the proof.

Proof: We interpolate $\tilde{u}'(t)$ with the help of the Lagrange basis ℓ_j

$$\begin{aligned}\tilde{u}'(t) &= \sum_{j=1}^{s} \tilde{u}'(c_jh)\ell_j(\tau) + r(t) = \\ &= \sum_{j=1}^{s} f(\tilde{u}(c_jh))\ell_j(\tau) + r(t),\end{aligned}$$

where $\tau := \frac{t}{h} \in [0,1]$ and $r(t)$ is the interpolation error. The interpolation error satisfies $\|r^{(\kappa)}(t)\|_2 \leq \mathcal{C}h^{s-\kappa}$ for $\kappa = 0, \ldots, s$ and $t \in [0,h]$. Integration of $\tilde{u}'(t)$ yields

$$\tilde{u}(t) = \bar{p} + \sum_{j=1}^{s} f(\tilde{u}(c_jh))h\int_0^\tau \ell_j(\sigma)\,d\sigma + R(t) \tag{5.79}$$

where $\|R(t)\|_2 \leq \mathcal{C}h^{s+1-\kappa}$ for $\kappa = 0, \ldots, s+1$ and $t \in [0,h]$. Analoguously we formulate the collocation polynomial with the help of the Lagrange basis

$$\begin{aligned}\check{\eta}(t) &= \bar{\eta} + \sum_{j=1}^{s} f(Y_j)h\int_0^\tau \ell_j(\sigma)\,d\sigma = \\ &= \bar{p} + \sum_{i=1}^{k}\bar{d}_i\bar{\pi}_i + \sum_{j=1}^{s} f(P_j)h\int_0^\tau \ell_j(\sigma)\,d\sigma + \\ &\quad + \sum_{j=1}^{s}\sum_{i=1}^{k}[\Lambda_jD_j]_i\Pi_{ji}h\int_0^\tau \ell_j(\sigma)\,d\sigma.\end{aligned} \tag{5.80}$$

The norm of the difference of (5.79) and (5.80) can be estimated via

$$\begin{aligned}
&\|\tilde{u}(t)-\check{\eta}(t)\|_2 \;\leq \\
&\quad\leq \|\bar{d}\|_2+\|\sum_{j=1}^{s}\Big(f(\underbrace{p(\tilde{u}(c_jh))}_{=\tilde{u}(c_jh)})-f(\underbrace{p(Y_j)}_{P_j})\Big)h\int_0^\tau \ell_j(\sigma)\,d\sigma\|_2+\\
&\qquad+\|\sum_{j=1}^{s}\sum_{i=1}^{k}[\Lambda_j d_j]_i\Pi_{ji}h\int_0^\tau \ell_j(\sigma)\,d\sigma\|_2+\|R(t)\|_2\;\leq\\
&\quad\leq \left\{\begin{array}{c}\mathcal{C}\varepsilon h^{q_o}\\ \mathcal{C}h^{q_o}\end{array}\right\}+\mathcal{C}\max_{t\in[0,h]}\|\tilde{u}(t)-\check{\eta}(t)\|_2 hB_0+\\
&\qquad+\left\{\begin{array}{c}\mathcal{C}(h+\varepsilon)h^{q_o}\\ \mathcal{C}h^{q_o}\end{array}\right\}B_0+\mathcal{C}h^{s+1}\;\leq\\
&\quad\leq h\mathcal{C}B_0\max_{t\in[0,h]}\|\tilde{u}(t)-\check{\eta}(t)\|_2+\left\{\begin{array}{ll}\mathcal{C}(\varepsilon+h)h^{q_o} & \text{Radau IIa}\\ \mathcal{C}h^{q_o} & \text{Radau Ia, Gauss}\end{array}\right.
\end{aligned}$$

where

$$\Big\|f\Big(p(\tilde{u}(c_jh))\Big)-f\Big(p(Y_j)\Big)\Big\|_2\leq L_{f|_{\mathcal{M}}}\underbrace{\max_{y\in\mathcal{G}}\|Jp(y)\|_2}_{\leq\mathcal{C}}\|\tilde{u}(c_jh)-Y_j\|_2$$

and

$$B_0:=\max_{\tau\in[0,1]}\Big(\sum_{j=1}^{s}|\int_0^\tau \ell_j(\sigma)d\sigma|\Big)$$

is used. The stepsize restriction

$$h\leq\frac{1}{2\mathcal{C}B_0} \tag{5.81}$$

implies

$$\|\tilde{u}(t)-\check{\eta}(t)\|_2\;\leq\;\left\{\begin{array}{ll}2\mathcal{C}(\varepsilon+h)h^{q_o}=\mathcal{C}(\varepsilon+h)h^{q_o} & \text{Radau IIa}\\ 2\mathcal{C}h^{q_o}=\mathcal{C}h^{q_o} & \text{Radau Ia, Gauss}\end{array}\right. \tag{5.82}$$

for all $t\in[0,h]$. This is the reason for defining

$$\bar{h}_{\max}:=\min\Big\{h_{\max},\frac{1}{2\mathcal{C}B_0}\Big\},$$

where $\mathcal{C}$ comes from (5.81). Differentiation of the difference $\tilde{u}(t)-\check{\eta}(t)$ yields

$$\begin{aligned}
\frac{d^\kappa}{dt^\kappa}\Big(\tilde{u}(t)-\check{\eta}(t)\Big) \;&=\; \sum_{j=1}^{s}\Big(f(p(\tilde{u}(c_jh)))-f(p(Y_j))\Big)h^{1-\kappa}\frac{d^\kappa}{d\tau^\kappa}\int_0^\tau \ell_j(\sigma)\,d\sigma-\\
&\quad-\sum_{j=1}^{s}\sum_{i=1}^{k}[\Lambda_jD_j]_i\Pi_{ji}h^{1-\kappa}\frac{d^\kappa}{d\tau^\kappa}\int_0^\tau \ell_j(\sigma)\,d\sigma+R^{(\kappa)}(t)
\end{aligned}$$

for $\kappa=1,\ldots,s$. Now we estimate

$$\|\tilde{u}^{(\kappa)}(t)-\check{\eta}^{(\kappa)}(t)\|_2\;\leq$$

$$
\begin{aligned}
&\leq L_{f|_{\mathcal{M}}}\mathcal{C}\max_{t\in[0,h]}\|\tilde{u}(t)-\check{\eta}(t)\|_2 h^{1-\kappa}B_\kappa + \left\{\begin{array}{c}\mathcal{C}(\varepsilon+h)h^{q_o-\kappa}\\ \mathcal{C}\varepsilon h^{q_o-\kappa}\end{array}\right\}B_\kappa + \\
&\quad + \mathcal{C}h^{s+1-\kappa} \leq \\
&\leq L_{f|_{\mathcal{M}}}\mathcal{C}\left\{\begin{array}{c}\mathcal{C}(\varepsilon+h)h^{q_o}\\ \mathcal{C}h^{q_o}\end{array}\right\}h^{1-\kappa}B_\kappa + \left\{\begin{array}{c}\mathcal{C}(h+\varepsilon)h^{q_o-\kappa}\\ \mathcal{C}h^{q_o-\kappa}\end{array}\right\}B_\kappa + \\
&\quad + \mathcal{C}h^{s+1-\kappa} \leq \\
&\leq \left\{\begin{array}{ll}\mathcal{C}(\varepsilon+h)h^{q_o-\kappa} & \text{Radau IIa}\\ \mathcal{C}h^{q_o-\kappa} & \text{Radau Ia, Gauss}\end{array}\right.
\end{aligned}
\tag{5.83}
$$

where

$$
B_\kappa := \max_{\tau\in[0,1]}\Big(\sum_{j=1}^{s}\Big|\frac{d^{\kappa-1}}{d\tau^{\kappa-1}}\ell_j(\tau)\Big|\Big)
$$

for $\kappa = 1,\ldots,s$. □

In the case of the method Radau Ia there is $q_o = s-1$, i.e. $\max_{t\in[0,1]}\|\tilde{u}^{(s)}(t)-\check{\eta}^{(s)}(t)\|_2 \leq O(h^{-1})$. This means that the estimate is useless in the context of the following estimates and therefore we restrict our considerations to the methods Radau IIa and Gauss.

