Applied Numerical Mathematics 9 (1992) 91–109
North-Holland

An extension of B-convergence for Runge–Kutta methods

W. Auzinger, R. Frank and G. Kirlinger

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

Abstract

Auzinger, W., R. Frank and G. Kirlinger, An extension of B-convergence for Runge–Kutta methods, Applied Numerical Mathematics 9 (1992) 91–109.

The well-known concepts of B-stability and B-convergence for the analysis of one-step methods applied to stiff initial value problems are based on the notion of one-sided Lipschitz continuity. In a recent paper (Auzinger et al. (1990)) the authors have pointed out that the one-sided Lipschitz constant m must often be expected to be very large (positive and of the order of magnitude of the stiff eigenvalues) despite a (globally) well-conditioned behavior of the underlying problem. As a consequence, the existing B-theory suffers from considerable restrictions; e.g., not even linear systems with time-dependent coefficients are satisfactorily covered. The purpose of the present paper is to fill this gap; for implicit Runge–Kutta methods we extend the B-convergence theory such as to be valid for a class of non-autonomous weakly nonlinear stiff systems; reference to the (potentially large) one-sided Lipschitz constant is avoided. Unique solvability of the system of algebraic equations is shown, and global error bounds are derived.

1. Historical remarks and motivation

In the last two decades a number of convergence results have been derived for discretizations of nonlinear stiff initial value problems. In particular, the notion of G-stability turned out to be crucial for the analysis of linear multistep methods (see Dahlquist [8]). For Runge–Kutta methods, the concept of B-stability was essential (see Butcher [6], Burrage and Butcher [5] and Crouzeix [7]). Further concepts like BS-stability and BSI-stability were introduced by Frank, Schneid and Ueberhuber (see [13]) and led to the so-called "B-convergence" theory (see [14]) which enabled the derivation of rigorous, quantitative convergence results for implicit Runge–Kutta methods. Subsequently, a large number of papers have appeared dealing with B-stability and B-convergence. Furthermore, a number of papers have appeared dealing with the question of the solvability of the Runge–Kutta equations (see e.g. [9,16]).

All these results are valid for stiff problems $y' = f(t, y)$ satisfying a one-sided Lipschitz condition

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

where $\langle \cdot, \cdot \rangle$ denotes a scalar product in $\mathbb{R}^n$ and $\| \cdot \|$ the corresponding norm. In particular, we shall use the Euclidean norm and denote it by $\| \cdot \|_2$. The parameter m in (1.1) is called a "one-sided Lipschitz constant" for f. The B-convergence theory leads to satisfactory error bounds for problems admitting a one-sided Lipschitz constant which is not strongly positive. Whether this requirement is typically satisfied or not has hardly been examined in the respective papers. Rather, (1.1) with m not strongly positive is used as a kind of natural "axiom" without further discussion. This point of view has been motivated by the fact that, locally, the optimal (smallest possible) one-sided Lipschitz constant characterizes the condition of the initial value problem w.r.t. the corresponding norm (see e.g. [11]).

On the other hand, examples have been well known for a long time for which $m \gg 0$ (such that, locally, the condition is very bad) but which are well conditioned in a *global* sense. Consider for instance the simple linear system $y' = Jy$ with

$$J = \begin{pmatrix} -1 & 0 \\ \frac{1}{\varepsilon} - 1 & -\frac{1}{\varepsilon} \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 \\ 0 & -\frac{1}{\varepsilon} \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix} \tag{1.2}$$

with $0 < \varepsilon \ll 1$. Here the best possible one-sided Lipschitz constant w.r.t. the Euclidean norm, which can be expressed as the logarithmic matrix norm $\mu_2(J) = \lambda_{\max}((J + J^{\mathrm{T}})/2)$, is given by

$$\mu_2(J) = \frac{\sqrt{2} - 1}{2} \cdot \frac{1}{\varepsilon} - \frac{\sqrt{2} + 1}{2} = \mathrm{O}\left(+\frac{1}{\varepsilon}\right) \gg 0 \tag{1.3}$$

despite the fact that the eigenvalues of J are negative and its eigensystem matrix is very well conditioned. Thus, the parameter $\mu_2(J)$ does not reflect the good global condition of the initial value problem.

