BIT 36:4 (1996), 635–652.

EXTENDING CONVERGENCE THEORY FOR NONLINEAR STIFF PROBLEMS PART I *

W. AUZINGER, R. FRANK, and G. KIRLINGER

Institut für Angewandte und Numerische Mathematik, Technische Universität Wien
Wiedner Hauptstrasse 8–10/1152, A-1040 Wien, Austria
email: winfried@uranus.tuwien.ac.at

Abstract.

Existing convergence concepts for the analysis of discretizations of nonlinear stiff problems suffer from considerable drawbacks. Our intention is to extend the convergence theory to a relevant class of nonlinear problems, where stiffness is axiomatically characterized in natural geometric terms.

Our results will be presented in a series of papers. In the present paper (Part I) we motivate the need for such an extension of the existing theory, and our approach is illustrated by means of a convergence argument for the Implicit Euler scheme.

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

Key words: Stiff differential equations, convergence.

1 Introduction.

Existing theoretical approaches concerning the convergence of discretization methods for stiff initial value problems suffer from considerable limitations: The theory of *B-convergence*, for instance (cf., e.g., [3, 5, 6, 7]), relies on the one-sided Lipschitz constant m as the essential problem-characterizing parameter and yields satisfactory error bounds for problems where m is not strongly positive. This latter requirement, however, is typically only satisfied for stiff problems where the Jacobian is very close to normal throughout, whereas the 'generic case' is $m \gg 0$. This means that, in general, stiff problems are locally ill-conditioned, i.e., for each point y there exist points z near y such that the difference of the solutions starting in these points is locally strongly increasing (like $e^{mt}\|z-y\|$). However, such an effect occurs only on a very short interval and does not contradict the fact that a stiff problem is typically well-conditioned in a global sense. For a detailed discussion of this essential observation cf. [2].

ODEs in *singular perturbation form* have also been considered as models for stiff problems (cf., e.g., [8, 9, 10, 11]). However, they have a very special structure; in particular, the variation of the 'stiff eigendirections' is only $O(\varepsilon)$ ($\varepsilon \ll 1$),

*Received January 1995.

such that the phase portrait resembles that of a linear constant coefficient problem. (For a more detailed discussion cf. [1].)

With this background in mind, we are striving for a more comprehensive theory. In particular, we shall consider a class of problems where stiffness is axiomatically described in a way which can be considered as a natural formalization of the qualitative, phenomenological notion of stiffness with which most authors agree: The essence of stiffness is that a stiff differential equation admits smooth solutions with moderate derivatives, together with nonsmooth ('transient') solutions rapidly converging towards smooth solutions. Stiff problems are assumed to be well-conditioned in a global sense (in contrast to the bad local condition which often occurs as a consequence of the transient behavior of nonsmooth solutions).

A mathematical formalization of stiffness should reflect such a behavior and enable a satisfactory convergence theory. In such a theory, the essential goal is to derive error bounds ensuring that a particular method is able to accurately follow a solution on grids adjusted to its smoothness only, unaffected by large problem parameters like a large (conventional or one-sided) Lipschitz constant or like the moduli of stiff eigenvalues of the Jacobian.

Historically, the first attempt to 'define' stiffness was to consider the spectrum of the local Jacobians, which is called stiff if there exist eigenvalues with large moduli and if all these eigenvalues have strongly negative real parts. Such a definition of stiffness has a long tradition and frequently appears in the existing literature. However, it is oriented on the linear constant coefficient case only and is simply too naive for nonlinear situations: For nonlinear problems there is no reason to assume that the requirement of a stiff spectrum of the Jacobian in each point is equivalent to a stiff behavior. The notion of stiffness is rather related to the global solution behavior (i.e., on the way how transient solutions converge towards smooth ones) and can therefore not be reduced to the solution structure in the infinitesimal neighborhood of single points.

Our present approach is based on a more natural, semi-global way to linearize. Here, instead of considering the Jacobians in single points y, each point y is associated with another point u on the smooth solution manifold such that $y-u$ can be considered as a stiff eigendirection in a generalized sense. For the details of this approach cf. Section 2, (2.8).

For an illustration of the above discussion, let us consider the following example.

EXAMPLE 1.1.

$$\begin{aligned} y_1' &= -y_2 - \lambda y_1(1-y_1^2-y_2^2), \\ y_2' &= y_1 - \rho\lambda y_2(1-y_1^2-y_2^2), \end{aligned} \tag{1.1}$$

where $\lambda \ll 0$ characterizes the stiffness and ρ is a moderate positive parameter. For $\lambda = -10$ and $\rho = 3$ the phase portrait is shown in Fig. 1.1 (here the rather moderate value $\lambda = -10$ has been chosen to make the behavior of transient solutions more visible). Obviously, equation (1.1) shows the typical qualitative stiffness properties: The circle $y_1^2 + y_2^2 = 1$ represents an invariant manifold $\tilde{\mathcal{M}}$ on which smooth solutions persist; any other solution starting from an arbitrary

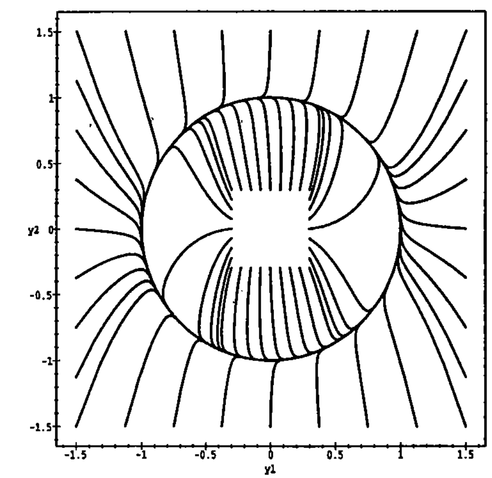

Figure 1.1: Phase portrait for Example 1.1.

point $(y_1, y_2) \in \mathbb{R}^2 \backslash (0,0)$ rapidly converges towards such a smooth solution. In this sense (1.1) is stiff, with $\mathbb{R}^2 \backslash (0,0)$ as basin of attraction for $\tilde{\mathcal{M}}$.

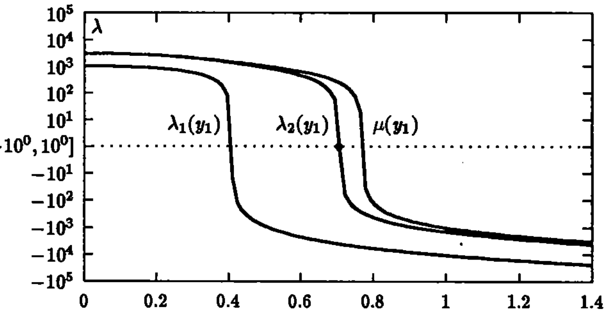

Figure 1.2: Eigenvalues and logarithmic norm of the Jacobian for Example 1.1.