5.4.1 The Smooth Error Component for the Methods Radau IIa and Gauss in the Strongly Stiff Case

Proposition 5.4.4 and the smoothness of $\tilde{u}(t)$ imply that under a moderate stepsize restriction there holds $\check{\eta}(t)\in\mathcal{V}$ for $t\in[0,h]$. This means that we can define $\check{p}(t) = p(\check{\eta}(t))$ and $\check{x}(t)$ with $\phi(\check{x}(t)) = \check{p}(t)$ for all $t\in[0,h]$. Now we estimate the smooth error component $p(\eta_\nu) - \tilde{u}_{\nu-1}(t_\nu) = \check{p}(h) - \tilde{u}(h)$. We consider the local coordinates and estimate

$$
\|\check{p}(t)-\tilde{u}(t)\|_2 \;\leq\; L_\phi\|\check{x}(t)-\tilde{x}(t)\|_2. \tag{5.84}
$$

Then the Gröbner-Alekseev Theorem (compare theorem A.4.1 in the appendix) is applied to the differential equations

$$
\begin{aligned}
\tilde{x}'(t) &= \hat{\psi}(\tilde{x}(t)),\\
\check{x}'(t) &= \hat{\psi}(\check{x}(t)) + \Big(\check{x}'(t) - \hat{\psi}(\check{x}(t))\Big)
\end{aligned}
$$

with initial values $\tilde{x}(0) = \check{x}(0) = 0$. This yields

$$
\check{x}(t)-\tilde{x}(t) \;=\; \int_0^t \frac{\partial\tilde{x}}{\partial\check{x}}\Big(t,\sigma,\check{x}(\sigma)\Big)\Big(\check{x}'(\sigma)-\hat{\psi}(\check{x}(\sigma))\Big)d\sigma, \tag{5.85}
$$

where $\tilde{x}(t,\sigma,\check{x})$ denotes the solution of the initial value prblem

$$
\frac{\partial\tilde{x}}{\partial t}(t,\sigma,\check{x}) \;=\; \hat{\psi}(\tilde{x}(t,\sigma,\check{x})), \quad \tilde{x}(\sigma,\sigma,\check{x}) = \check{x}
$$

with initial time σ and initial value $\check{x}$. This means that the partial derivative of the solution with respect to the initial value which is denoted by $\frac{\partial\tilde{x}}{\partial\check{x}}(t,\sigma,\check{x})$ can be computed from the variational

equations

$$\begin{aligned}\frac{\partial}{\partial t}\frac{\partial \tilde{x}}{\partial \check{x}}(t,\sigma,\check{x}) &= J\hat{\psi}(\tilde{x}(t,\sigma,\check{x}))\frac{\partial \tilde{x}}{\partial \check{x}}(t,\sigma,\check{x}),\\ \frac{\partial \tilde{x}}{\partial \check{x}}(\sigma,\sigma,\check{x}) &= I_{n-k}.\end{aligned}$$

In order to compute the integral in (5.85) we use the quadrature formula with the wights b_i from the implicit Runge-Kutta method and obtain at $t = h$

$$\check{x}(h) - \tilde{x}(h) \;=\; h\sum_{i=1}^{s} b_i \frac{\partial \tilde{x}}{\partial \check{x}}\Big(h, c_i h, \check{x}(c_i h)\Big)\Big(\check{x}'(c_i h) - \hat{\psi}(\check{x}(c_i h))\Big) + \text{Rest}, \tag{5.86}$$

with

$$\|\text{Rest}\|_2 \le \mathcal{C} h^{p_o+1} \sup_{h\in(0,\bar{h}_{\max}]} \max_{\sigma\in[0,h]} \|\rho^{(p_o)}(\sigma)\|_2$$

and

$$\rho(\sigma) := \frac{\partial \tilde{x}}{\partial \check{x}}\Big(h,\sigma,\check{x}(\sigma)\Big)\Big(\check{x}'(\sigma) - \hat{\psi}(\check{x}(\sigma))\Big).$$

The next step is to estimate the terms in (5.86).

1. The smoothness assumption for $\hat{\psi}(x)$ in $\mathcal{U}$ implies that the solution of the variational equations is bounded by a moderate constant

$$\Big\|\frac{\partial \tilde{x}}{\partial \check{x}}\Big(h, c_i h, \check{x}(c_i h)\Big)\Big\|_2 \;\le\; \mathcal{C}. \tag{5.87}$$

2. Proposition 5.4.3 implies

$$\begin{aligned}\|\check{x}'(c_j h) - \hat{\psi}(\check{x}(c_j h))\|_2 &\le \|\check{p}'(c_j h) - f(\check{p}(c_j h))\|_2 \le\\ &\le \begin{cases} \mathcal{C}\varepsilon h^{q_o} & \text{Radau IIa}\\ \mathcal{C}\min\{h,\varepsilon\}h^{q_o-1} & \text{Gauss.}\end{cases}\end{aligned} \tag{5.88}$$

The last step is to derive a h-independent bound for $\|\rho^{(p_o)}(\sigma)\|_2$. Therefore we estimate the following terms.

I. The derivatives $\check{x}^{(\kappa)}(\sigma)$ for $\kappa = 1,\dots,p_o+1$ satisfy

$$\|\check{x}^{(\kappa)}(\sigma)\|_2 \;\le\; \|\check{p}^{(\kappa)}(\sigma)\|_2.$$

Since $\check{p}^{(\kappa)}(\sigma) = \frac{d^\kappa}{d\sigma^\kappa} p(\check{\eta}(\sigma))$ can be written as a finite sum of products of the following terms

$$\text{i)}\ \frac{\partial^{|l|} p_i}{\partial y_1^{l_1}\cdots\partial y_n^{l_n}}(\check{\eta}(\sigma)) \quad\text{and}\quad \text{ii)}\ \check{\eta}^{(|l|)}(\sigma)$$

for $1 \le |l| \le \kappa$, $i = 1,\dots,n$ it suffices to show that the terms i) and ii) are moderately bounded. The norms of the terms i) are moderately bounded due to the smoothness assumption on $p(y)$. The terms ii) can be estimated by

$$\begin{aligned}\|\check{\eta}^{(|l|)}(\sigma)\|_2 &\le \|\tilde{u}^{(|l|)}(\sigma)\|_2 + \|\tilde{u}^{(|l|)}(\sigma) - \check{\eta}^{(|l|)}(\sigma)\|_2 \le\\ &\le \mathcal{C} + \begin{Bmatrix} \mathcal{C}(h+\varepsilon)h^{q_o-|l|}\\ \mathcal{C}h^{q_o-|l|}\end{Bmatrix} \le \mathcal{C}\begin{cases}\text{Radau IIa}\\ \text{Gauss}\end{cases}\end{aligned}$$

for $|l| = 1, \dots, s$. Furthermore the terms $\check{\eta}^{(|l|)}(\sigma)$ vanish for $|l| \geq s+1$. This implies the existence of a moderate constant $\mathcal{C} \geq 0$ such that

$$\|\check{x}^{(\kappa)}(\sigma)\|_2 \leq \mathcal{C} \tag{5.89}$$

for $\kappa = 1, \dots, p_o + 1$ independent of $h \in (0, \bar{h}_{\max}]$.