The question arises whether such a strong discrepancy between local and global condition occurs frequently or is rather the exception. This question was recently studied in [4], and it turned out that for linear two-dimensional problems $y' = Jy$ with one stiff and one nonstiff eigenvalue the logarithmic norm $\mu_2(J)$ is of moderate size *if and only if* J is nearly symmetric, which means that the angle between the eigenvectors of J must be $\pi/2 + \mathrm{O}(\varepsilon^{1/2})$ where $-1/\varepsilon$ is the order of magnitude of the stiff eigenvalue. Also for higher-dimensional systems, numerical experience indicates that nonnormality is usually incompatible with the existence of a moderate-sized one-sided Lipschitz constant.

In all these cases where $m \gg 0$, the B-theory is not successfully applicable.[1] In view of these observations, the axiomatic assumption "m not strongly positive" appears much less natural and the B-theory appears more restrictive than usually believed.

[1] For the special case of a constant coefficient problem $y' = Jy$ there is a well-known remedy: An appropriate change of norm ("Euclidean norm" → "elliptic norm") usually will lead to a moderate one-sided Lipschitz constant (not significantly larger than the spectral abscissa of J), entailing satisfactory B-convergence bounds w.r.t. this new norm. However, there is no such simple remedy for more difficult nonlinear and/or time-dependent stiff problems: Already for problems of the type $y' = J(t)y$ with an eigensystem varying smoothly in time there is *no fixed* elliptic norm with the property that the corresponding one-sided Lipschitz constant is moderate-sized for all t in the integration interval (cf. the discussion in [4]).

There also exist alternative approaches towards a convergence analysis for stiff problems. *Singular perturbation theory* is of major importance here (cf. for instance Veldhuizen [18], Auzinger, Frank and Macsek [1], Auzinger and Frank [2,3], Hairer, Lubich and Roche [15] and Lubich [17]). Since the concept of one-sided Lipschitz continuity is not relevant here, this approach does not suffer from the requirement "m not strongly positive". However, also singular perturbation theory has its drawbacks and does not cover general stiff situations: The main restriction of this approach is that special assumptions about the nature of the stiff spectrum are required (there is only one "cluster" of stiff eigenvalues of magnitude $-1/\varepsilon$, ε a small parameter). Furthermore, the large stiffness parameter $1/\varepsilon$ is not allowed to appear in a nonlinear way but only as a factor on the right-hand side.

The above discussion shows that only certain classes of stiff problems are satisfactorily covered by the existing theoretical approaches: Remarkably enough, not even simple linear time-dependent problems of the form $y' = J(t)y$ with a smoothly varying eigensystem have been covered so far (for nonsymmetric $J(t)$, the B-theory is usually not applicable, cf. [4]; singular perturbation theory is restricted to special spectral structures). Our present aim is to close this gap; we consider weakly nonlinear stiff problems of the form $y' = J(t)y + \varphi(t, y)$, with an "arbitrary" stiff spectrum and with a smooth, Lipschitz continuous function [2] $\varphi(t, y)$. The main objective of the present paper is to extend the B-convergence theory to this problem class; in particular, we study the convergence properties of *implicit Runge–Kutta methods.*

Highly nonlinear stiff problems, i.e., problems where also the nonlinear terms on the right-hand side are affected by large parameters, frequently arise in applications, e.g., in the modelling of phenomena from reaction kinetics. In their general form, such problems are not satisfactorily covered so far, neither by existing convergence theories nor by the present paper. A more powerful theoretical concept would be desirable in order to cover highly nonlinear stiff problems in a sufficiently comprehensive way.