Fig. 1.2 illustrates the spectral behavior of the local Jacobians for this example. Here, for $\lambda = -10^3$ and $\rho = 3$ the eigenvalues λ_1, λ_2 and the logarithmic norm μ[1] of the Jacobians are considered as functions of y_1, along the straight line $y_2 = y_1$ $(0 \le y_1 \le 1.4)$. In Fig. 1.2 the point $y_1 = \sqrt{2}/2$, where $y_2 = y_1$ intersects the smooth manifold $\tilde{\mathcal{M}}$, is marked by a bullet. Up to $y_1 \approx 0.4$ all quantities are strongly positive, and the Jacobian does not possess a stiff spectrum, although

[1] Here, μ denotes the logarithmic norm of the Jacobian w.r.t. the Euclidean norm. Note that μ is the locally optimal one-sided Lipschitz constant w.r.t. this norm.

this neighborhood of $(0,0)$ belongs to the basin of attraction of $\tilde{\mathcal{M}}$. Moreover, $\lambda_2(y_1)$ remains strongly positive even very close to $\tilde{\mathcal{M}}$. Thus, not even in a very close neighborhood of $\tilde{\mathcal{M}}$ the Jacobian possesses a stiff spectrum. Furthermore, $\mu(y_1)$ remains strongly positive in a $O(1)$-neighborhood of $\tilde{\mathcal{M}}$, up to $y_1 \approx 0.8$.

Numerical evidence shows that the common stiff solvers integrate equation (1.1) without any difficulties; but none of the existing theoretical concepts can be used to explain this behavior: Due to $m \gg 0$, the B-theory is not successfully applicable; furthermore, due to the significant variation of the stiff eigendirection, (1.1) cannot be considered as a problem in singular perturbation form.

Our present approach is sufficiently comprehensive to cover problems with a general nonlinear stiff geometry as in the above example. Furthermore, we will not only present error bounds for particular methods but also reveal certain structural information about the discretization error. In particular, we shall decompose the error into 'smooth' and 'stiff' components and it will turn out that the well-known order reduction phenomena (cf. [7, 12]) observed for implicit Runge-Kutta (IRK) methods occur in the stiff component only. Whether this fact can be exploited algorithmically to obtain high accuracy in an efficient way is an open question.

In the present paper (Part I of a series of papers) our basic ideas are introduced and illustrated under simplified assumptions. This should be convenient for those readers who are not too interested in technical details. Only the Implicit Euler scheme is studied, and certain simplifications are made concerning the problem class. In particular, only autonomous systems are considered for which the dimension of the smooth solution manifold is $n-1$ (for an n-dimensional ODE system). Furthermore, the solvability of the algebraic equations is out of the scope of Part I. In forthcoming papers we shall study the convergence properties of higher-order IRK schemes, together with the locally unique solvability of the respective algebraic equations. A more general problem class, with an $(n-k)$-dimensional smooth solution manifold, will be considered.

The performance of codes based on IRK schemes crucially depends on the way how the algebraic equations are solved. An efficient approach is Iterated Defect-Correction (IDeC), an iterative procedure the fixed point of which is the collocating polynomial corresponding to the Runge-Kutta solution (cf. [4]). The implicit equations involved in the IDeC iteration are only of the dimension n (number of given ODEs) in contrast to $n \times s$ for direct solution of the IRK equations ($s \dots$ number of IRK stages). However, up to now the applicability of the IDeC approach was restricted to schemes with equidistant internal abscissas, because otherwise the IDeC iteration as proposed in [4] does not converge. This excludes relevant methods like Gauss or Radau schemes. However, in the meantime we have found a way to modify the IDeC algorithm entailing rapid convergence also for non-equidistant collocation abscissas. The theoretical foundation of this observation is still incomplete; however, together with the superconvergence properties of the smooth error components this gives rise to the hope that such an algorithm may solve a stiff problem in an extremely efficient way.

2 A class of nonlinear stiff problems.

We consider initial value problems for autonomous stiff ODEs

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

where $f: \mathcal{G} \to \mathbb{R}^n$ ($\mathcal{G} \subset \mathbb{R}^n$). Throughout, $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$ denote the Euclidean inner product and norm in $\mathbb{R}^n$, respectively.

In order to formalize the notion of stiffness in geometrical terms we make the following assumptions:

• Existence of an invariant manifold $\tilde{\mathcal{M}}$ containing smooth solutions; smoothness assumptions concerning $\tilde{\mathcal{M}}$

By definition, it is an inherent property of stiff problems that there exist smooth solutions, i.e., solutions with moderate derivatives. Our particular assumption is that there exists an *invariant manifold* $\tilde{\mathcal{M}} \subset \mathcal{G}$ on which the smooth solutions persist. The convergence result presented in Section 3 refers to the case where $\tilde{\mathcal{M}}$ is an $(n-1)$-dimensional manifold.

Concerning $\tilde{\mathcal{M}}$, certain smoothness assumptions are natural. Basically, we assume that $\tilde{\mathcal{M}}$ is moderately curved. Further smoothness assumptions w.r.t. $\tilde{\mathcal{M}}$ refer to appropriate local parametrizations: For any point $u \in \tilde{\mathcal{M}}$, let $\tilde{\mathcal{M}}$ locally be represented by

$$\xi_n = \varphi(\xi_1, \dots, \xi_{n-1}; u), \tag{2.2}$$

where $(\xi_1, \dots, \xi_n)$ denotes a local Cartesian coordinate system with origin in u, chosen in such a way that $\xi_n \equiv 0$ represents the hyperplane $\mathcal{T}(u)$ tangent to $\tilde{\mathcal{M}}$ in u. Concerning the functions φ in (2.2) we assume that those derivatives exist and are bounded which are involved in the proofs of the following propositions and influence the quantities σ_1, σ_2 and σ_3 in (2.3), (2.4) and (2.5).

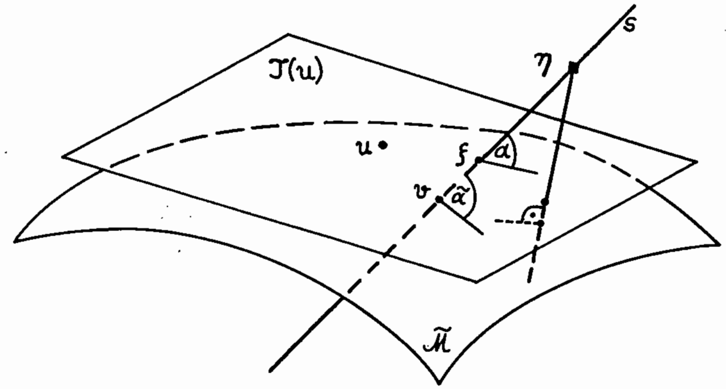

Figure 2.1: Propositions 2.1 and 2.2.

Under these smoothness assumptions the following assertions are valid:[2]

PROPOSITION 2.1. *Consider a point* $u \in \tilde{\mathcal{M}}$ *and another point* $\zeta \in \mathcal{T}(u)$ *satisfying* $\|\zeta - u\| \le \rho$. *Let* s *denote a straight line through* ζ, *with* $\alpha := \sphericalangle(s, \mathcal{T}(u))$, *and let* $v \in \tilde{\mathcal{M}}$ *denote the point where* s *intersects* $\tilde{\mathcal{M}}$.

Then,

$$\|\zeta - v\| \le \sigma_1 \rho^2. \tag{2.3}$$

Furthermore,

$$|\alpha - \tilde{\alpha}| \le \sigma_2 \rho, \tag{2.4}$$

where $\tilde{\alpha} := \sphericalangle(s, \tilde{\mathcal{M}})$.