II. The terms $\frac{d^\kappa}{d\sigma^\kappa}\hat{\psi}(\check{x}(\sigma))$ for $\kappa = 1, \dots, p_o$ can be written as a finite sum of products of

$$\text{i) } \frac{\partial^{|l|}\hat{\psi}_i}{\partial x_1^{l_1} \cdots \partial x_{n-k}^{l_{n-k}}}(\check{x}(\sigma)) \quad \text{and} \quad \text{ii) } \check{x}^{(|l|)}(\sigma)$$

for $1 \leq |l| \leq \kappa$ and $i = 1, \dots, n-k$. The norms of the terms i) are moderately bounded due to the smoothness assumptions on $\hat{\psi}$. For the terms ii) there holds (5.89). This implies

$$\Big\| \frac{d^\kappa}{d\sigma^\kappa}\hat{\psi}(\check{x}(\sigma)) \Big\|_2 \leq \mathcal{C} \tag{5.90}$$

for $\kappa = 1, \dots, p_o$.

III. The derivatives of $\frac{\partial \tilde{x}}{\partial \check{x}}(h, \sigma, \check{x}(\sigma))$ with respect to σ up to the order p_o consist of a finite sum of products of the terms

$$\text{i) } \frac{\partial^{|l|}}{\partial \sigma^{l_0} \partial \check{x}_1^{l_1} \cdots \partial \check{x}_{n-k}^{l_{n-k}}} \frac{\partial \tilde{x}_i}{\partial \check{x}_j}(h, \sigma, \check{x}(\sigma)) \quad \text{and} \quad \text{ii) } \check{x}^{(|l|)}(\sigma)$$

where $1 \leq |l| \leq \kappa$ and $i, j = 1, \dots n-k$. The terms i) can be computed via linear differential equations where the right hand sides consist of finite sums of products of partial derivatives of $\hat{\psi}$ and $\tilde{x}(h, \sigma, \check{x})$ which are of a lower order. The appearing matrices and inhomogeneous terms of the right hand side as well as the initial values are moderately bounded such that the solutions of these differential equations are moderately bounded. This means that the terms i) are moderately bounded independent of h. We conclude that there holds

$$\Big\| \frac{d^\kappa}{d\sigma^\kappa} \frac{\partial \tilde{x}}{\partial \check{x}}(h, \sigma, \check{x}(\sigma)) \Big\|_2 \leq \mathcal{C} \tag{5.91}$$

for $\kappa = 1, \dots, p_o$.

The bounds (5.89), (5.90), (5.91) imply (h-independent)

$$\|\rho^{(p_o)}(\sigma)\|_2 \leq \mathcal{C}. \tag{5.92}$$

This leads to the estimate

$$\begin{aligned} \|\check{x}(h) - \tilde{x}(h)\|_2 &\leq h\|b_1\|_1 \left\{ \begin{array}{c} \mathcal{C}\varepsilon h^{q_o} \\ \mathcal{C}\min\{h, \varepsilon\} h^{q_o - 1} \end{array} \right\} + \mathcal{C}h^{p_o+1} \leq \\ &\leq \left\{ \begin{array}{ll} \mathcal{C}(\varepsilon + h^{p_o - q_o})h^{q_o+1} & \text{Radau IIa} \\ \mathcal{C}(\min\{\varepsilon, h\} + h^{p_o+1-q_o})h^{q_o} & \text{Gauss} \end{array} \right. \end{aligned}$$

which implies

$$\begin{aligned}\|\check{p}(h)-\tilde{u}(h)\|_2 &\leq L_\phi \left\{ \begin{array}{l} \mathcal{C}(\varepsilon+h^{p_o-q_o})h^{q_o+1} \\ \mathcal{C}(\min\{\varepsilon,h\}+h^{p_o+1-q_o})h^{q_o} \end{array} \right\} \leq \\ &\leq \left\{ \begin{array}{ll} \mathcal{C}(\varepsilon+h^{p_o-q_o})h^{q_o+1} & \text{Radau IIa} \\ \mathcal{C}(\min\{\varepsilon,h\}+h^{p_o+1-q_o})h^{q_o} & \text{Gauss.} \end{array} \right. \end{aligned} \tag{5.93}$$

Now we use the index notation , i.e. $p(\eta_\nu)=\check{p}(h)$ and $\tilde{u}_{\nu-1}$ instead of $\tilde{u}$ and estimate the smooth error component

$$\begin{aligned}&\|p(\eta_\nu)-u(t_\nu)\|_2 \leq \\ &\leq \|p(\eta_\nu)-\tilde{u}_{\nu-1}(t_\nu)\|_2+\|\tilde{u}_{\nu-1}(t_{\nu-1}+h)-u(t_\nu-1+h)\|_2 \leq \\ &\leq \left\{ \begin{array}{l} \mathcal{C}(\varepsilon+h^{p_o-q_o})h^{q_o+1} \\ \mathcal{C}(\min\{\varepsilon,h\}+h^{p_o+1-q_o})h^{q_o} \end{array} \right\} + e^{L_{f|_{\mathcal{M}}}h}\|p(\eta_{\nu-1})-u(t_{\nu-1})\|_2 \leq \\ &\leq e^{L_{f|_{\mathcal{M}}}t_\nu}\|p(\eta_0)-u_0\|_2+\frac{e^{L_{f|_{\mathcal{M}}}t_\nu}-1}{e^{L_{f|_{\mathcal{M}}}h}-1}\left\{ \begin{array}{l} \mathcal{C}(\varepsilon+h^{p_o-q_o})h^{q_o+1} \\ \mathcal{C}(\min\{\varepsilon,h\}+h^{p_o+1-q_o})h^{q_o} \end{array} \right\} \leq \\ &\leq e^{L_{f|_{\mathcal{M}}}t_\nu}\|p(\eta_0)-u_0\|_2+ \\ &\quad + (e^{L_{f|_{\mathcal{M}}}t_\nu}-1)\left\{ \begin{array}{ll} \mathcal{C}(\varepsilon+h^{p_o-q_o})h^{q_o} & \text{Radau IIa} \\ \mathcal{C}(\min\{\varepsilon,h\}+h^{p_o+1-q_o})h^{q_o-1} & \text{Gauss.} \end{array} \right. \end{aligned} \tag{5.94}$$

Under the assumption

$$\|p(\eta_0)-\tilde{u}_0\|_2 \leq \left\{ \begin{array}{ll} C_0(\varepsilon+h^{p_o-q_o})h^{q_o} & \text{Radau IIa} \\ C_0(\min\{\varepsilon,h\}+h^{p_o+1-q_o})h^{q_o-1} & \text{Gauss} \end{array} \right. \tag{5.95}$$

where $C_0 \geq 0$ and $h \leq \bar{h}_{\max}$ there holds

$$\begin{aligned}&\|p(\eta_\nu)-u(t_\nu)\|_2 \leq \\ &\leq e^{L_{f|_{\mathcal{M}}}T}\left\{ \begin{array}{l} C_0(\varepsilon+h^{p_o-q_o})h^{q_o} \\ C_0(\min\{\varepsilon,h\}+h^{p_o+1-q_o})h^{q_o-1} \end{array} \right\} + \\ &\qquad + (e^{L_{f|_{\mathcal{M}}}t_\nu}-1)\left\{ \begin{array}{l} \mathcal{C}(\varepsilon+h^{p_o-q_o})h^{q_o} \\ \mathcal{C}(\min\{\varepsilon,h\}+h^{p_o+1-q_o})h^{q_o-1} \end{array} \right\} \leq \\ &\qquad \leq \left\{ \begin{array}{ll} \mathcal{C}(\varepsilon+h^{p_o-q_o})h^{q_o} & \text{Radau IIa} \\ \mathcal{C}(\min\{\varepsilon,h\}+h^{p_o+1-q_o})h^{q_o-1} & \text{Gauss.} \end{array} \right. \end{aligned} \tag{5.96}$$