2. Problem class and Runge–Kutta discretization

2.1. The problem class

We consider stiff initial value problems of the form

$$y' = J(t)y + \varphi(t, y), \quad t \in [0, t_{\mathrm{end}}], \tag{2.1a}$$

$$y(0) = y_0, \tag{2.1b}$$

where $J(t) \in \mathbb{R}^{n \times n}$ and $\varphi : [0, t_{\mathrm{end}}] \times \mathbb{R}^n \to \mathbb{R}^n$. The exact solution of (2.1) will be denoted by $y(t)$. As mentioned in Section 1, the logarithmic norm $\mu_2(J(t))$ must be expected to be strongly positive and much larger than the spectral abscissa of $J(t)$ in the overwhelming majority of cases, and so the logarithmic norm is usually unrelated to the (global) condition of the

[2] The assumption that $\varphi(t, y)$ be Lipschitz continuous could easily be weakened by requiring only one-sided Lipschitz continuity with a moderate-sized one-sided Lipschitz constant m_φ. However, such an extension seems hardly to be of practical relevance.

underlying problem. In all these cases, B-convergence bounds based on $\mu_2(J(t))$ are of no use; but for globally well-conditioned stiff problems with not too strong nonlinearities there usually exist suitable—generally time-dependent—coordinate transformations such that the resulting transformed initial value problem is characterized by parameters reflecting the good condition. The basic idea in our convergence theory is to make use of such a transformation in order to derive satisfactory quantitative global error bounds.

Thus we transform the given problem (2.1) according to

$$y = S(t)z \tag{2.2}$$

with an appropriately chosen smooth linear transformation $S(t)$ (cf. the discussion below), leading to

$$z' = \Gamma(t)z + D(t)z + \psi(t, z), \quad t \in [0, t_{\text{end}}], \tag{2.3a}$$

$$z(0) = z_0 := S^{-1}(0)y_0 \tag{2.3b}$$

with the exact solution $z(t) = S^{-1}(t)y(t)$. Here we have introduced

$$\begin{aligned} &\Gamma(t) := S^{-1}(t)J(t)S(t), \qquad D(t) := -S^{-1}(t)S'(t), \quad \text{and} \\ &\psi(t, z) := S^{-1}(t)\varphi(t, S(t)z). \end{aligned} \tag{2.4}$$

We make the following assumptions about the problem data:

(i) Concerning $J(t) = S(t)\Gamma(t)S^{-1}(t)$ we assume that the logarithmic norm $\mu_2(\Gamma(t))$ is uniformly bounded on $[0, t_{\text{end}}]$:

$$\mu_2(\Gamma(t)) \leqslant \alpha \quad \text{for } t \in [0, t_{\text{end}}]. \tag{2.5}$$

Furthermore, we assume that $S(t)$ is continuously differentiable on $[0, t_{\text{end}}]$, and denote by σ, $\hat{\sigma}$, δ the following bounds:

$$\|S(t)\|_2 \leqslant \sigma, \qquad \|S^{-1}(t)\|_2 \leqslant \hat{\sigma}, \qquad \|S'(t)\|_2 \leqslant \delta \tag{2.6}$$

for all $t \in [0, t_{\text{end}}]$. Note that from (2.6),

$$\|S^{-1}(t_1)S(t_2) - I\|_2 \leqslant |t_1 - t_2|\hat{\sigma}\delta, \quad t_1, t_2 \in [0, t_{\text{end}}]. \tag{2.7}$$