For sufficiently small values of ρ, *the quantities* σ_1 *and* σ_2 *depend smoothly on certain derivatives of* φ (cf. (2.2)) *and on* $1/\alpha$, *remaining moderate provided these quantities are.*

PROPOSITION 2.2. *Consider a straight line* s *intersecting* $\tilde{\mathcal{M}}$ *in a point* v, *with* $\tilde{\alpha} := \sphericalangle(s, \tilde{\mathcal{M}})$. *Let* η *be a point on* s, *in a sufficiently close* $O(1)$*-neighborhood of* $\tilde{\mathcal{M}}$. *Then,*

$$\|\eta - v\| \le \sigma_3 \cdot \text{distance}(\eta, \tilde{\mathcal{M}}). \tag{2.5}$$

Here the quantity σ_3 *depends smoothly on certain derivatives of* φ (cf. (2.2)) *and on* $1/\tilde{\alpha}$, *remaining moderate provided these quantities are.*

Furthermore, let $\mathcal{P}$ *denote an arbitrary plane containing* s. *Then, in a neighborhood of* v, *the intersection curve* $\mathcal{P} \cap \tilde{\mathcal{M}}$ *remains smooth provided* $\tilde{\alpha}$ *is not too small.*

• Behavior of the flow of the given ODE restricted to $\tilde{\mathcal{M}}$

For the flow of the given ODE (2.1) restricted to $\tilde{\mathcal{M}}$ we assume one-sided Lipschitz continuity

$$\langle u - v, f(u) - f(v) \rangle \le \tilde{m}\|u - v\|^2 \quad \text{for } u, v \in \tilde{\mathcal{M}} \tag{2.6}$$

with a one-sided Lipschitz constant $\tilde{m} = m(f|_{\tilde{\mathcal{M}}})$ which is not strongly positive. (2.6) implies that the difference between two smooth solutions on $\tilde{\mathcal{M}}$ cannot grow faster than $e^{\tilde{m}t}$. Note that the assumption that $\tilde{m}$ is not strongly positive is natural and significantly less restrictive than the usual one in the B-theory, where one-sided Lipschitz continuity with a moderate one-sided Lipschitz constant m is required for *arbitrary* pairs $u, v \in \mathcal{G}$ (and not only on $\tilde{\mathcal{M}}$); cf. the discussion at the beginning of Section 1.

By $\tilde{M}_1$ and $\tilde{M}_2$ we shall denote bounds for the first and second derivatives, uniformly valid for arbitrary smooth solutions $u(t)$ on $\tilde{\mathcal{M}}$:

$$\|u'\| = \|f(u)\| \le \tilde{M}_1, \quad \|u''\| = \|f_y(u)f(u)\| \le \tilde{M}_2 \quad \text{for all } u \in \tilde{\mathcal{M}}. \tag{2.7}$$

[2]Here and in the sequel, the angle $\sphericalangle(s, \mathcal{H}) \in [0, \frac{\pi}{2}]$ between a straight line s and an $(n-1)$-dimensional hypersurface $\mathcal{H}$ is understood as the minimal angle between s and any curve κ on $\mathcal{H}$ which is the intersection of $\mathcal{H}$ with any (2-dimensional) plane containing s.

• Transversality condition

For each $\eta \notin \tilde{\mathcal{M}}$ in a sufficiently close $O(1)$-neighborhood of $\tilde{\mathcal{M}}$ there exists a point[3] $u(\eta) \in \tilde{\mathcal{M}}$ such that

$$f(\eta) - f(u(\eta)) = \lambda(\eta)(\eta - u(\eta)) \quad \text{with } \lambda(\eta) \le \bar{\lambda} \ll 0. \tag{2.8}$$

The corresponding difference will be denoted by

$$\pi(\eta) := \eta - u(\eta). \tag{2.9}$$

The quantities $\lambda(\eta)$ and $\pi(\eta)$ may be thought of as generalizations of the notions 'stiff eigenvalue' and 'stiff eigenvector', respectively: For linear problems $y' = Ay$ with a stiff eigenvalue $\lambda \ll 0$, a relation $A\eta - Au = \lambda(\eta - u)$ holds whenever $\eta - u$ is a corresponding eigenvector, and the difference between the solutions $\eta(t)$ and $u(t)$ starting from η and u, respectively, is given by $e^{\lambda t}(\eta - u)$. The idea behind (2.8) is to describe stiffness for nonlinear problems by an analogous property, namely that for η sufficiently close to $\tilde{\mathcal{M}}$ and their associated $u(\eta) \in \tilde{\mathcal{M}}$, the corresponding f-difference is a scalar multiple of $\eta - u(\eta)$, with a factor $\lambda(\eta) \ll 0$. This implies that, locally, the difference between the transient solution starting from η and the smooth solution starting from $u(\eta)$ is strongly exponentially decreasing. Thus, (2.8)—together with (2.6)—reflects essential structural properties of stiff problems in a natural way.

In our convergence analysis, further natural assumptions concerning the behavior of the stiff eigendirections $\pi(\eta)$ are crucial:

- The angle $\alpha(\eta)$ between $\pi(\eta)$ and $\tilde{\mathcal{M}}$ is required to be significantly away from 0, such that

  $$\sin\alpha(\eta) \ge s_0 > 0, \quad \text{with } \frac{1}{s_0} \text{ moderate-sized.} \tag{2.10}$$

  Together with Proposition 2, this implies

  $$\|\pi(\eta)\| \le \sigma_3 \cdot \text{distance}(\eta, \tilde{\mathcal{M}}) \tag{2.11}$$

  for all η in a sufficiently close $O(1)$-neighborhood of $\tilde{\mathcal{M}}$.

- Continuous variation of the stiff eigendirection:

  The angle $\sphericalangle(\pi(\eta), \pi(\zeta))$ between $\pi(\eta)$ and $\pi(\zeta)$ is assumed to satisfy

  $$\sphericalangle(\pi(\eta), \pi(\zeta)) \le \beta(\eta, \zeta)\|\eta - \zeta\| \tag{2.12}$$

  where $\beta(\eta, \zeta) \le \bar{\beta}$ for all η, ζ in a sufficiently close $O(1)$-neighborhood of $\tilde{\mathcal{M}}$.

[3]Due to nonlinearity, equation (2.8) may not be solvable or may admit more than one solution $u \in \tilde{\mathcal{M}}$ for an arbitrary $\eta \in \mathcal{G} \subset \mathbb{R}^n$ (cf. (2.1)). For stiff equations, however, it is natural to assume that in a 'semi-local' sense, i.e., for η in a sufficiently close $O(1)$-neighborhood of $\tilde{\mathcal{M}}$, a solution $u(\eta) \in \tilde{\mathcal{M}}$ exists and is locally unique.

So far we have introduced a number of problem-characterizing parameters as for instance $\tilde{m}$, $\tilde{M}_1$, $\tilde{M}_2$, $\bar{\lambda}$, etc. In the sequel we distinguish between

- the *stiff parameter* $\bar{\lambda} \ll 0$;
- the *non-stiff parameters* $\tilde{m}$, $\tilde{M}_1$, $\sigma_1, \ldots$, which are assumed to be of moderate size compared to $|\bar{\lambda}|$.

