THE STATE OF THE ART IN THE CONVERGENCE THEORY FOR STIFF ODES

W. Auzinger, R. Frank and G. Kirlinger

Institut für Angewandte und Numerische Mathematik
Technische Universität Wien
Wiedner Hauptstrasse 8–10/115
A-1040 Wien, Austria

Abstract - In this paper we give a brief review on available theoretical results about convergence and error structures for discretizations of stiff initial value problems. We point out limitations of the various approaches and discuss some recent developments.

I. CURRENTS IN THE EXISTING CONVERGENCE THEORY

More or less efficient codes have been used more or less successfully for the numerical solution of stiff ODEs; at present backward differentiation schemes (BDF) are without doubt among the most widely used methods. However, as it is often the case in numerical mathematics and its applications, the theoretical foundations are still incomplete. Many problems which arise in practice - and which are solved numerically - are of such a complexity that the available theory (e.g. concerning convergence properties or error structures of the methods used) is not applicable. Algorithms and control mechanisms used within the respective codes are often based on model concepts or heuristic principles only. This is obviously reasonable from a practical point of view; but a convergence theory as universal as possible is of course desirable.

Below we give a brief review on existing theoretical approaches and results about convergence and error structures for discretizations of stiff initial value problems

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

$$y(0) = y_0. \tag{1.1b}$$

We point out some limitations of the various approaches and discuss some recent developments.

Historically, the first theoretical concept especially suited for the assessment of numerical methods for stiff problems was A-stability (cf. [13]), resp. A(α)-stability (cf. [32]). Based on this model concept, rigorous conclusions can only be drawn for linear constant coefficient problems $y' = Jy$. Nevertheless it has been used, with doubtful reliability, as a guideline to select methods for the solution of more difficult problems. First serious attempts to develop a rigorous, comprehensive convergence theory for nonlinear stiff problems go back to the early Seventies. In 1975, G. Dahlquist [14] presented a stability concept (G-stability) which is in a sense natural, as explained in the following.

Let us start by considering condition estimates for (1.1). Already in the Fifties (cf. for instance [12]) it had been pointed out that the so-called *logarithmic norm* $\mu(f)$ is a canonical parameter in the assessment of the condition behavior of an initial value problem (1.1). (For the definition of $\mu(f)$ see e.g. [16], and section II. below.) Namely, it enables estimates for the effect of perturbations (of the initial value or of the direction field) which are, in a certain sense, sharp (cf. the discussion in section II.). Consequently, these estimates have been considered as a natural improvement over the well-known 'classical' estimates based on a Lipschitz constant $L(f)$ for f, the improvement being dramatic for such stiff problems where $L(f) \gg \mu(f)$ (recall that, inevitably, $L(f) \gg 0$ for stiff problems in spite of their good condition). In the analysis of discretization methods, it is equally natural to strive for stability estimates also based on the problem-characterizing parameter $\mu(f)$ to overcome the obvious fact that classical stability inequalities based on $L(f) \gg 0$ are of no use in stiff situations. In the concept of G-stability, this idea was realized for multistep methods. For one-step methods (in particular implicit Runge-Kutta methods), Butcher [10] introduced an analogous stability concept called B-stability.

Norms induced by a scalar product $\langle \cdot, \cdot \rangle$ turned out to be particularly convenient. In this case the logarithmic norm $\mu(f)$ can be expressed as the best possible *one-sided Lipschitz constant* for f, i.e. the smallest real number m for which

$$\langle f(t,y_1) - f(t,y_2),\, y_1 - y_2 \rangle \le m \, \|y_1 - y_2\|^2 \tag{1.1}$$

(for all y_1, y_2 in the domain under consideration). Let us now, for example, recall the precise definition of B-stability: Assume that m is a one-sided Lipschitz constant for f, and consider two 'parallel' steps $(t_{\nu-1}, \eta_{\nu-1}) \to (t_\nu, \eta_\nu)$ and $(t_{\nu-1}, \tilde{\eta}_{\nu-1}) \to (t_\nu, \tilde{\eta}_\nu)$ of a one-step method applied to (1.1). The method is called B-stable if there holds

$$\|\eta_\nu - \tilde{\eta}_\nu\| \le \Phi(hm) \, \|\eta_{\nu-1} - \tilde{\eta}_{\nu-1}\| \tag{1.2}$$

where $h = t_\nu - t_{\nu-1}$ is the steplength, and where Φ a smooth function satisfying $\Phi(0) = 1$. Note that (1.2) is an immediate discrete analogon to the condition inequality

$$\|y(t+h) - \tilde{y}(t+h)\| \le e^{hm} \, \|y(t) - \tilde{y}(t)\| \tag{1.3}$$

where $y, \tilde{y}$ denote a pair of solutions to (1.1a) (see e.g. [17]).