5.4.2 Error Bounds for the Methods Radau IIa and Gauss in the Strongly Stiff Case

Now we consider the whole error as a combinaton of the stiff and the smooth error components:

$$\|\eta_\nu-u(t_\nu)\|_2 \leq \underbrace{\|\eta_\nu-p(\eta_\nu)\|_2}_{=\|d(\eta_\nu)\|_2}+\|p(\eta_\nu)-u(t_\nu)\|_2 \leq$$

$$\leq \left\{ \begin{array}{l} K_0 \varepsilon h^{q_o} \\ K_\infty h^{q_o} \end{array} \right\} + \left\{ \begin{array}{l} \mathcal{C}(\varepsilon + h^{p_o - q_o}) h^{q_o} \\ \mathcal{C}(\min\{h, \varepsilon\} + h^{p_o + 1 - q_o}) h^{q_o - 1} \end{array} \right\} \leq$$

$$\leq \left\{ \begin{array}{ll} \mathcal{C}(\varepsilon + h^{p_o - q_o}) h^{q_o} & \text{Radau IIa} \\ \mathcal{C} h^{q_o} & \text{Gauss} \end{array} \right. \tag{5.97}$$

for all $\nu \in \mathbb{N}_0$ with $t_\nu \in [0, T]$.

Theorem 5.4.1 (Convergence of the methods Radau IIa and Gauss – strongly stiff case) *Let $B_0, C_0 \geq 0$, $K_0 := B_0 M_{V^{-1}}^{\frac{1}{2}}$ and*

$$\|p(\eta_0) - u_0\|_2 \leq \left\{ \begin{array}{ll} C_0(\varepsilon + h^{s-1}) h^s & \textit{Radau IIa} \\ C_0(\min\{h, \varepsilon\} + h^{s+1}) h^{s-1} & \textit{Gauss} \end{array} \right.$$

$$\|\eta_0 - p(\eta_0)\|_2 \leq \left\{ \begin{array}{ll} B_0 \varepsilon h^s & \textit{Radau IIa} \\ B_0 h^s & \textit{Gauss} \end{array} \right.$$

where $K_0 \in [K_{\min}, \frac{1}{\varepsilon}]$ for the method Radau IIa and $h \in (0, \bar{h}_{\max}]$. Then there holds

$$\|\eta_\nu - u(t_\nu)\|_2 \leq \left\{ \begin{array}{ll} \mathcal{C}(\varepsilon + h^{s-1}) h^s & \textit{Radau IIa} \\ \mathcal{C} h^s & \textit{Gauss} \end{array} \right. \tag{5.98}$$

$$\|p(\eta_\nu) - u(t_\nu)\|_2 \leq \left\{ \begin{array}{ll} \mathcal{C}(\varepsilon + h^{s-1}) h^s & \textit{Radau IIa} \\ \mathcal{C}(\min\{h, \varepsilon\} + h^{s+1}) h^{s-1} & \textit{Gauss} \end{array} \right. \tag{5.99}$$

$$\|\eta_\nu - p(\eta_\nu)\|_2 \leq \left\{ \begin{array}{ll} \mathcal{K}_0 \max\{B_0, B_1\} \varepsilon h^s & \textit{Radau IIa} \\ M_V^{\frac{1}{2}} K_\infty h^s & \textit{Gauss} \end{array} \right. \tag{5.100}$$

for all $\nu \in \mathbb{N}_0$ with $t_\nu \in [0, T]$. The appearing constants have to be chosen like in the considerations above.

Remark: In the strongly stiff case, i.e. where ε is very small in comparision to h, the smooth component is nearly spuerconvergent. This means that for the method Radau IIa we have $O(\varepsilon h^s + h^{2s-1})$ and for the method Gauss we have $O(\varepsilon h^{s-1} + h^{2s})$.

Chapter 6

Example of a Nonlinear Stiff ODE

We cosider a nonlinear autonomous ODE in $\mathbb{R}^3$, where the right hand side $f(y)$ is defined as follows:

$$f(y) := \begin{pmatrix} -\alpha y_2 & + & \lambda(y_1^2 + y_2^2 - 1)y_1 & + & \mu y_3 y_1 \\ \alpha y_1 & + & \lambda(y_1^2 + y_2^2 - 1)\rho y_2 & + & \mu y_3 y_2 \\ & & \lambda(y_1^2 + y_2^2)\sigma y_3 & + & \mu(1 - y_1^2 - y_2^2) \end{pmatrix} \tag{6.1}$$

There exists a onedimensional smooth invarinat manifold

$$\mathcal{M} = \{y \in \mathbb{R}^3 \colon y_1^2 + y_2^2 = 1 \wedge y_3 = 0\}. \tag{6.2}$$

Compare figure 6.1 concerning the flow of the ODE, where the parameters in $f(y)$ are set as:

λ	α	ρ	σ	μ
-10	5	3	1	20

The parameter λ describes the stiffness of the problem. The parameters α and μ describe the rotation of the solutions.

We hope that we can use the ODE (2.11) with the right hand side (6.1) as a model-problem for our nonlinear stiff model class. Therefore we have to ask whether this model-problem satisfies the assumptions of section 2.2. Since the analysis of this model problem is in work we close our considerations at this open question.

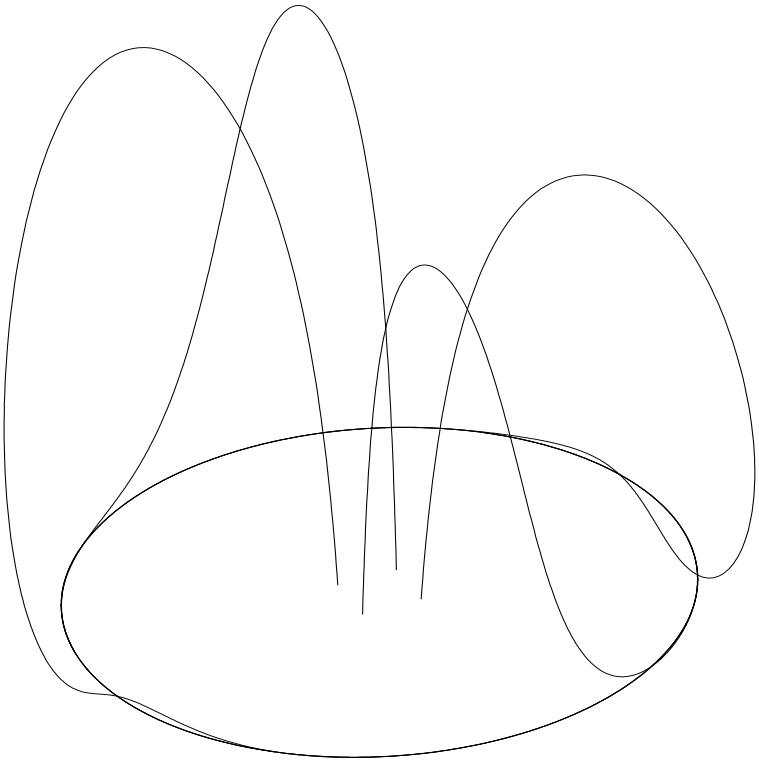


Figure 6.1: Some solutions in a neighbourhood of $\mathcal{M}$

Appendix A

A.1 A Fixed Point Theorem

Let $n \in \mathbb{N}$, $\mathcal{D} \subseteq \mathbb{R}^n$ and $F\colon \mathcal{D} \to \mathbb{R}^n$ a function. We consider the fixed point equation

$$\xi \;=\; F(\xi) \tag{A.1}$$

and ask for existence and local uniqueness of a solution.