REMARK 2.1. For a particular problem at hand, the verification of the above assumptions (smoothness of $\tilde{\mathcal{M}}$, transversality condition, ...) or the explicit determination of the problem-characterizing parameters ($\tilde{m}$, $\tilde{M}_i$, $\sigma_i, \ldots$) will be impossible or will require an extremely high computational effort—except in trivial cases. In particular, $\tilde{\mathcal{M}}$ would have to be explicitly known for this purpose. However, this is typical in the context of an a-priori theory. A-priori bounds are mathematical assertions concerning the method considered, yielding structural information about the discretization error. The assumptions involved simply characterize the problem class under consideration; it makes no sense to require that the occurring parameters should be necessarily computable in practice. (A-posteriori error estimation— based on quantities already computed and serving for control purposes— is a different topic.)

3 Convergence of the Implicit Euler scheme.

Let us now consider the simplest implicit one-step method, the Implicit Euler scheme with stepsize h,

$$\frac{1}{h}(\eta_\nu - \eta_{\nu-1}) = f(\eta_\nu). \tag{3.1}$$

As mentioned in Section 1, we assume that for any $\eta_{\nu-1}$ in a sufficiently close $O(1)$-neighborhood of $\tilde{\mathcal{M}}$ there exists a unique η_ν in this neighborhood satisfying (3.1).

• **Remark concerning error bounds of the form Ch^p:** Traditionally, error bounds for discretization schemes are written in the form $O(h^p)$, in the sense that there exists a bound Ch^p (with some problem-dependent factor C) which is valid uniformly for $h \to 0$. We prefer a more explicit denotation, where the factor C is specified as an expression in problem parameters (like Lipschitz constants, ...). In fact, the unreflecting use of the shorthand O-denotation is one of the main reasons for the historical fact that in the analysis of stiff problems some important aspects have been consistently overlooked. For a long time, in particular, the only aspect taken into consideration in the analysis of stiff problems was stability. The traditional local error estimates for Runge-Kutta schemes, however, are based on elementary differentials, and are therefore inevitably influenced by the norm of the Jacobians, or the (conventional) Lipschitz constant $L \gg 0$. This essential difficulty was concealed behind the unprecise O-denotation and therefore not taken notice of. Such a misunderstanding appears even in the recent literature, e.g., in conjunction with remainder estimates for asymptotic error expansions.

In our quantitative convergence concept, error bounds are denoted in the form Ch^p, where C is a well-defined expression in problem parameters. A further important difference to the conventional order concept is the following: A conventional bound of the form $O(h^p)$, i.e., Ch^p with C unspecified, makes no sense for fixed values of h, because for fixed h and p, the mere existence of an appropriate C is trivial. A nontrivial mathematical assertion is only obtained if the bound is shown to be uniformly valid for $h \to 0$. But if the factor C is explicitly specified in the sense described above, a bound Ch^p makes sense for any fixed h or for $h \in [h_{min}, h_{max}]$—and not only asymptotically for $h \to 0$.

Concerning the following convergence theorem some further preliminary remarks are in order:

• In order not to overload the denotation, we often write expressions of the type $C(\textit{problem-characterizing parameters}) \cdot h^p$ in the generic form $\mathcal{C} \cdot h^p$. Here the symbol $\mathcal{C}$ stands for well-defined expressions in non-stiff problem parameters which remain moderate—in particular, significantly smaller than $|\bar{\lambda}|$. As far as the O-symbol is used, it has to interpreted in exactly the same sense.

• The theorem refers to the case that the Implicit Euler scheme (3.1) follows a smooth solution $\tilde{u}(t)$ of (2.1), determined by an initial value $\tilde{u}(0) \in \tilde{\mathcal{M}}$, such that $\tilde{u}(t) \in \tilde{\mathcal{M}}$ for $t \geq 0$. Sufficiently small initial perturbations are admitted. In particular, we assume that η_0 satisfies

$$\|\eta_0 - \tilde{u}(0)\| \leq \mathcal{C}h, \tag{3.2a}$$

$$\|\pi(\eta_0)\| \leq Kh/|\bar{\lambda}|, \tag{3.2b}$$

with a constant K satisfying

$$K_{min} \leq K \leq |\bar{\lambda}| \tag{3.3}$$

(for the definition of $\bar{\lambda}$ and π cf. (2.8), (2.9)). At first sight, it seems to make no sense to bound K from below ($K \geq K_{min}$). However, the point is that—analogously to (3.2b)—all quantities $\|\pi(\eta_\nu)\|$ will turn out to be bounded by $Kh/|\bar{\lambda}|$, with the *same* constant K (cf. (3.6) below). But this can only be ensured for K not too small ($K = 0$ in (3.2b) and (3.6) would mean that all η_ν stay in $\tilde{\mathcal{M}}$ provided $\eta_0 \in \tilde{\mathcal{M}}$). It will turn out in the proofs below that K_{min} is a certain well-defined moderate generic constant satisfying $K_{min} \ll |\bar{\lambda}|$. Thus, by (3.3) an interval of admissible constants K is well-defined.

Note that the theorem provides more than a mere $O(h)$-convergence result: (3.6) shows that the stiff error component $\pi(\eta_\nu)$ is $O(h/|\bar{\lambda}|)$ provided $\pi(\eta_0)$ is.

• Taking initial perturbations into account is important with regard to nonequidistant grids; here the actual integration interval—to which the theorem refers and where h is kept constant—is considered as a subinterval of the whole integration interval (discretized by a piecewise equidistant grid). In this context, $\eta_0 - \tilde{u}(0)$ is interpreted as the global error accumulated in the preceding subintervals. The theorem is formulated in such a way that its inductive applicability w.r.t. several subintervals is straightforward. Only for the first subinterval, (3.2)/(3.3) is required to start this induction. If this initial error results from an integration along a transient phase, sufficiently fine grids must be used in the transients to ensure that distance$(\eta_0, \tilde{\mathcal{M}})$ is sufficiently small; then, (2.11) ensures that $\|\pi(\eta_0)\|$ is sufficiently small. The resulting value for K, i.e., for the magnitude of the stiff error components $\pi(\eta_\nu)$, depends on the accuracy of this transient integration.

THEOREM 3.1. Convergence of the Implicit Euler scheme.
For the problem class introduced in Section 2 and under the assumptions (3.2) *and* (3.3), *the Implicit Euler scheme* (3.1) *is convergent of first order, i.e.,*

$$\|\eta_\nu - \tilde{u}(t_\nu)\| \le Ch \tag{3.4}$$

under a mild stepsize restriction

$$h \in (0, h_{max}], \tag{3.5}$$

where h_{max} is not affected by the stiffness of the underlying problem, i.e., not prohibitively small. Furthermore,

$$\|\pi(\eta_\nu)\| \le K\frac{h}{|\bar{\lambda}|} \tag{3.6}$$

at all grid points $t_\nu = \nu h$, with the same value for K as in (3.2b).