B-stability turned out to be a very successful concept; in particular, several classes of implicit Runge-Kutta methods have been shown to be B-stable. The essential technical tool for the derivation of stability results are certain algebraic conditions on the Runge-Kutta coefficients like *algebraic stability* and *diagonal stability* (cf. for instance [9], [11] and [19]). For multistep methods it turns out that, unfortunately, G-stability is a rather restrictive requirement in the following sense: A-stability is of course a necessary condition and therefore, due to well-known order barriers, a G-stable multistep method cannot have an order of consistency higher than 2.

For higher stage one-step methods it is important to notice that - apart from stability questions - also the local error analysis is by no means trivial: For implicit Runge-Kutta methods the local error (i.e. the error induced by a single integration step) is a complicated expression involving various derivatives of the right hand side f, the norm of which is inevitably affected by large problem parameters like $L(f)$, and therefore the actual magnitude of the local error is not a priori obvious. This difficulty had almost consistently be ignored in the earlier literature (see, however, [27]). Frank, Schneid & Ueberhuber were the first to give a systematical local error analysis (see [18],[20]). The point is that, from a reasonable *quantitative* standpoint, an assertion like 'the order of the local error is $O(h^p)$' makes only sense if the O-constant is not influenced by prohibitively large problem parameters - but only by the local smoothness of the ODE solution (the latter dependence is quite natural). It turns out that, in this quantitative order concept, the local error of an implicit Runge-Kutta usually suffers from a *order reduction* (compared to the classical, non-stiff order). In general, the order actually observed reduces to the so-called stage order, i.e. the order of the local truncation error. But even that the stage order can be achieved is by no means simple to prove and requires a careful 'internal' stability analysis (concept of BS-stability). The concepts of B-stability and BS-stability led to a quantitative convergence theory (B-convergence) for stiff problems satisfying a one-sided Lipschitz con-

dition (see [18], [20]). A typical B-convergence result reads

$$\|\eta_\nu - y(t_\nu)\| \le C(m, M_\ell)h^p \tag{1.5}$$

where the quantitity $C(m, M_\ell)$ depends only on the one-sided Lipschitz constant m and on bounds M_ℓ for the derivatives of $y(t)$ entering the expression for the local error. In the last years, a large number of results concerning B-stability (and related stability concepts) and B-convergence have been derived for various types of implicit Runge-Kutta methods (for an overview cf. for instance [16], [21]).

An alternative approach towards a convergence analysis for stiff problems is based on the concepts of *singular perturbations*. A standard problem considered in singular perturbation theory is

$$u' = \varphi(t, u, v), \tag{1.6a}$$
$$\varepsilon v' = \psi(t, u, v) \tag{1.6b}$$

with

$$\mu\Big(\frac{\partial \psi}{\partial v}\Big) \le -\kappa < 0. \tag{1.7}$$

$0 < \varepsilon \ll 1$ is a small parameter characterizing the stiffness. A problem of this form has a quite particular structure: There is only one 'cluster' of stiff eigenvalues (of magnitude $-1/\varepsilon$); furthermore, the dependence of the right hand side on the stiffness parameter ε is rather special. The behavior of solutions to (1.6) is well understood; it can be analyzed by asymptotic methods and the necessary condition estimates can e.g. be derived using contraction arguments, without relying on a one-sided Lipschitz condition for the right hand side (cf. e.g. [8]). In the last years, convergence results have been derived for implicit Runge-Kutta methods and, recently, also for BDF methods applied to (1.6) (see [23], [25]). These results constitute a nontrivial extension of the convergence theory (note that the problem class (1.6) is not satisfactorily covered by the B-convergence theory - cf. the discussion in section II.).