Theorem A.1.1 (Fixed point theorem) *Let $n \in \mathbb{N}$, $\xi_0 \in \mathbb{R}^n$ a start value, $0 < r$,*

$$\mathcal{B}_r(\xi_0) := \{\xi \in \mathbb{R}^n \colon \|\xi - \xi_0\| < r\}$$

and $F\colon \overline{\mathcal{B}_r(\xi_0)} \to \mathbb{R}^n$ a function. If there exists $0 \le K < 1$ such that

$$\|F(\xi) - F(\tilde{\xi})\| \;\le\; K\|\xi - \tilde{\xi}\|, \tag{A.2}$$

for all $\xi, \tilde{\xi} \in \overline{\mathcal{B}_r(\xi_0)}$ and if the first iteration step satisfies

$$\|F(\xi_0) - \xi_0\| \;\le\; (1-K)r, \tag{A.3}$$

then there holds:

$$\xi_k \;:=\; F(\xi_{k-1}) \in \mathcal{B}_r(\xi_0) \quad \forall k \in \mathbb{N}, \tag{A.4}$$

$$\xi^* \;:=\; \lim_{k\to\infty} \xi_k \in \overline{\mathcal{B}_r(\xi_0)} \quad \textit{exists where}$$

$$\xi^* \;=\; F(\xi^*) \quad \textit{is the unique fixed point in} \quad \overline{\mathcal{B}_r(\xi_0)}. \tag{A.5}$$

Proof: can be found in [19, p.248]. □

If there exists a closed ball $\overline{\mathcal{B}_r(\xi_0)} \in \mathcal{D}$ such that the assumptions of the fixed point theorem A.1.1 are fulfilled, then the fixed point problem (A.1) restricted to $\overline{\mathcal{B}_r(\xi_0)}$ has a unique solution.

A.1.1 Equivalent Systems of Nonlinear Equations

In order to analyze the solvability of a nonlinear equation with the help of a fixed point theorem an appropriate fixed point form is required. Therefore we give some equivalent formulations of a nonlinear equation, equivalent in the sense that the solution set stays invariant.

- Let $F(\xi), H(\xi)$ be two vector valued functions.[1)] Then the equations

$$F(\xi) \;=\; 0 \tag{A.6}$$

and

$$F(\xi) + H(\xi) \;=\; H(\xi) \tag{A.7}$$

for ξ have the same solution set.

- Let $F_1(\xi, \zeta), F_2(\xi), H(\xi)$ be vector valued functions. Then the Equations

$$F_1(\xi, H(\xi)) \;=\; 0, \tag{A.8}$$

$$F_2(\xi) \;=\; H(\xi) \tag{A.9}$$

and

$$F_1(\xi, F_2(\xi)) \;=\; 0, \tag{A.10}$$

$$F_2(\xi) \;=\; H(\xi) \tag{A.11}$$

for ξ have the same solution set.

- Let $F(\xi)$ be a vector valued function and $G(\xi)$ a matrix valued function such that $G(\xi)$ is regular for all ξ. Then $F(\xi) = 0$ and

$$G(\xi)F(\xi) \;=\; 0 \tag{A.12}$$

have the same solution set.

A.2 The Square Root of Positive Definite Matrices

Proposition A.2.1 *Let $k \in \mathbb{N}$, $M_V, M_{V^{-1}} \geq 0$ and*

$$\mathcal{P} \;:=\; \{V \in \mathbb{R}^{k\times k} \text{ positive definite}\colon \|V\|_2 \leq M_V \;\wedge\; \|V^{-1}\|_2 \leq M_{V^{-1}}\}.$$

Then there exists a positive real constant $L_{\surd}$ such that

$$\|V^{\frac{1}{2}} - \tilde{V}^{\frac{1}{2}}\|_2 \;\leq\; L_{\surd}\|V - \tilde{V}\|_2 \tag{A.13}$$

for all $V, \tilde{V} \in \mathcal{P}$. The constant $L_{\surd}$ depends on the constants $M_V, M_{V^{-1}}$ and on k.

Proof: The spectrum of $V \in \mathcal{P}$ is contained in the intervall $[\frac{1}{M_{V^{-1}}}, M_V]$. Since the main branche of the root function is analytic in the right complex half plane we can use the Cauchy integral formula to express $V^{\frac{1}{2}}, \tilde{V}^{\frac{1}{2}}$ for $V, \tilde{V} \in \mathcal{P}$. This yields

$$\|V^{\frac{1}{2}} - \tilde{V}^{\frac{1}{2}}\|_2 \;=\; \|\oint_{\kappa} \sqrt{z}(R(z) - \tilde{R}(z))\,dz\|_2 \tag{A.14}$$

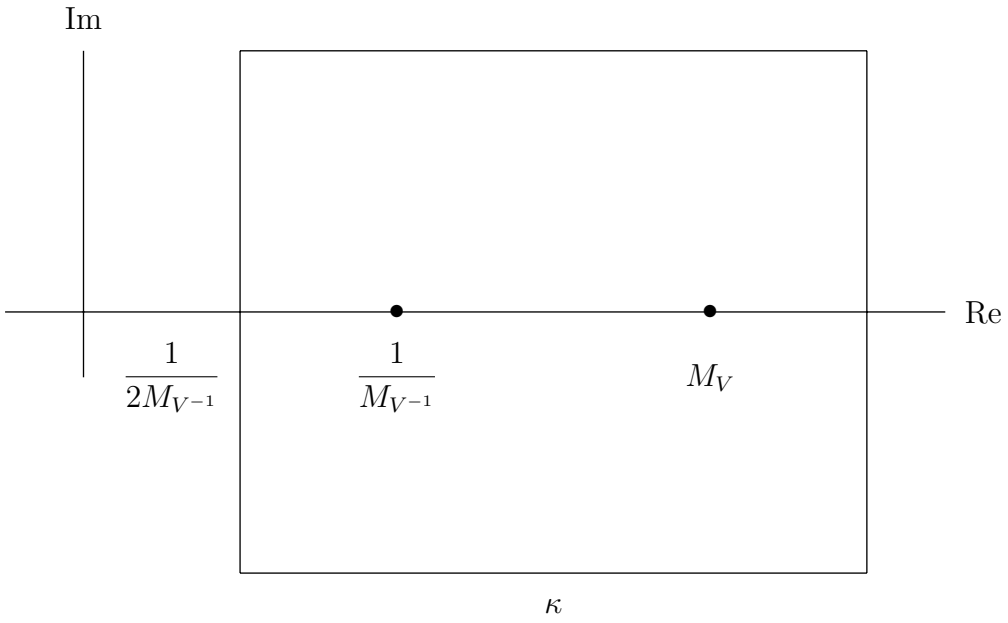


Figure A.1: The path κ

where $R(z) := (z - V)^{-1}$, $\tilde{R}(z) := (z - \tilde{V})^{-1}$ are the corresponding resolvents and κ is the path as described in figure A.1.
The length $|\kappa|$ of the path κ is

$$|\kappa| \;=\; 2\Big(M_V + \frac{1}{M_{V^{-1}}}\Big) =: M_\kappa. \tag{A.15}$$

The function $|\sqrt{z}|$ on the path κ can be estimated by

$$|\sqrt{z}| \;\leq\; \sqrt{M_V + \frac{1}{2M_{V^{-1}}}} =: M_{\sqrt{}}, \quad \forall z \in \kappa. \tag{A.16}$$