In the sequel we will give two proofs:

• Proof 1 is shorter, more transparent and obtained under milder assumptions w.r.t. the given problem, e.g., without explicit use of (2.12). However, Proof 1 is only valid under the following restrictions:

- The stepsize h is required to satisfy

$$h \in [h_{min}, h_{max}], \tag{3.7}$$

  - where the precise characterization of h_{min} will be given in Proof 1 below; the order of magnitude of h_{min} is $O(1/|\bar{\lambda}|) \ll 1$. Thus by (3.7) those cases are covered where the accuracy requirements are not too severe, such that the required stepsizes are not extremely small. This standard situation is often called 'strongly stiff'.
  - The parameter K in assumption (3.2) is required to be moderate but not at the $O(|\bar{\lambda}|)$-level (cf. (3.32) in Proof 1 below). The general case where $K = O(|\bar{\lambda}|)$ is admitted is covered by Proof 2.

• Proof 2 refers to the full stepsize interval (3.5) and is valid for arbitrary K in (3.2)/(3.3), at the cost of being more technical. Proof 2 covers not only strongly stiff situations but also 'mildly stiff', 'weakly stiff', or ....[4] stiff situations.

PROOF 1. To estimate the global error $\eta_\nu - \tilde{u}(t_\nu)$ we split it up in the following way:

[4]Please fill in your favorite term.

$$\|\eta_\nu - \tilde{u}(t_\nu)\| \le \|\eta_\nu - u_\nu\| + \|u_\nu - \tilde{u}(t_\nu)\|, \tag{3.8}$$

where

$$u_\nu := u(\eta_\nu) \in \tilde{\mathcal{M}} \quad \text{in the sense of (2.8).} \tag{3.9}$$

The quantities

$$\pi_\nu := \pi(\eta_\nu) = \eta_\nu - u_\nu, \qquad \delta_\nu := u_\nu - \tilde{u}(t_\nu) \tag{3.10}$$

may be called the 'stiff' and the 'smooth component' of the global error, respectively. Below, we shall study a pair of coupled recursions for the δ_ν and π_ν. To this end we introduce the following terminology: Let $u_\nu(t)$ denote that local solution of $u' = f(u)$ which satisfies $u_\nu(t_\nu) = u_\nu$; this is of course a smooth solution remaining in $\tilde{\mathcal{M}}$. We define the auxiliary point

$$\zeta_{\nu-1} := u_\nu - h\underbrace{f(u_\nu)}_{=u_\nu'(t_\nu)}, \tag{3.11}$$

such that u_ν can be interpreted as the result of an implicit Euler step starting from $\zeta_{\nu-1}$. The quantity

$$\tau_{\nu-1} := \zeta_{\nu-1} - u_\nu(t_{\nu-1}) = u_\nu - u_\nu(t_{\nu-1}) - hu_\nu'(t_\nu) \tag{3.12}$$

is simply a smooth $O(h^2)$-Taylor remainder term satisfying $\|\tau_{\nu-1}\| \le \frac{\tilde{M}_2}{2}h^2 = \mathcal{C}h^2$ (cf. (2.7)).

Furthermore we introduce the auxiliary quantity

$$\chi_{\nu-1} := \zeta_{\nu-1} - u_{\nu-1} = \zeta_{\nu-1} - u(\eta_{\nu-1}). \tag{3.13}$$

The relevant quantities are visualized in Fig. 3.1 for the case $n = 3$. Note that Fig. 3.1 is not quantitatively correct but has to be understood in a schematic sense. For $1/|\bar{\lambda}| \ll h$, in particular, distance$(\eta_{\nu-1}, \tilde{\mathcal{M}})$ and distance$(\eta_\nu, \tilde{\mathcal{M}})$ are very small compared to $\|\zeta_{\nu-1} - v_{\nu-1}\| = O(h^2)$ (cf. (3.22) below), and $\eta_\nu - u_\nu$ is significantly smaller than $\eta_{\nu-1} - \zeta_{\nu-1}$.

• *Recursion for the smooth error component δ_ν:*
Using (2.6), we obtain

$$\begin{aligned} \|\delta_\nu\| &= \|u_\nu - \tilde{u}(t_\nu)\| \le e^{\tilde{m}h}\|u_\nu(t_{\nu-1}) - \tilde{u}(t_{\nu-1})\| \\ &\le e^{\tilde{m}h}\Big(\|u_\nu(t_{\nu-1}) - \zeta_{\nu-1}\| + \|\zeta_{\nu-1} - u_{\nu-1}\| + \|u_{\nu-1} - \tilde{u}(t_{\nu-1})\|\Big) \\ &= e^{\tilde{m}h}\|\delta_{\nu-1}\| + e^{\tilde{m}h}\Big(\|\tau_{\nu-1}\| + \|\chi_{\nu-1}\|\Big) \end{aligned} \tag{3.14}$$

(cf. (3.11)–(3.13)). A standard argument leads us to[5]

$$\begin{aligned} \|\delta_\nu\| &\le e^{\tilde{m}t_\nu}\|\delta_0\| + \frac{e^{\tilde{m}t_\nu} - 1}{e^{\tilde{m}h} - 1} e^{\tilde{m}h} \max_{0\le k<\nu}\Big(\|\tau_k\| + \|\chi_k\|\Big) \\ &\le e^{\tilde{m}t_\nu}\|\delta_0\| + \frac{e^{\tilde{m}t_\nu} - 1}{\tilde{m}}\frac{\mathcal{C}}{h} \max_{0\le k<\nu}\Big(\|\tau_k\| + \|\chi_k\|\Big), \end{aligned} \tag{3.15}$$

[5]For $\tilde{m}$ positive, $\mathcal{C}$ in (3.15) is a bound for $e^{\tilde{m}h}$ and depends on the maximal admitted stepsize h_{max}. Note that h_{max} need not be chosen prohibitively small because $\tilde{m}$ is assumed to be moderate-sized.

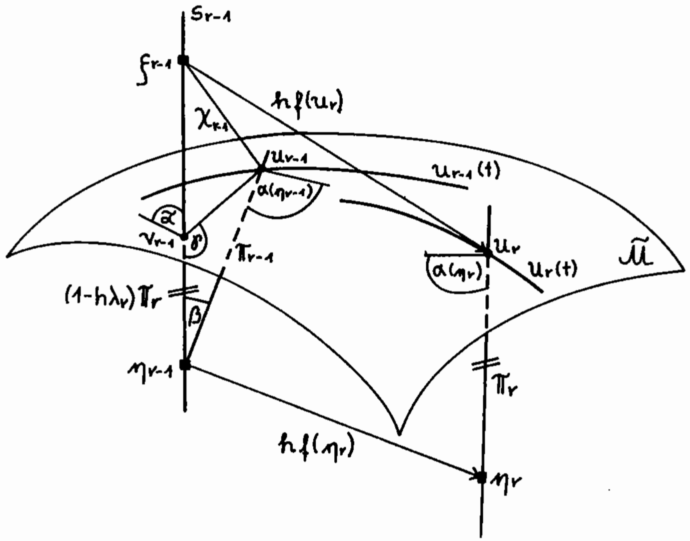

Figure 3.1: Local visualization of the error induction.

where $\|\delta_0\| = \|u(\eta_0) - \tilde{u}(0)\| \le Ch$ as a consequence of (3.2), and $\|\tau_k\| \le Ch^2$ for all k (cf. (3.12)).