Stiffness also frequently occurs in problems connected with partial differential equations (e.g. for parabolic initial/boundary value problems); often a stiff ODE system arises after a PDE has been discretized in space. The one-sided Lipschitz constant m frequently appears in a natural way (often it has a direct physical meaning, e.g., $m \le 0$ may indicate dissipation of energy). Therefore the concepts and results of the B-theory are often useful here; for an overview of results cf. for instance [28]. Also certain modifications of the one-sided Lipschitz condition are of relevance; in the analysis of non-selfadjoint parabolic equations, for instance, a strengthened one-sided Lipschitz condition (the so-called 'sectorial condition') plays an important role (cf. for instance [26], [2] for its significance in the analysis of numerical methods). As a word of warning, however, it must be said that all these concepts are likely to fail (or, at least, be too weak) for PDEs involving significant nonlinearities.

Concerning the *structure* of discretization errors (i.e., asymptotic error expansions in powers of the stepsize), useful results have been derived only recently for special classes of stiff problems (e.g. for class (1.6) and simple one-step methods (see [1]–[5], [22], [31]). It turned out that the usual arguments (following the lines of the general procedure described by Stetter [30]) are of no use in the stiff case. Already for very simple schemes as e.g. implicit Euler or implicit midpoint rule, the global error cannot be described by smooth functions, but strongly varying (decaying or oscillating) components may play a dominant role. The particular error structure strongly depends on the type of problem under consideration, on problem parameters like for instance the magnitude of the stiff eigenvalues, and on the method and stepsize actually used.

These results about error structures form the basis for the analysis of extrapolation or defect correction methods and for a sound justification of stepsize control mechanisms; but the theory is far from complete. For multistep schemes, results about error structures in the stiff case do not seem to exist.

II. NORMS, LOGARITHMIC NORMS AND ONE-SIDED LIPSCHITZ CONSTANTS: A CRITICAL DISCUSSION

The overview given in section I. shows that one-sided Lipschitz constants are an essential tool in the analysis of stiff problems; a large number of results are based on this concept. As already mentioned, the use of one-sided Lipschitz constants is, in a sense, natural because typical stability estimates like (1.3) are obvious discrete analoga of condition estimates of the form (1.4) for the solutions of the given ODE. Note that the estimate (1.4) is even *optimal in a local sense*, i.e., for arbitrary $y(t)$ there exists a perturbed $\bar{y}(t)$ such that the estimate (1.4) is asymptotically sharp for $h \to 0$ (cf. [16], [17]).

On the other hand, it is easy to write down examples which show an uncritical error sensitivity - in the sense that the global effect of perturbations remains moderately bounded - despite $m \gg 0$. Consider for instance the simple constant coefficient problem

$$y' = J\,y \qquad \text{with} \quad J = \begin{pmatrix} -1 & 0 \\ \frac{1}{\varepsilon} - 1 & -\frac{1}{\varepsilon} \end{pmatrix} \tag{2.1}$$

where $0 < \varepsilon \ll 1$. This is a linear problem of the type (1.6). The logarithmic norm $\mu_2(J)$ w.r.t. the Euclidean norm $\|\cdot\|_2$ (i.e. the best possible one-sided Lipschitz constant m for $f(t, y) = J\,y$ can be expressed as the spectral abscissa of $(J + J^T)/2$ (cf. e.g. [16]), and we easily obtain

$$\mu_2(J) = \frac{\sqrt{2}-1}{2} \cdot \frac{1}{\varepsilon} - \frac{\sqrt{2}+1}{2} = O\Big(+\frac{1}{\varepsilon}\Big). \tag{2.2}$$

Consequently, (2.1) is locally ill-conditioned: The worst propagation of a perturbation δ of the initial value y_0, measured in $\|\cdot\|_2$, qualitatively behaves like $e^{+t/\varepsilon}\|\delta\|_2$ for small t, despite the fact the eigenvalues of J are -1 and $-1/\varepsilon$. On the other hand, the global effect of such a perturbation is easily seen to be uniformly bounded by $\sqrt{2}\,\|\delta\|_2$ for