The differece of the resolvents can be estimated by

$$\|R(z) - \tilde{R}(z)\|_2 \;\leq\; \|R(z)\|_2 \|\tilde{R}(z)\|_2 \|V - \tilde{V}\|_2.$$

For the resolvents restricted to the path κ there holds [2)]

$$\begin{aligned} \|R(z)\|_2 \;&\leq\; \frac{\|z - V\|_2^{k-1}}{|\det(z - V)|} \;\leq \\ &\leq\; (2M_{V^{-1}})^k (2M_V + \frac{1}{2M_{V^{-1}}})^{k-1} =: M_R, \quad \forall z \in \kappa \end{aligned}$$

[1)] We additionally assume that the appearing functions are of the form such that the following equations are well defined, i.g. that the appearing operations make sense.

[2)] The first bound is due to Kato's inequality which says that

$$\|B^{-1}\|_2 \;\leq\; \frac{\|B\|_2^{k-1}}{|\det B|}$$

for all regular $k \times k$ matrices B.

as well as $\|\tilde{R}(z)\|_2 \leq M_R$ for all $z \in \kappa$. We conclude

$$\|R(z) - \tilde{R}(z)\|_2 \;\leq\; M_R^2 \|V - \tilde{V}\|_2, \quad \forall z \in \kappa. \tag{A.17}$$

We estimate (A.14) by the use of (A.15)-(A.17) and obtain

$$\|V^{\frac{1}{2}} - \tilde{V}^{\frac{1}{2}}\|_2 \;\leq\; M_\kappa M_{\surd} M_R^2 \|V - \tilde{V}\|_2.$$

Now we set $L_{\surd} := M_\kappa M_{\surd} M_R^2$ which completes the proof. □

A.3 The Direct Product

The *direct product* of matrices or vectors (*Kronecker Product*) is defined as follows

$$A \otimes B := \begin{pmatrix} a_{11}B & \cdots & a_{1s}B \\ \vdots & & \vdots \\ a_{s1}B & \cdots & a_{ss}B \end{pmatrix}, \quad a \otimes b := \begin{pmatrix} a_1 b \\ \vdots \\ a_s b \end{pmatrix}, \tag{A.18}$$

where $A \in \mathbb{R}^{s\times s}$, $B \in \mathbb{R}^{n\times n}$ and $a \in \mathbb{R}^s$, $b \in \mathbb{R}^n$.

Proposition A.3.1 *Let $\|\cdot\|_\circ$ be the vectornorm defined as*

$$\|X\|_\circ := \|(\|X_1\|_2, \ldots, \|X_s\|_2)\|_\infty$$

for $X = (X_1, \ldots, X_s)$ where $X_i \in \mathbb{R}^n$. Then in the corresponding matrixnorm $\|\cdot\|_\circ$ there holds

$$\|A \otimes I\|_\circ = \|A\|_\infty \tag{A.19}$$

where $A \in \mathbb{R}^{s\times s}$ and I is the identity matrix in $\mathbb{R}^n$.

Proof: The estimate

$$\begin{aligned} \|A \otimes I\|_\circ &= \sup_{\|X\|_\circ = 1} \|(A \otimes I)X\|_\circ = \\ &= \sup_{\|X\|_\circ = 1} \max_i \|\sum_{j=1}^{s} a_{ij} X_j\|_2 \;\leq \\ &\leq \sup_{\|X\|_\circ = 1} \underbrace{\max_j \|X_j\|_2}_{=\|X\|_\circ = 1} \max_i \sum_{j=1}^{s} |a_{ij}| = \|A\|_\infty. \end{aligned}$$

The bound is sharp since

$$\|(A \otimes I)X\|_\circ = \|A\|_\infty$$

for

$$X = \frac{1}{\sqrt{n}} (\underbrace{\operatorname{sgn}(a_{i_{\max}1}), \ldots, \operatorname{sgn}(a_{i_{\max}1})}_{n \text{ times}}, \ldots, \underbrace{\operatorname{sgn}(a_{i_{\max}s}), \ldots, \operatorname{sgn}(a_{i_{\max}s})}_{n \text{ times}})^\top$$

where $i_{\max}$ is choosen such that

$$\sum_{j=1}^{s} |a_{i_{\max}j}| = \max_i \sum_{j=1}^{s} |a_{ij}|.$$

□

A.4 The Nonlinear Variation of Constants Formula

Theorem A.4.1 (Alekseev, Gröbner) *Let $y(t)$ be the solution of the IVP*

$$\begin{aligned} y'(t) &= f(t, y(t)), \\ y(t_0) &= y_0 \end{aligned} \tag{A.20}$$

and $z(t)$ the solution of the perturbed IVP

$$\begin{aligned} z'(t) &= f(t, z(t)) + g(t, z(t)), \\ z(t_0) &= z_0 \end{aligned} \tag{A.21}$$

then the solutions are connected by

$$z(t) - y(t) = \int_{t_0}^{t} \frac{\partial y}{\partial z}(t, \sigma, z(\sigma)) g(\sigma, z(\sigma)) d\sigma. \tag{A.22}$$

Remark: The functions f, g are vector valued, i.e. $f, g \colon \mathbb{R}^{n+1} \to \mathbb{R}^n$. We denote by $y(t, t_0, y_0)$ the solution of the differential equation (A.20) dependent on the initial value y_0 at initial time t_0. Then the matrix

$$\frac{\partial y}{\partial y_0}(t, t_0, y_0)$$

is the solution of the matrix IVP

$$\begin{aligned} \frac{\partial}{\partial t}\frac{\partial y}{\partial y_0}(t, t_0, y_0) &= \frac{\partial f}{\partial y}(t, y(t, t_0, y_0))\frac{\partial y}{\partial y_0}(t, t_0, y_0), \\ \frac{\partial y}{\partial y_0}(t_0, t_0, y_0) &= I. \end{aligned} \tag{A.23}$$

Proof: Compare [12, pp.97-99]. □

Appendix B

B.1 The Implicit Runge-Kutta Methods Radau Ia, IIa and Gauss

We consider the initial value problem

$$\begin{aligned} y'(t) &= f_\bullet(t, y(t)), \\ y(t_0) &= y_0. \end{aligned} \tag{B.1}$$

The s-stage implicit Runge-Kutta method is defined via the algebraic equations

$$Y_i = \eta_{\nu-1} + h\sum_{j=1}^{s} a_{ij} f_\bullet(t_{\nu-1} + c_j h, Y_j), \tag{B.2}$$

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

$$\eta_\nu = \eta_{\nu-1} + h\sum_{i=1}^{s} b_i f_\bullet(t_{\nu-1} + c_i h, Y_i), \tag{B.3}$$

where $t_\nu = t_0 + \nu h$.
The abscissas $c_1, \ldots, c_s$ of the Runge-Kutta methods Radau Ia, IIa, and Gauss are the abscissas of the Radau resp. Gauss quadrature transformed to the intervall $[0, 1]$:

Gauss: Let P_s^* be the Legendre polynomial of degree s transformed to the intervall $[0, 1]$. The abscissas c_i are the zeros of the P_s^*.

Radau Ia: There is $c_0 = 0$ and the remaining c_i's are the zeros of $P_{s-1}^* + P_s^*$.