• *Recursion for the stiff error component π_ν:*

By (2.8) and by definition of π_ν (cf. (3.10)) we have

$$f(\eta_\nu) - f(u_\nu) = \underbrace{\lambda(\eta_\nu)}_{=:\lambda_\nu}\pi_\nu. \tag{3.16}$$

Together with (3.11) this leads to

$$\begin{aligned} \pi_\nu &= \eta_\nu - u_\nu = \eta_{\nu-1} + hf(\eta_\nu) - \zeta_{\nu-1} - hf(u_\nu) \\ &= \underbrace{\eta_{\nu-1} - u_{\nu-1}}_{=\pi_{\nu-1}} + \underbrace{u_{\nu-1} - \zeta_{\nu-1}}_{=-\chi_{\nu-1}} + h\lambda_\nu\pi_\nu; \end{aligned} \tag{3.17}$$

thus we obtain a recursion for the π_ν in the form

$$\pi_\nu = \theta_\nu\pi_{\nu-1} - \theta_\nu\chi_{\nu-1}, \tag{3.18}$$

where (cf. (2.8))

$$\theta_\nu := \frac{1}{1-h\lambda_\nu} \le \frac{1}{1-h\bar{\lambda}} =: \bar{\theta}. \tag{3.19}$$

• *Estimation of the auxiliary quantity $\chi_{\nu-1}$:*

The recursions (3.15) and (3.18) for the smooth and stiff error components are coupled via the quantities $\chi_{\nu-1}$ (cf. (3.13)). To estimate $\chi_{\nu-1}$ we define the auxiliary point $v_{\nu-1} \in \tilde{\mathcal{M}}$ as the intersection of $\tilde{\mathcal{M}}$ with the straight line $s_{\nu-1}$ connecting $\eta_{\nu-1}$ and $\zeta_{\nu-1}$ (cf. Fig. 3.1), and split according to

$$\|\chi_{\nu-1}\| = \|\zeta_{\nu-1} - u_{\nu-1}\| \le \|\zeta_{\nu-1} - v_{\nu-1}\| + \|v_{\nu-1} - u_{\nu-1}\|. \tag{3.20}$$

To estimate the terms on the right hand side of (3.20) we make use of the smoothness of $\tilde{\mathcal{M}}$ (cf. Section 2, Propositions 1 & 2):

- To estimate $\|\zeta_{\nu-1} - v_{\nu-1}\|$ with the help of Proposition 1, we identify the quantities $\zeta_{\nu-1}$ and $v_{\nu-1}$ with ζ and v from (2.3). By definition (cf. (3.11)), $\zeta_{\nu-1} \in \mathcal{T}(u_\nu)$ and $\|\zeta_{\nu-1} - u_\nu\| = \|hf(u_\nu)\| \le h\tilde{M}_1$ since $u_\nu \in \tilde{\mathcal{M}}$ (cf. (2.7)). The point u_ν corresponds to u from Proposition 1. Furthermore, $\sphericalangle(s_{\nu-1}, \mathcal{T}(u_\nu))$ corresponds to α in Proposition 1. But (3.1), (3.11) and (2.8) imply that

$$\eta_{\nu-1} - \zeta_{\nu-1} = (1-h\lambda_\nu)(\eta_\nu - u_\nu) = (1-h\lambda_\nu)\pi_\nu$$

holds; hence $s_{\nu-1}$ (corresponding to s in Proposition 1) is parallel to π_ν, and thus,

$$\sphericalangle(s_{\nu-1}, \mathcal{T}(u_\nu)) = \sphericalangle(\pi_\nu, \tilde{\mathcal{M}}) = \alpha(\eta_\nu), \tag{3.21}$$

which is significantly away from 0 due to assumption (2.10).

Thus, with the above identifications the situation is exactly as described in Proposition 1, with $\rho = h\tilde{M}_1$, and (2.3) implies

$$\|\zeta_{\nu-1} - v_{\nu-1}\| \le Ch^2. \tag{3.22}$$

- To estimate $\|v_{\nu-1} - u_{\nu-1}\|$ we split into

$$\|v_{\nu-1} - u_{\nu-1}\| \le \|\eta_{\nu-1} - v_{\nu-1}\| + \|\underbrace{\eta_{\nu-1} - u_{\nu-1}}_{=\pi_{\nu-1}}\|. \tag{3.23}$$

In order to verify

$$\|\eta_{\nu-1} - v_{\nu-1}\| \le C\,\|\pi_{\nu-1}\| \tag{3.24}$$

we apply Propositions 1 and 2, identifying $\sphericalangle(s_{\nu-1}, \tilde{\mathcal{M}})$ with $\tilde{\alpha}$ from (2.4) and $\eta_{\nu-1}$ with η from (2.5):

Due to (2.4) and the fact that $\alpha = \alpha(\eta_\nu) = \sphericalangle(s_{\nu-1}, \mathcal{T}(u_\nu))$ is significantly away from 0 (cf. (3.21)), the same is true for $\tilde{\alpha} = \sphericalangle(s_{\nu-1}, \tilde{\mathcal{M}})$. Thus the assumptions of Proposition 2 are satisfied, and (2.5) yields

$$\|\eta_{\nu-1} - v_{\nu-1}\| \le C\cdot\text{distance}(\eta_{\nu-1}, \tilde{\mathcal{M}}), \tag{3.25}$$

from which (3.24) trivially follows. Together with (3.23) this yields

$$\|v_{\nu-1} - u_{\nu-1}\| \le C\|\pi_{\nu-1}\|. \tag{3.26}$$

(3.20), (3.22) and (3.26) imply the desired estimate for $\chi_{\nu-1}$:

$$\|\chi_{\nu-1}\| \le C\|\pi_{\nu-1}\| + Ch^2. \tag{3.27}$$

- *Estimation of the global error:*

  - First we estimate the stiff error component π_ν: Together with (3.27), the recursion (3.18) yields (for the definition of $\bar\theta$ cf. (3.19)):

$$\|\pi_\nu\| \le \bar\theta\Big(\|\pi_{\nu-1}\| + C\|\pi_{\nu-1}\| + Ch^2\Big) \le C\bar\theta\,\|\pi_{\nu-1}\| + C\bar\theta\,h^2. \tag{3.28}$$

Thus,

$$\|\pi_\nu\| \le (C\bar\theta)^\nu\|\pi_0\| + (C\bar\theta)h^2\sum_{k=0}^{\nu-1}(C\bar\theta)^k. \tag{3.29}$$

From (3.29), the desired estimate for $\|\pi_\nu\|$ can be derived for stepsizes $h \ge h_{min}$ where h_{min} has to be chosen such that (with the maximal C occurring in (3.29))

$$C\bar\theta = C\frac{1}{1-h\bar\lambda} \le C\frac{1}{1-h_{min}\bar\lambda} =: \rho \le \frac{1}{2}. \tag{3.30}$$