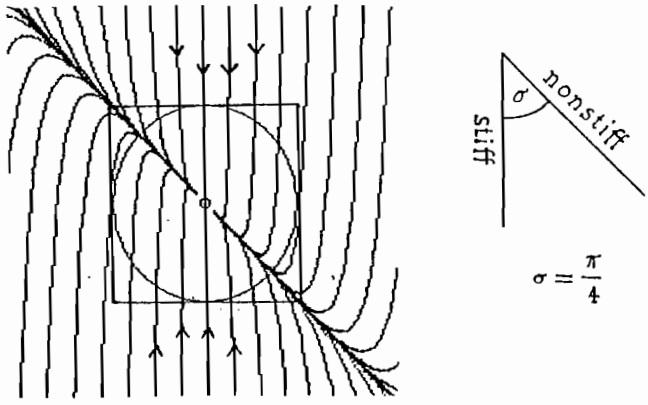

Figure 1: Phase portrait for example (2.1) (J nonsymmetric)

arbitrary $t > 0$. I.e., a rapid growth of perturbations can only occur during a very short transient phase.

To illustrate this situation we study the behavior of solutions to (2.1) by means of the phase portrait (cf. Figure). (By linearity, it is not necessary here to distinguish between solutions and perturbations.) W.l.o.g. we consider all y with $\|y\|_2 = 1$. For certain y the direction field points outward the unit circle; hence for a solution trajectory $y(t)$ passing through such a point the norm $\|y(t)\|_2$ increases. Moreover, the local growth of $\|y(t)\|_2$ is enormous due to the large size of the stiff eigenvalue. However, after a very short time interval the rapid variation vanishes and the solution trajectory smoothly follows the direction corresponding to the non-stiff eigenvalue.

This situation, namely that there is a dramatic discrepancy between local and global condition, is not exceptional but, in a sense, *typical:* It has been shown in [6] that for two-dimensional stiff pro-

blems

$$y' = J\,y = S\Lambda S^{-1}y \qquad \text{with} \quad \Lambda = \begin{pmatrix} c_1 & 0 \\ 0 & -\frac{c_2}{\varepsilon} \end{pmatrix} \qquad (2.3)$$

($c_2 \geq \kappa > 0$), the logarithmic norm $\mu_2(J)$ is moderate-sized (i.e., not affected by a factor $+1/\varepsilon$) if and only if J is 'nearly symmetric' in the sense that the angle between the eigendirections of J must be $\pi/2 + O(\varepsilon^{1/2})$. This can also directly be seen from Figure : Only if the eigendirections would be orthogonal (symmetric case) or at least 'nearly orthogonal', a locally strong increase of perturbations could not occur. (In example (2.1) we have $c_1 = -1$, $c_2 = 1$ and $\sigma = \pi/4$.) Thus, in the nonsymmetric case we inevitably must have $\mu_2(J) \gg 0$ due to the potentially strong local increase; but the good global condition - i.e., the fact that such a strong increase can only occur on a very short time interval - is not reflected by this local concept. (Only in the symmetric case there is no such difficulty.) Numerical experience indicates that a similar discrepancy is also typical for higher-dimensional stiff systems; however, to derive rigorous characterizations as for the two-dimensional case would be a highly nontrivial task. In the PDE field (initial/boundary value problems, method of lines context) examples are known where a strongly positive logarithmic norm is caused by certain types of boundary conditions (cf. e.g. [16], section 10.6). A comprehensive *general* characterization of cases where the logarithmic norm concept is reasonably applicable or not does not exist; but in view of the results of [6] we must exect that $\mu_2(f) \gg 0$ is often the case, i.e., that a strong discrepancy between local and global condition is often present. Naturally, B-convergence bounds based on $\mu_2(f)$ are of no use in all these situations.

This restricted applicability of results from the B-theory is not too surprising in the light of a result given in [9]: There it was shown for non-confluent Runge-Kutta methods that B-stability for problems with $m \leq 0$ is equivalent to AN-stability (the generalization of A-stability for scalar problems $y' = \lambda(t)y$). Thus a stability property for nonlinear problems is directly related to a stability property for scalar linear problems. A similar equivalence also holds in the context of multistep methods: A-stability is not only necessary but even sufficient for G-stability (cf. [15]). With this background one may say (and some people indeed do so) that B-stability and G-stability are scalar concepts which do not take account of time-dependent or nonlinear coupling.