Radau IIa: There is $c_s = 1$ and the remaining c_i's are the zreos of $P_{s-1}^*(1-t) + P_s^*(1-t)$.

In connection with order results the so-called simplifying conditions

$$\begin{aligned} B(\xi): &\quad \sum_{i=1}^{s} b_i c_i^{k-1} = \frac{1}{k}, && k = 1(1)\xi, \\ C(\xi): &\quad \sum_{j=1}^{s} a_{ij} c_j^{k-1} = \frac{1}{k} c_i^k, && k = 1(1)\xi,\ i = 1(1)s, \\ D(\xi): &\quad \sum_{i=1}^{s} b_i c_i^{k-1} a_{ij} = \frac{1}{k} b_j (1 - c_j^k), && k = 1(1)\xi,\ j = 1(1)s \end{aligned} \tag{B.4}$$

appear (compare [7]). The vector $b = (b_1, \ldots, b_s)^\top$ of the methods Radau Ia, IIa and Gauss is defined via the condition $B(s)$. In the case of the Radau Ia method the matrix $A = (a_{ij})_{ij}$ is defined

via the condition $D(s)$. The matrix A for the methods Radau IIa and Gauss is defined via $C(s)$. Furthermore there holds

$$\begin{array}{llll} \text{Radau Ia} & B(2s-1) & C(s-1) & D(s) \\ \text{Radau IIa} & B(2s-1) & C(s) & D(s-1) \\ \text{Gauss} & B(2s) & C(s) & D(s) \end{array} \tag{B.5}$$

The methods Radau Ia, IIa and Gauss are so-called collocation methods, i.e. there exists a collocation polynomial $\check{\eta}(t)$ of degree s. The polynomial $\check{\eta}(t)$ fulfills the conditions

$$\check{\eta}(c_i h) \;=\; f_\bullet(t_{\nu-1} + c_i h, \check{\eta}(c_i h)) \tag{B.6}$$

and $\check{\eta}(0) = \eta_{\nu-1}$, $\check{\eta}(c_i h) = Y_i$, $\check{\eta}(h) = \eta_\nu$, where $i = 1, \ldots, s$.

B.2 Stability Results for the Methods Radau Ia, IIa and Gauss

We consider the differential equation (B.1), where $f_\bullet : \mathbb{R}^{n+1} \to \mathbb{R}^n$ fulfills the one-sided Lipschitz condition

$$\langle f_\bullet(t,y) - f_\bullet(t,\tilde{y}), y - \tilde{y}\rangle_\bullet \;\leq\; m\|y-\tilde{y}\|_\bullet^2, \tag{B.7}$$

for all $y, \tilde{y} \in \mathcal{D} \subseteq \mathbb{R}^n$. Here $m \in \mathbb{R}$ is the so-called one-sided Lipschitz constant. $\langle\cdot,\cdot\rangle_\bullet$ denotes an arbitrary inner product and $\|\cdot\|_\bullet$ the corresponding induced norm.

B.2.1 BSI-Stability

A Runge-Kutta method is called BSI-stable if there exists a monotone increasing function Φ_I and a real constant $\bar{q} > 0$ such that the function $F := (F_1, \ldots, F_s)$ defined via

$$F_i(Z;\eta) \;:=\; Z_i - \eta - h\sum_{j=1}^{s} a_{ij} f_\bullet(t_{\nu-1} + c_j h, Z_j), \tag{B.8}$$

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

$$|||Z - \tilde{Z}|||_\bullet \;\leq\; \Phi_I(hm)|||F(Z;\eta) - F(\tilde{Z};\eta)|||_\bullet, \tag{B.9}$$

for all $Z, \tilde{Z} \in \mathcal{D}^s$, $\eta \in \mathcal{D}$ and $hm \leq \bar{q}$. The norm $|||\cdot|||_\bullet$ is defined as

$$|||Z|||_\bullet \;=\; \|(\|Z_1\|_\bullet, \ldots, \|Z_s\|_\bullet)^\top\|_2$$

for $Z = (Z_1, \ldots, Z_s)^\top \in \mathbb{R}^{s\times n}$.

Remark: In chapter 4 for example the function $f_\bullet(t,y) := \Lambda(t)y$ is used, where $\Lambda(t)$ is chosen such that there holds

$$\Lambda(t_{\nu-1} + c_i h) \;=\; \Lambda_i \tag{B.10}$$

for $i = 1, \ldots, s$. This special $f_\bullet$ satisfies a one sided Lipschitz condition with $m = O(-\frac{1}{\varepsilon})$ corresponding to the scalarproduct $\langle\cdot,\cdot\rangle_\bullet = \langle\cdot,\cdot\rangle_{\bar{V}}$. Note that in each step from $t_{\nu-1} \to t_\nu$ a separate $f_\bullet$ is defined on $\{t_{\nu-1} + c_i h : i = 1, \ldots, s\} \times \mathbb{R}^k$ with a corresponding inner product.

Theorem B.2.1 *The methods Radau Ia, IIa and Gauss are BSI-stable with a BSI-stability function*

$$\Phi_I(hm) \;=\; \max_{i,j} \sqrt{\frac{d_i}{d_j}} \frac{\|A\|_D}{\bar{q} - hm} \tag{B.11}$$

and $D = \mathrm{diag}(d_1, \ldots, d_s)$, $\bar{q}$ *from*

Verfahren	D	$\bar{q}$
Gauss	$\mathrm{diag}(b^\top)(\mathrm{diag}(c^\top)^{-1} - I_s)$	$\frac{1}{2}\min_i(c_i - c_i^2)^{-1}$
Radau Ia	$\mathrm{diag}(b^\top A)$	$\frac{1}{2}(1-c_2)^{-1}$
Radau IIa	$\mathrm{diag}(b^\top)\mathrm{diag}(c^\top)^{-1}$	$\frac{1}{2c_{s-1}} \quad (s \geq 2)$ $1 \quad (s = 1)$

where $\|\cdot\|_D$ *applied to matrices is the operatornorm corresponding to the vectornorm*

$$\|\xi\|_D \;=\; \sqrt{\sum_{j=1}^{s} d_j \xi_j^2}.$$

Remark: For our purposes (compare chapter 4) the BSI-stability function can be expressed in the form

$$\Phi_I(hm) \;=\; \frac{\mathcal{C}}{1 - \phi_I hm}, \tag{B.12}$$

where $\phi_I = \bar{q}^{-1}, m = -\frac{1}{\varepsilon \mathcal{K}_1}$ which in our generic notation means

$$\Phi_I(-\frac{h}{\varepsilon \mathcal{K}_1}) \;\leq\; \mathcal{C}, \tag{B.13}$$

$$\Phi_I(-\frac{h}{\varepsilon \mathcal{K}_1}) \cdot \frac{h}{\varepsilon} \;\leq\; \mathcal{C} \tag{B.14}$$

for all $\frac{h}{\varepsilon} > 0$.