Clearly, $h_{min} = O(1/|\bar\lambda|)$. For such $h \ge h_{min}$ we may argue as follows: From (3.29),

$$\begin{aligned}\|\pi_\nu\| &\le \rho^\nu\|\pi_0\| + C\bar\theta\,h^2\frac{1-\rho^\nu}{1-\rho}\\ &\le \Big(\frac{1}{2}\Big)^\nu\|\pi_0\| + C\frac{1}{1+h|\bar\lambda|}h^2\cdot 2 \le \Big(\frac{1}{2}\Big)^\nu\|\pi_0\| + C\frac{h}{|\bar\lambda|}.\end{aligned} \tag{3.31}$$

In (3.3) we have introduced the quantity K_{min}. However, the precise choice for K_{min} has not been a-priori specified. The present Proof 1 covers only the case where K from the inductive assumption (3.2) is chosen not too large: Let

$$K := K_{min} := 2C; \tag{3.32}$$

here, C has to be understood as that particular generic constant which has appeared as a factor in the last term of (3.31). With this choice for K, the inductive assumption (3.2b) and (3.31) yield

$$\|\pi_\nu\| \le \Big(\frac{1}{2}\Big)^\nu K\frac{h}{|\bar\lambda|} + \frac{K}{2}\frac{h}{|\bar\lambda|} \le K\frac{h}{|\bar\lambda|}, \quad \nu = 1,2,\ldots, \quad h \ge h_{min}, \tag{3.33}$$

which proves (3.6).

Since $h_{min} = O(1/|\bar\lambda|)$ (cf. (3.30)), this also yields

$$\|\pi_\nu\| \le KCh_{min}h \le Ch^2, \quad \nu = 0,1,2,\ldots. \tag{3.34}$$

  - Due to (3.27), (3.34) also immediately yields

$$\|\chi_\nu\| \le Ch^2, \quad \nu = 0,1,2,\ldots, \tag{3.35}$$

and using (3.15) the smooth error component δ_ν can thus be estimated by

$$\|\delta_\nu\| \le e^{\tilde m t_\nu}\|\delta_0\| + \frac{e^{\tilde m t_\nu}-1}{\tilde m}Ch. \tag{3.36}$$

  - Finally, (3.34) and (3.36) yield

$$\|\eta_\nu - \tilde u(t_\nu)\| \le \|\pi_\nu\| + \|\delta_\nu\| \le Ch, \tag{3.37}$$

which proves (3.4) for $h \ge h_{min}$ ($h_{min} = O(1/|\bar\lambda|)$ as specified above; cf. (3.30)). □

PROOF 2: Essentially, Proof 2 is an extension of Proof 1. Most arguments used in Proof 1 remain valid, except those where the assumption 'strongly stiff' (i.e., the requirement $h \ge h_{min}$) was used (for the precise definition of h_{min} cf. (3.30)). In the sequel we only present those arguments necessary to show that our results are also valid for $0 < h < h_{min}$.

In particular, (3.8)–(3.19) remain valid. We use the same splitting of the global error into a smooth component δ_ν and a stiff component π_ν satisfying the inductive relations (3.15) and (3.18), coupled by $\chi_{\nu-1}$ (cf. (3.13)).

- *Estimation of the auxiliary quantity $\chi_{\nu-1}$:*

Basically, the estimates (3.20)–(3.27) remain valid, too; but the estimate (3.27) for $\chi_{\nu-1}$ is not sufficient to prove convergence in the present context (mildly stiff case). In the following, (3.27) will be used as an a-priori relation enabling a more refined estimate for $\chi_{\nu-1}$ on the basis of assumption (2.12).

As in Proof 1, let $s_{\nu-1}$ denote the straight line connecting $\eta_{\nu-1}$ and $\zeta_{\nu-1}$, intersecting $\tilde{\mathcal M}$ in a point denoted by $v_{\nu-1}$. Furthermore, the notation indicated in Fig. 3.1 is used. We denote

$$\begin{aligned}\beta &:= \sphericalangle(\pi_{\nu-1},\pi_\nu),\\ \gamma &:= \sphericalangle(\eta_{\nu-1}-v_{\nu-1}, v_{\nu-1}-u_{\nu-1}) = \sphericalangle(s_{\nu-1}, v_{\nu-1}-u_{\nu-1}).\end{aligned} \tag{3.38}$$

Basically, $\|\chi_{\nu-1}\|$ is estimated on the basis of (3.20), (3.22); to obtain a sharper estimate for $\|v_{\nu-1} - u_{\nu-1}\|$ than in (3.26) we proceed from the relation

$$\|v_{\nu-1} - u_{\nu-1}\| = \left|\frac{\sin\beta}{\sin\gamma}\right|\,\|\pi_{\nu-1}\| \le \left|\frac{\beta}{\sin\gamma}\right|\,\|\pi_{\nu-1}\|, \tag{3.39}$$

and estimate the factor $|\frac{\beta}{\sin\gamma}|$ in the following way:

  - By definition, β is the angle between the stiff directions $\pi_{\nu-1}$ and π_ν; thus, assumption (2.12) yields

$$\begin{aligned}|\beta| &\le C\|\eta_\nu - \eta_{\nu-1}\| = C\|hf(\eta_\nu)\|\\ &\le C\Big(h\,\|f(u_\nu)\| + h|\lambda_\nu|\,\|\pi_\nu\|\Big) \le C\Big(h\tilde M_1 + h|\lambda_\nu|\,\|\pi_\nu\|\Big),\end{aligned} \tag{3.40}$$

where (cf. (3.18), (3.19), (3.27))

$$\|\pi_\nu\| \le \theta_\nu \|\pi_{\nu-1}\| + \theta_\nu \|\chi_{\nu-1}\| \le \theta_\nu (C\|\pi_{\nu-1}\| + Ch^2). \tag{3.41}$$

Now (3.40) and (3.41) yield

$$|\beta| \le Ch + C\|\pi_{\nu-1}\|. \tag{3.42}$$

– To estimate $1/|\sin\gamma|$ we may argue as follows: With $\tilde{\alpha} = \sphericalangle(s_{\nu-1}, \tilde{\mathcal{M}})$, (2.4) from Proposition 1 (again identifying ρ with $h\tilde{M}_1$) and (2.10) yield

$$\frac{1}{|\sin\tilde{\alpha}|} \le C. \tag{3.43}$$

Consider the plane spanned by $s_{\nu-1}$ and $v_{\nu-1} - u_{\nu-1}$, and denote its intersection curve with $\tilde{\mathcal{M}}$ by κ. Proposition 2 together with (3.43) imply that κ is smooth. Let $\hat{\alpha} := \sphericalangle(s_{\nu-1}, \kappa)$. Clearly,[6] $\sin\hat{\alpha} \ge \sin\tilde{\alpha}$, so (3.43) implies

$$\frac{1}{|\sin\hat{\alpha}|} \le C. \tag{3.44}$$

The desired estimate for $1/|\sin\gamma|$ can now be obtained as follows: $v_{\nu-1} - u_{\nu-1}$ is a secant to the smooth curve κ, and therefore the difference between $\hat{\alpha}$ and γ is bounded by $C \times$ the length of this secant, which has already been estimated in Proof 1, (3.26). Together with (3.44) this leads us to

$$\frac{1}{|\sin\gamma|} \le \frac{1}{|\sin\hat{\alpha}|} + C\|v_{\nu-1} - u_{\nu-1}\| \le C + C\|\pi_{\nu-1}\|. \tag{3.45}$$

From (3.39), (3.42) and (3.45) we conclude that the quantity $v_{\nu-1} - u_{\nu-1}$ can be estimated by

$$\|v_{\nu-1} - u_{\nu-1}\| \le Ch\|\pi_{\nu-1}\| + C\|\pi_{\nu-1}\|^2 + C\|\pi_{\nu-1}\|^3, \tag{3.46}$$

which is sharper than (3.27).