(ii) Concerning the nonlinearity we assume that $\psi(t, z) = S^{-1}(t)\varphi(t, S(t)z)$ is Lipschitz continuous w.r.t. z, with a Lipschitz constant L_ψ, in an appropriate "tube" $\mathcal{S}$ of width ρ about the transformed solution, i.e., $\mathcal{S}$ is a subset of $[0, t_{\text{end}}] \times \mathbb{R}^n$ where $(t, z) \in \mathcal{S}$ if the corresponding component $z \in \mathbb{R}^n$ is contained in the ball

$$\|z - z(t)\|_2 \leqslant \rho. \tag{2.8}$$

The convergence bounds in our analysis will turn out to be expressions in the parameters α, σ, $\hat{\sigma}$, δ and L_ψ (they will also depend on the actual stepsize h and on bounds for certain derivatives of the exact solution $y(t)$). These bounds become large only if at least one of these parameters becomes strongly positive. A problem (2.1) is therefore satisfactorily covered by our theory if there exists a smooth transformation $S(t)$ such that all these parameters are moderate-sized.

By $\mathcal{F}_{\alpha,\sigma,\hat{\sigma},\delta,L_\psi}$ we denote the class of all problems (2.1) for which there exists a transformation $S(t)$ such that the above assumptions are satisfied. In many cases the transformation to real canonical form [3] of $J(t)$ is the appropriate choice; then,

$$\alpha = \sup_{t \in [0, t_{\text{end}}]} \text{spectral abscissa}(J(t)) \tag{2.9}$$

(or slightly larger in the case of nonlinear elementary divisors), and so this transformation is natural and optimal w.r.t. the parameter α. Nevertheless, there exist examples where such a transformation is not satisfactory w.r.t. the other parameters σ, $\hat{\sigma}$, ... Cf. for instance [10, Example 1.4.13], where for the transformation to real diagonal form the product $\sigma\hat{\sigma}$ becomes inevitably large, whereas there exists a nondiagonalizing transformation leading to a set of moderate parameters α, σ, $\hat{\sigma}$.

2.2. The Runge–Kutta discretization

We consider implicit s-stage Runge–Kutta methods characterized by the coefficient scheme

$$\begin{array}{c|ccc} c_1 & a_{11} & \dots & a_{1s} \\ \vdots & \vdots & & \vdots \\ c_s & a_{s1} & \dots & a_{ss} \\ \hline & b_1 & \dots & b_s \end{array} = \begin{array}{c|c} c & A \\ \hline & b^{\mathrm{T}} \end{array}, \tag{2.10}$$

i.e., for (2.1) one step $(t_{\nu-1}, \eta_{\nu-1}) \to (t_\nu, \eta_\nu)$ with stepsize $h = t_\nu - t_{\nu-1}$ reads

$$Y_i = \eta_{\nu-1} + h\sum_{j=1}^{s} a_{ij}\big[J(\tau_j)Y_j + \varphi(\tau_j, Y_j)\big], \quad i = 1, \dots, s, \tag{2.11a}$$

$$\eta_\nu = \eta_{\nu-1} + h\sum_{j=1}^{s} b_j\big[J(\tau_j)Y_j + \varphi(\tau_j, Y_j)\big], \tag{2.11b}$$

where $\tau_j := t_{\nu-1} + c_j h$ $(j = 1, \dots, s)$.

By ℓ_i we denote the local truncation errors w.r.t. (2.11), i.e.,

$$\ell_i := y(\tau_i) - y(t_{\nu-1}) - h\sum_{j=1}^{s} a_{ij}\underbrace{\big[J(\tau_j)y(\tau_j) + \varphi(\tau_j, y(\tau_j))\big]}_{y'(\tau_j)}, \quad i = 1, \dots, s, \tag{2.12a}$$

$$\ell_{s+1} := y(t_\nu) - y(t_{\nu-1}) - h\sum_{j=1}^{s} b_j\underbrace{\big[J(\tau_j)y(\tau_j) + \varphi(\tau_j, y(\tau_j))\big]}_{y'(\tau_j)}. \tag{2.12b}$$

[3] Since we have assumed the problem data to be real, it is convenient to avoid complex denotation. "Real canonical form" means that complex conjugate eigenvalues are represented by real 2×2-blocks, and that a real eigensystem is chosen.