The question arises how to circumvent this drawback. For constant coefficient problems like (2.3) this is simple: Instead of the Euclidean norm one chooses another scalar product norm, the so-called 'elliptic norm' $\|y\|_S := \|S^{-1}y\|_2$. This corresponds to a complete decoupling and, obviously, the corresponding logarithmic norm is simply $\mu_S(J) = \mu_2(\Lambda)$, the spectral abscissa of J. However, already for two-dimensional linear systems $y' = J(t)y$ with time dependent coefficients there is no such simple remedy because it turns out that, even for smoothly varying eigendirections, there is *no fixed elliptic norm* for which the corresponding logarithmic norm remains moderate-sized on a time interval of length $O(1)$ (cf. [6]). In other words: The logarithmic norm $\mu_S(J)$ is extremely sensitive w.r.t. rotations of the eigendirections of J; unless the norm $\|\cdot\|_S$ (or, equivalently, the unit ellipse $\{y : \|y\|_S = 1\}$) is chosen precisely adapted to these eigendirections, there always exist points y on that unit ellipse where for a solution trajectory $y(t)$ passing through this point $\|y(t)\|_S$ locally rapidly increases.

This latter observation causes surprising complications e.g. in the analysis of Runge-Kutta methods. Cf. section III., where we briefly review a recent result concerning the convergence of implicit Runge-Kutta methods applied to stiff problems with time dependent coefficients and weak nonlinearities. However, these results are not applicable to the more difficult case where also the nonlinear terms on the right hand side are affected by large parameters (such problems are, in their general form, neither covered by the B-theory).

As indicated above, the trick of working with the logarithmic norm corresponding to an appropriately modified scalar product norm (elliptic norm) is successful only for very simple stiff problems. It is, for instance, not applicable in the case of (1.6), where the Jacobian

$$J = \begin{pmatrix} \frac{\partial\varphi}{\partial u} & \frac{\partial\varphi}{\partial v} \\ \frac{1}{\varepsilon}\frac{\partial\psi}{\partial u} & \frac{1}{\varepsilon}\frac{\partial\psi}{\partial v} \end{pmatrix} \qquad (2.4)$$

is varying and nonsymmetric in general. Due to this 'lack of robustness' of scalar product norms the question arises whether *non-scalar product norms* are more suitable.

For an illustration of this point, let us now consider singularly perturbed equations of the form (1.6). This problem class is of course well understood in singular perturbation theory and there is, so to speak, nothing to show. Nevertheless we reconsider problems (1.6) here; they simply serve as as illustrative examples for our present discussion concerning the choice of norms.

Let us, for simplicity of presentation, assume that the problem is two-dimensional, i.e., φ and ψ in (1.6) are scalar functions. We also assume for the moment that $|\partial\psi/\partial u| \leq \kappa$ in some domain, where $\partial\psi/\partial v \leq -\kappa < 0$ (cf: condition (1.7)). Then, obviously, the logarithmic norm $\mu_\infty(J)$ w.r.t. the maximum norm is moderate-sized, i.e., not affected by a factor $+1/\varepsilon$. (Note that $\mu_\infty(A) = \max_i (a_{ii} + \sum_{j\neq i} |a_{ij}|)$, cf. [16].) Also the case where $|\partial\psi/\partial u|$ becomes larger than κ causes no difficulty: It is easy to find an appropriately scaled maximum norm $\|y\|_{\infty,D} := \|Dy\|_\infty$, $D > 0$ diagonal, such that the corresponding logarithmic norm $\mu_{\infty,D}(J)$ is moderate-sized (note that for $A = (a_{ij})$ and $D = \mathrm{Diag}(d_i)$, $\mu_{\infty,D}(A) = \mu_\infty(DAD^{-1}) = \max_i (a_{ii} + \sum_{j\neq i} |a_{ij}d_i/d_j|)$). The minimal amount of scaling necessary depends on the actual size of $|\partial\psi/\partial u|$ and is related to the global condition of the problem.

Geometrically, this can be interpreted as follows. Consider e.g. the simple linear example (2.1). Obviously, $\mu_\infty(J) = -1$, and therefore the problem is contractive in the maximum norm (i.e., the norm of or perturbations never grows). In Figure this property is reflected by the fact that at each point y with $\|y\|_\infty = 1$ the direction field points inward the unit square. A re-scaling of the norm would be necessary in cases where the angle σ between the eigenvectors of J would be smaller than $\pi/4$; geometrically speaking, the direction field points inward any rectangle sufficiently stretched in the vertical direction. In other words: The norm has to be chosen in such a way that a strong local increase of a solution component is 'invisible'.