Together with (3.20) and (3.22), (3.46) implies

$$\|\chi_{\nu-1}\| \le Ch\|\pi_{\nu-1}\| + C\|\pi_{\nu-1}\|^2 + C\|\pi_{\nu-1}\|^3 + Ch^2. \tag{3.47}$$

• *Estimation of the global error:*

– The essential point is the inductive estimation of the stiff error component π_ν: The goal is to show that

$$\|\pi_{\nu-1}\| \le K\frac{h}{|\bar{\lambda}|} \tag{3.48}$$

[6] Note that $\tilde{\alpha}$ is the minimum of all angles between $s_{\nu-1}$ and arbitrary smooth curves on $\tilde{\mathcal{M}}$ transversing $v_{\nu-1}$.

implies the same estimate for π_ν, i.e.,

$$\|\pi_\nu\| \le K\frac{h}{|\bar{\lambda}|} \tag{3.49}$$

with the *same* constant $K \in [K_{min}, |\bar{\lambda}|]$, where the precise value of K_{min} remains to be specified. Thus—once (3.49) has been verified—we can conclude that the π_ν are uniformly bounded at the $K\frac{h}{|\bar{\lambda}|}$-level provided π_0 satisfies $\|\pi_0\| \le K\frac{h}{|\bar{\lambda}|}$ with this particular constant K. To obtain such a recursive estimate for π_ν we proceed from (3.18) and use (3.47):

$$\|\pi_\nu\| \le \bar{\theta}\Big((1+Ch)\|\pi_{\nu-1}\| + C\|\pi_{\nu-1}\|^2 + C\|\pi_{\nu-1}\|^3 + Ch^2\Big). \tag{3.50}$$

Using $K \le |\bar{\lambda}|$ and the inductive assumption (3.48) we obtain

$$\begin{aligned}\|\pi_\nu\| &\le \bar{\theta}\Big((1+Ch)K\frac{h}{|\bar{\lambda}|} + C\Big(K\frac{h}{|\bar{\lambda}|}\Big)^2 + C\Big(K\frac{h}{|\bar{\lambda}|}\Big)^3 + Ch^2\Big)\\ &\le \frac{1}{1-h\bar{\lambda}}\Big((1+Ch)K\frac{h}{|\bar{\lambda}|} + Ch^2\Big)\\ &= K\frac{h}{|\bar{\lambda}|}\Big(\frac{1+Ch}{1-h\bar{\lambda}} + \frac{|\bar{\lambda}|}{K}\frac{Ch}{1-h\bar{\lambda}}\Big) = K\frac{h}{|\bar{\lambda}|}\frac{1+h|\bar{\lambda}|(\frac{C}{|\bar{\lambda}|}+\frac{C}{K})}{1+h|\bar{\lambda}|}.\end{aligned} \tag{3.51}$$

Now we are in a position to specify the quantity K_{min}: We choose

$$K_{min} := 2C; \tag{3.52}$$

here, C has to be understood as the maximum of those generic constants which have appeared in the last term of (3.51). With this choice for K_{min}, both factors $C/|\bar{\lambda}|$ and C/K appearing in the last term of (3.51) can be bounded by $1/2$ for all K satisfying $K_{min} \le K \le |\bar{\lambda}|$, showing that (3.49) is indeed satisfied for such K. This proves (3.6) for all $K \in [K_{min}, |\bar{\lambda}|]$, for $0 < h \le h_{max}$.

– Due to (3.49), (3.47) also immediately yields

$$\|\chi_{\nu-1}\| \le CK\frac{h^2}{|\bar{\lambda}|} + C\Big(K\frac{h}{|\bar{\lambda}|}\Big)^2 + C\Big(K\frac{h}{|\bar{\lambda}|}\Big)^3 + Ch^2 \le Ch^2 \tag{3.53}$$

for any choice of K admitted above.

The rest of the proof is analogous as in Proof 1 (cf. (3.36) ff.): Due to (3.15) and (3.53), the smooth error component δ_ν can be estimated by

$$\|\delta_\nu\| \le e^{\tilde{m}t_\nu}\|\delta_0\| + \frac{e^{\tilde{m}t_\nu} - 1}{\tilde{m}}Ch. \tag{3.54}$$

– Finally, (3.49) and (3.54) yield the global error estimate

$$\|\eta_\nu - \tilde{u}(t_\nu)\| \le \|\delta_\nu\| + \|\pi_\nu\| = O(h) \tag{3.55}$$

for all $0 < h \le h_{max}$. □

REFERENCES

1. W. Auzinger, R. Frank, and G. Kirlinger, *Extending convergence theory for nonlinear stiff problems. Part I*, Report 116/94, Institute for Applied and Numerical Mathematics, Vienna University of Technology (extended version of the present paper).
2. W. Auzinger, R. Frank, and G. Kirlinger, *A note on convergence concepts for stiff problems*, Computing 44 (1990), pp. 197–208.
3. K. Dekker and J. G. Verwer, *Stability of Runge-Kutta Methods for Stiff Nonlinear Differential Equations*, North-Holland, 1984.
4. R. Frank and C. W. Ueberhuber, *Iterated defect-correction for the efficient solution of stiff systems of ordinary differential equations*, BIT 17 (1977), pp. 146–159.
5. R. Frank, J. Schneid, and C. W. Ueberhuber, *The concept of B-convergence*, SIAM J. Numer. Anal. 18 (1981), pp. 753–780.
6. R. Frank, J. Schneid, and C. W. Ueberhuber, *Stability properties of implicit Runge-Kutta methods*, SIAM J. Numer. Anal. 22 (1985), pp. 497–515.
7. R. Frank, J. Schneid, and C. W. Ueberhuber, *Order results for implicit Runge-Kutta methods applied to stiff systems*, SIAM J. Numer. Anal. 22 (1985), pp. 515–534.
8. E. Hairer, C. Lubich, M. Roche, *Error of Runge-Kutta methods for stiff problems studied via differential algebraic equations*, BIT 28 (1988), pp. 678–700.
9. E. Hairer and G. Wanner, *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems*. Springer, 1991.
10. Ch. Lubich, *On the convergence of multistep methods for nonlinear stiff differential equations*, Numer. Math. 58 (1991), pp. 839–853; Erratum, 61 (1992), pp. 277–279.
11. K. Nipp and D. Stoffer, *Invariant manifolds and global error estimates of numerical integration schemes applied to stiff systems of singular perturbation type*, Numer. Math., 70 (1995), pp. 245–257.
12. A. Prothero and A. Robinson, *On the stability and accuracy of one-step methods for solving stiff systems of ordinary differential equations*, Math. Comp. 28 (1974), pp. 145–162.