Taylor expansion [4] of $y(t)$ and use of the so-called *simplifying conditions* (cf. e.g. [10]) lead to estimates

$$\| \ell_i \|_2 \leqslant C(M_l) h^{p_i+1}, \quad i = 1, \ldots, s+1, \tag{2.13}$$

where the $C(M_l)$ denote expressions depending on bounds M_l for certain derivatives of y, i.e.,

$$\left\| \frac{\mathrm{d}^l y(t)}{\mathrm{d}t^l} \right\|_2 \leqslant M_l, \quad t \in [0, t_{\mathrm{end}}]. \text{ [5]}$$

The *stage order* p of the Runge–Kutta method is defined as

$$p := \min_{1 \leqslant i \leqslant s+1} p_i. \tag{2.14}$$

Now we transform the Runge–Kutta equations (2.11) using the same linear transformation $S(t)$ as above. We multiply the ith stage equation in (2.11a) by $S^{-1}(\tau_i)$, analogously for (2.11b) with $S^{-1}(t_\nu)$. This leads to the transformed system

$$Z_i = S_i^{-1} S_0 \zeta_{\nu-1} + h \sum_{j=1}^{s} a_{ij} S_i^{-1} S_j \left[\Gamma_j Z_j + \psi_j(Z_j) \right], \quad i = 1, \ldots, s, \tag{2.15a}$$

$$\zeta_\nu = S_{s+1}^{-1} S_0 \zeta_{\nu-1} + h \sum_{j=1}^{s} b_j S_{s+1}^{-1} S_j \left[\Gamma_j Z_j + \psi_j(Z_j) \right], \tag{2.15b}$$

where

$$S_0 := S(t_{\nu-1}), \qquad S_i := S(\tau_i), \quad i = 1, \ldots, s, \qquad S_{s+1} := S(t_\nu), \tag{2.16a}$$

$$\Gamma_i := \Gamma(\tau_i), \qquad \psi_i(Z_i) := \psi(\tau_i, Z_i), \qquad i = 1, \ldots, s, \tag{2.16b}$$

$$\zeta_{\nu-1} := S_0^{-1} \eta_{\nu-1}, \qquad Z_i := S_i^{-1} Y_i, \quad i = 1, \ldots, s, \qquad \zeta_\nu := S_{s+1}^{-1} \eta_\nu. \tag{2.16c}$$

The local truncation errors (2.12) transform into

$$\tilde{\ell}_i := S_i^{-1} \ell_i = z(\tau_i) - S_i^{-1} S_0 z(t_{\nu-1}) - h \sum_{j=1}^{s} a_{ij} S_i^{-1} S_j \left[\Gamma_j z(\tau_j) + \psi_j(z(\tau_j)) \right], \quad i = 1, \ldots, s, \tag{2.17a}$$

$$\tilde{\ell}_{s+1} := S_{s+1}^{-1} \ell_{s+1} = z(t_\nu) - S_{s+1}^{-1} S_0 z(t_{\nu-1}) - h \sum_{j=1}^{s} b_j S_{s+1}^{-1} S_j \left[\Gamma_j z(\tau_j) + \psi_j(z(\tau_j)) \right]. \tag{2.17b}$$

[4] Note that, due to their definition, the order of magnitude of the ℓ_i depends on h and on the (local) smoothness of $y(t)$; it does not depend on $\| J(t) \|_2$ and is thus not affected by the stiff eigenvalues.

[5] Since for stiff problems the local smoothness of the solution often strongly varies (in particular in the presence of transient terms), it is useful to interpret the interval $[0, t_{\mathrm{end}}]$ not as the whole integration interval but only as one subinterval where the smoothness is not strongly varying, such that the different sets of bounds M_l reflect the local smoothness in each such a subinterval. Our global error bounds will be formulated in a way enabling their inductive application over a sequence of such subintervals. Hence the general case of non-equidistant grids is covered by our analysis.