The *really* essential point is the following. Recall that for the problem class under consideration, the logarithmic norm w.r.t. any scalar product norm inevitably gets positive and large unless the norm is very precisely adapted to the actual orientation of the eigendirections of the Jacobian J; this makes scalar product norms unsuitable for nonconstant J. (Note that the logarithmic norm of f in a certain domain is the supremum of the logarithmic norms of its Jacobian in that domain.) On the other hand, for each problem of the form (1.6) it is easy to find an appropriately scaled maximum norm such that the corresponding logarithmic norm remains moderate-sized not only for a single $J = J(t,u,v)$ but for certain $O(1)$-domains of arguments (t,u,v); one simply has to choose a sufficiently distorting scaling matrix (i.e., a sufficiently stretched unit rectangle), which is essentially determined by the maximal occurring value of $|\partial\psi/\partial u|$.

Geometrically, this can be interpreted as follows. A simple eigensystem analysis of (2.4) shows that a smooth variation of the occurring partial derivatives corresponds to a smooth ($O(1)$) rotation of the 'nonstiff eigendirection'; the 'stiff eigendirection' also rotates but only with a speed $O(\varepsilon)$ (the eigenvector of (2.4) corresponding to the stiff eigenvalue is always of the form $(O(\varepsilon),1)^T$). Now the amount of scaling necessary is simply determined by the smallest occurring angle between these eigendirections, and we finally can express the good problem condition with the help of the logarithmic norm w.r.t. a certain *fixed*, re-scaled maximum norm. It is important to notice here that the 'inertia' of the stiff eigendirection is essential; the above way of choosing the norm would break down in the case of a $O(1)$- (and

not only $O(\varepsilon)$-) rotation of the stiff eigendirection.

In view of the above consideration it may be hoped that there also exist other interesting nontrivial classes of stiff problems (e.g., certain multiparameter generalizations of (1.6)) for which a careful choice of norm leads to a moderate logarithmic norm; this is a question worth investigating. However, such a remedy will not always exist; it appears to be inpossible for highly nonlinear problems where, in contrast to (1.6), also the stiff eigendirections significantly vary.

Now, for such problem classes where an appropriate norm in the above sense exists, the question is whether it is possible to develop a convergence theory for discretization methods based on such a (non-scalar product) norm. Some results in this direction can be found in [29] and [24]. Unfortunately it turns out that non-scalar product norms are much less convenient to use in the convergence theory than scalar product norms. Furthermore, the results of [29] and [24] are of a rather negative flavour: It is shown that a one-step method that behaves contractively (w.r.t. $\|\cdot\|_\infty$) for any $\|\cdot\|_\infty$-contractive ODE cannot have an order higher than 1.

These results are disappointing; but one should bear in mind that they are formulated in a very general setting and that the requirement of strict contractivity (i.e. step-by-step stability with a stability threshold ≤ 1) is rather strong. Indeed, it is simple to write down examples where, for a concrete $\|\cdot\|_\infty$-contractive problem, a given Runge-Kutta method of an order higher than 1 (e.g. a 2-stage Radau Ia method) does also behave contractively - not necessarily w.r.t. the maximum norm itself but w.r.t. an appropriately re-scaled maximum norm. Thus it may be advisable to modify the norm for the analysis of the discrete scheme.

Summarizing all that, one may say that the choice of norm is crucial in the analysis of stiff problems. One may hope that a careful choice of norms will lead to interesting convergence results for some particular classes of stiff problems and particular discretization schemes.

III. AN EXTENSION OF B-CONVERGENCE

As explained in section II., the applicability of results from the B-theory is more restricted than usually believed. In particular, not even simple linear stiff systems $y' = J(t)y$ with time-dependent coefficients are satisfactorily covered: For any scalar product norm, the supremum of $\mu_2(J(t))$ on some time interval of length $O(1)$ is inevitably strongly positive (affected by the moduli of the stiff eigenvalues) except in special cases ($J(t)$ symmetric). But a theoretical concept like B-stability heavily relies on the existence of a suitable scalar product norm; therefore an immediate application of the concepts of the B-theory to such problems does not lead to reasonable convergence results. (Also the arguments from singular perturbation theory are only applicable in special cases.) In the following we briefly discuss this point further and review a recent result, given in [7], concerning the convergence of Runge-Kutta methods. The approach used in [7] is not based on non-scalar product norms but on time dependent coordinate transformations.