B.2.2 BS-Stability

A Runge-Kutta method is called BS-stable if there exists a monotone increasing function Φ_{BS} and a constant $\hat{q} > 0$ such that for an unperturbed Runge-Kuttat step

$$\begin{aligned} Y_i &= \eta_{\nu-1} + h \sum_{j=1}^{s} a_{ij} f_\bullet(t_{\nu-1} + c_j h, Y_j), \quad i = 1, \ldots, s \\ \eta_\nu &= \eta_{\nu-1} + h \sum_{i=1}^{s} b_i f_\bullet(t_{\nu-1} + c_i h, Y_i) \end{aligned}$$

and a perturbed Runge-Kutta step

$$\begin{aligned} \tilde{Y}_i &= \eta_{\nu-1} + h \sum_{j=1}^{s} a_{ij} f_\bullet(t_{\nu-1} + c_j h, \tilde{Y}_j) + \Delta_i, \quad i = 1, \ldots, s \\ \tilde{\eta}_\nu &= \eta_{\nu-1} + h \sum_{i=1}^{s} b_i f_\bullet(t_{\nu-1} + c_i h, \tilde{Y}_i) + \delta \end{aligned}$$

there holds

$$\|\tilde{\eta}_\nu - \eta_\nu\|_\bullet \leq \Phi_{BS}(hm)(|||\Delta|||_\bullet + \|\delta\|_\bullet) \tag{B.15}$$

for $hm < \hat{q}$.

Theorem B.2.2 *The methos Radau Ia, IIa and Gauss are BS-stabel.*

Remarks:

- For the method Radau IIa there holds $b_j = a_{sj}$ for $j = 1, \ldots, s$. This implies that the last line of the Radau IIa method is redundant, which means that for the BS-stability δ can be set equal to Δ_s and therefore
$$\Phi_{BS} = \Phi_I.$$
- For our special function $f_\bullet(t, y) := \tilde{\Lambda}(t)y$ in chapter 5, where we choose
$$\tilde{\Lambda}(t_{\nu-1} + c_i h) = \mathfrak{V}_i \Lambda_i \mathfrak{V}_i^{-1} \tag{B.16}$$
we have a one sided Lipschitz constant $m = -\frac{1}{\varepsilon \mathcal{K}_1}$ corresponding to the inner product $\langle \cdot, \cdot \rangle_V$. In the case of the Radau IIa method this implies a BS-stability function of the form
$$\Phi_{BS}(-\frac{h}{\varepsilon \mathcal{K}_1}) = \frac{\mathcal{C}}{1 + \phi_I \frac{h}{\varepsilon \mathcal{K}_1}} \tag{B.17}$$
(compare page 63).

B.2.3 B-Stability

A Runge-Kutta method is called B-stable if there exists a monotone increasing function Φ_B with $\Phi_B(0) = 1$ and a constant $q > 0$ such that for two parallel Runge-Kutta steps

$$\begin{aligned}
Y_i &= \eta_{\nu-1} + h \sum_{j=1}^{s} a_{ij} f_\bullet(t_{\nu-1} + c_j h, Y_j), \quad i = 1, \ldots, s \\
\eta_\nu &= \eta_{\nu-1} + h \sum_{i=1}^{s} b_i f_\bullet(t_{\nu-1} + c_i h, Y_i), \\
\tilde{Y}_i &= \tilde{\eta}_{\nu-1} + h \sum_{j=1}^{s} a_{ij} f_\bullet(t_{\nu-1} + c_j h, \tilde{Y}_j), \quad i = 1, \ldots, s \\
\tilde{\eta}_\nu &= \tilde{\eta}_{\nu-1} + h \sum_{i=1}^{s} b_i f_\bullet(t_{\nu-1} + c_i h, \tilde{Y}_i)
\end{aligned}$$

there holds

$$\|\tilde{\eta}_\nu - \eta_\nu\|_\bullet \leq \Phi_B(hm) \|\tilde{\eta}_{\nu-1} - \eta_{\nu-1}\|_\bullet \tag{B.18}$$

for all $hm < q$.

Theorem B.2.3 *The methods Radau Ia, IIa and Gauss are B-stable with a B-stability function*

$$\Phi_B(hm) = \begin{cases} 1 & \textit{Gauss} \\ \frac{1}{\sqrt{1-2hm}} & \textit{Radau Ia} \\ \frac{1}{\sqrt{1-2b_s hm}} & \textit{Radau IIa} \end{cases} \tag{B.19}$$

Remark: In our case (i.e. on page 63) the B-stability function for the method Radau IIa corresponding to the function $f_\bullet(t,y) = \tilde{\Lambda}(t)y$ reads as

$$\Phi_B(-\frac{h}{\varepsilon\mathcal{K}_1}) \;=\; \frac{1}{\sqrt{1+\phi_B\frac{h}{\varepsilon\mathcal{K}_1}}}\,, \tag{B.20}$$

where $\phi_B := 2b_s$.

Bibliography

[1] Auzinger W., Frank R., Kirlinger G.: *A Note on Convergence Concepts for Stiff Problems.* Computing 44, 197-208 (1990).

[2] Auzinger W., Frank R., Kirlinger G.: *Modern convergence theory for stiff initial-value problems.* Appl. Numer. Math. 45, 5-16 (1993).

[3] Auzinger W., Frank R., Kirlinger G.: *Extending Convergence Theory for Nonlinear Stiff Problems. Part I.* BIT 36:4, 635-652, (1996); Extendid version: Report 116/94, Institute for Applied and Numerical Mathematics, Vienna University of Technology.

[4] Auzinger W., Frank R., Stetter H.J.: *Vienna contributions to the development of RK-methods.* Appl. Numer. Math. 22, 35-49 (1996).

[5] Auzinger W., Frank R., Kirlinger G.: *An extension of B-convergence for Runge-Kutta methods.* Appl. Numer. Math. 9, 91-109 (1992).

[6] Auzinger W., Frank R., Eder A.: *Condition notes on a linear problem class with one stiff parameter.* Working notes (1998), Institute for Applied and Numerical Mathematics, Vienna University of Technology.

[7] Burrage K., Butcher J.C.: *Stability criteria for implicit Runge-Kutta methods.* SIAM J. Numer. Anal., 16, 46-57 (1979).

[8] Dekker K., Verwer J.G.: *Stability of Runge-Kutta methods for stiff nonlinear differential equations.* North-Holland 1984.

[9] Frank R., Schneid J., Überhuber C.W.: *The concept of B-convergence.* SIAM J. Numer. Anal., 18, 753-780 (1981).

[10] Frank R., Schneid J., Überhuber C.W.: *Stability Properties of implicit Runge-Kutta methods.* SIAM J. Numer. Anal., 22, 497-514 (1985).

[11] Frank R., Schneid J., Überhuber C.W.: *Order results for implicit Runge-Kutta methods applied to stiff systems.* SIAM J. Numer. Anal., 22, 515-534 (1985).

[12] Hairer E., Nørsett S.P., Wanner G.: *Solving Ordinary Differential Equations I.* Nonstiff Problems. Springer 1987.

[13] Hairer E., Wanner G.: *Solving Ordinary Differential Equations II.* Stiff and Differential-Algebraic Problems. Springer 1991.

[14] Hairer E., Lubich C., Roche M.: *Error of Runge-Kutta methods for stiff problems studied via differential algebraic equations.* BIT 28, 678-700 (1988).

[15] Horn R.A., Johnson C.R.: *Topics in matrix analysis.* Cambridge University Press 1991.

[16] Kågström B.: *Bounds and perturbation bounds for the matrix exponential.* BIT 17, 39-57 (1977).

[17] Nipp K.: *Invariant manifolds of singularly perturbed ordinary differential equations.* ZAMP 36, 309-320 (1985).

[18] Nipp K., Stoffer D.: *Invariant Manifolds and global error estimates of numerical intgration schemes applied to stiff systems of singular perturbation type.* Numer. Math., 70, 245-257, (1995).

[19] Stoer J.: *Numerische Mathematik I.* Springer 1989.

[20] Ström T.: *On logarithmic norms.* SIAM J. Numer. Anal. 12, 741-753 (1975)

[21] Strehmel K., Weiner R.: *Numerik gewöhnlicher Differentialgleichungen.* Teubner Studienbücher 1995.