Problems of the form $y' = J(t)y$ can be studied by means of a suitable time-dependent change of coordinates. Assume for instance that $J(t)$ is diagonalizable, $J(t) = S(t)\,\Lambda(t)\,S^{-1}(t)$, with a well-conditioned, smoothly varying eigensystem $S(t)$. Then the time-dependent transformation $\bar{y} := S^{-1}(t)y$ leads to the transformed ODE

$$\bar{y}' = \bar{J}(t)\bar{y} \qquad \text{with} \quad \bar{J}(t) := \Lambda(t) - S^{-1}(t)S'(t) \tag{3.1}$$

(the eigenvalues of $\bar{J}(t)$ are sometimes called the kinematic eigenvalues of $J(t)$). This time-dependent change of coordinates may also be viewed as a time-dependent adaption of the norm (generalization of the concept of elliptic norms). Under appropriate smoothness assumptions w.r.t. $S(t)$, the logarithmic norm $\mu_2(\bar{J}(t))$ (i.e. the largest kinematic eigenvalue of $J(t)$) is obviously of the magnitude of the spectral abscissa of $J(t)$; it is not affected by prohibitively large problem parameters like the moduli of the stiff eigenvalues.

Let us now consider Runge-Kutta methods applied to (3.1). One could think that for a study of stability one would simply have to apply an analogous time-dependent transformation to the discrete scheme (including the internal Runge-Kutta stages) and to apply the B-theory to this transformed scheme. Unfortunately, such a procedure does not lead us to a useful result because the Runge-Kutta scheme transformed in that way cannot be interpreted as a Runge-Kutta scheme applied to the transformed ODE (3.1). Still one could believe that a simple modification of the arguments from the B-theory would lead to a successful stability analysis of the transformed scheme; but this is also not the case. Actually, the derivation of satisfactory stability and convergence results requires a careful use of perturbation arguments; the respective technical details can be found in [7].

By the results of [7], stiff initial value problems of the form

$$y' = J(t)y + g(t,y), \tag{3.2a}$$

$$y(0) = y_0 \tag{3.2b}$$

are covered, with an arbitrary distribution of stiff eigenvalues and with a smooth, Lipschitz-continuous nonlinearity $g(t,y)$ (this excludes highly nonlinear problems). It is assumed that a (not necessarily diagonalizing) time dependent transformation $S(t)$ exists such that the right hand side of the correspondingly transformed problem has a logarithmic norm μ_2 which reflects the well-conditioned behavior of (3.2) in a realistic way. For implicit Runge-Kutta methods applied to such problems, stability and error bounds are derived in [7] which constitute an extension of the B-theory. Simply speaking, these results say that a B-stable resp. B-convergent Runge-Kutta method is also stable resp. convergent when applied to a problem (3.2).

Still, we are far from understanding all phenomena in highly nonlinear stiff situations. Problems in reaction kinetics, for instance, form a practically important class which is by no means completely understood. It is an interesting question whether further extensions of the convergence theory are possible for reasonably large classes of highly nonlinear stiff problems.

References

[1] W. Auzinger, *On the error structure of the implicit Euler scheme applied to stiff systems of differential equations*, Computing 43, 115–131 (1989).

[2] W. Auzinger, *On error structures and extrapolation for stiff systems, with application in the method of lines*, Computing 44, 331–356 (1990).

[3] W. Auzinger, R. Frank, F. Macsek, *Asymptotic error expansions for stiff equations: the implicit Euler scheme*, SIAM J. Numer. Anal. 27, 67–104 (1990).

[4] W. Auzinger, R. Frank, *Asymptotic expansions of the global discretization error for stiff problems*, SIAM J. Sci. Stat. Comput. 10, 950–963 (1989).

[5] W. Auzinger, R. Frank, G. Kirlinger, *Asymptotic error expansions for stiff equations: applications*, Computing 43, 223–253 (1989).

[6] W. Auzinger, R. Frank, G. Kirlinger, *A note on convergence concepts for stiff problems*, Computing 44, 197–208 (1990).

[7] W. Auzinger, R. Frank, G. Kirlinger, *An extension of B-convergence for Runge-Kutta methods*, Report Nr. 83/90, Institut für Angewandte und Numerische Mathematik, Technische Universität Wien, 1990. (Submitted)

[8] R. E. O'Malley, Jr., *Introduction to singular perturbations*, Academic Press, New York, London, 1974.

[9] K. Burrage, J. C. Butcher, *Stability criteria for implicit Runge-Kutta methods*, SIAM J. Numer. Anal. 16, 46–57 (1979).

[10] J. C. Butcher, *A stability property of implicit Runge-Kutta methods*, BIT 15, 358–361 (1975).

[11] M. Crouzeix, *Sur la B-stabilité des méthodes de Runge-Kutta*, Numer. Math. 32, 75–82 (1979).

[12] G. Dahlquist, *Stability and error bounds in the numerical integration of ordinary differential equations*, Ph. D. Thesis, Stockholm, 1958.

[13] G. Dahlquist, *A special stability problem for linear multistep methods*, BIT 3, 27–43 (1963).

[14] G. Dahlquist, *Error analysis for a class of methods for stiff nonlinear initial value problems*, in: Lecture Notes in Mathematics 506, G. A. Watson (Ed.), Springer-Verlag, Berlin, 1975.

[15] G. Dahlquist, *G-stability is equivalent to A-stability*, BIT 18, 384–401 (1978).

[16] K. Dekker, J. G. Verwer, *Stability of Runge-Kutta methods for stiff nonlinear differential equations*, North-Holland, Amsterdam, New York, Oxford, 1984.

[17] R. Frank, J. Schneid, C. W. Ueberhuber, *Einseitige Lipschitzbedingungen für gewöhnliche Differentialgleichungen*, Report Nr. 33/78, Institut für Numerische Mathematik, TU Wien, 1978.

[18] R. Frank, J. Schneid, C. W. Ueberhuber, *The concept of B-convergence*, SIAM J. Numer. Anal. 18, 753–780 (1981).

[19] R. Frank, J. Schneid, C. W. Ueberhuber, *Stability properties of implicit Runge-Kutta methods*, SIAM J. Numer. Anal. 22, 497–515 (1985).

[20] R. Frank, J. Schneid, C. W. Ueberhuber, *Order results for implicit Runge-Kutta methods applied to stiff systems*, SIAM J. Numer. Anal. 22, 515–534 (1985).

[21] R. Frank, J. Schneid, C. W. Ueberhuber, *B-convergence: A survey*, Appl. Numer. Math. 5, 51–61 (1989).

[22] E. Hairer, Ch. Lubich, *Extrapolation at stiff differential equations*, Numer. Math. 52, 377–400 (1988).

[23] E. Hairer, Ch. Lubich, M. Roche, *Error of Runge-Kutta methods for stiff problems studied via differential algebraic equations*, BIT 28, 678–700 (1988).

[24] J.F.B.M. Kraaijevanger, *Contractivity of Runge-Kutta methods*, Report TW-89-12, Leiden University, 1989.

[25] Ch. Lubich, *On the convergence of multistep methods for nonlinear stiff differential equations*, Report, Institut für Mathematik und Geometrie, Universität Innsbruck, 1989.

[26] M.-N. Le Roux, *Méthodes multipas pour des équations paraboliques non linéaires*, Numer. Math. 35, 143–162 (1980).

[27] A. Prothero, A. Robinson, *On the stability and accuracy of one step methods for solving stiff systems of ordinary differential equations*, Math. Comp. 28, 145–162 (1974).

[28] J. M. Sanz-Serna, J. G. Verwer, *Stability and convergence at the PDE / Stiff ODE interface*, Appl. Numer. Math. 5, 117–132 (1989).

[29] M. N. Spijker, *Contractivity in the numerical solution of initial value problems*, Numer. Math. 42, 271–290 (1983).

[30] H. J. Stetter, *Asymptotic expansions for the error of discretization algorithms for nonlinear functional equations*, Numer. Math. 7, 18–31 (1965).

[31] R. van Veldhuizen, *Asymptotic expansions of the global error for the implicit midpoint rule (stiff case)*, Computing 33, 185–192 (1984).

[32] O. B. Widlund, *A note on unconditionally stable linear multistep methods*, BIT 7, 65–70 (1967).