Note that the $\tilde{\ell}_i$ exactly coincide with the local truncation errors w.r.t. the transformed scheme (2.15) (i.e., the residual after inserting $z(t_{\nu-1})$, $z(t_\nu)$ and the $z(\tau_j)$ into (2.15)). Obviously,

$$\| \tilde{\ell}_i \|_2 \leqslant \hat{\sigma} C(M_l) h^{p+1}, \quad i = 1, \ldots, s+1, \tag{2.18}$$

where p is the stage order (cf. (2.14)).

Remark. The transformed differential equation (2.3a) admits a one-sided Lipschitz constant $\overline{m} = \alpha + \hat{\sigma}\delta + L_\psi$ which will often be moderate-sized for appropriately chosen $S(t)$. One could therefore think of applying the well-known B-convergence theory to (2.3) in order to derive global error bounds based on the moderate parameter $\overline{m}$. However, such a procedure would *not* lead to the desired estimate for $\eta_\nu - y(t_\nu) = S(t_\nu)(\zeta_\nu - z(t_\nu))$, because (2.15) does *not* coincide with the original Runge–Kutta scheme applied to the transformed problem (2.3).

Since (2.15) is of a similar nature as the given Runge–Kutta scheme, one might further think that the usual arguments from the B-theory can easily be modified. This is, however, not the case: If, for instance, one tries to adapt the well-known argument based on the concept of algebraic stability (cf. e.g. [5]) in order to show a stability inequality

$$\| \zeta_\nu - \bar{\zeta}_\nu \|_2 \leqslant (1 + Ch) \| \zeta_{\nu-1} - \bar{\zeta}_{\nu-1} \|_2,$$

one would have to estimate scalar products like $\langle S_i^{-1} S_j \Gamma_j (Z_j - \bar{Z}_j), S_i^{-1} S_j (Z_j - \bar{Z}_j) \rangle$ $(i \neq j)$. But even if Γ is diagonal and despite the fact that $S_i^{-1} S_j = I + O(h)$ is only slightly nonsymmetric in this case, these scalar products can become very large compared to $\| Z_j - \bar{Z}_j \|_2^2$ for certain directions $Z_j - \bar{Z}_j$ (cf. the discussion in [4]).

3. Stability concepts and results

In the following sections we shall derive global error bounds which depend on the stepsize h, on bounds M_l for certain derivatives of the solution $y(t)$ and on the problem-characterizing parameters α, σ, $\hat{\sigma}$, δ, L_ψ, and which uniformly hold for all problems from the class $\mathcal{F}_{\alpha,\sigma,\hat{\sigma},\delta,L_\psi}$ introduced in Section 2. For low-stage schemes it is simple to derive these bounds as explicit expressions in the parameters. For higher stage schemes the derivation of such explicit expressions is cumbersome and not necessary from a practical point of view: Usually it is sufficient to verify that the respective bounds are moderate-sized provided the underlying problem-characterizing parameters are; for this purpose it suffices to understand the quantitative behavior of these bounds. In this spirit, we shall use the following denotation:

Denotation. Let $p_1, p_2, \ldots$ represent certain problem-characterizing parameters (for example $\alpha, \sigma, \ldots$). Whenever an expression denoted by

$$\mathcal{B}(h; p_1, p_2, \ldots) \tag{3.1}$$

appears on the right-hand side of an estimate, this is meant to say that the estimated quantity depends on h and on the parameters $p_1, p_2, \ldots$, that it is well-defined for $h \leqslant h_0$, and that the estimate is valid for $h \leqslant h_0$, in a way such that

(i) the largest admissible stepsize $h_0 = h_0(p_1, p_2, \ldots)$ $(0 < h_0 \leqslant \infty)$ is not restrictively small ("mild stepsize restriction"),
(ii) $\mathscr{B}(h; p_1, p_2, \ldots) \leqslant \mathscr{B}(h_0; p_1, p_2, \ldots)$ for $h \leqslant h_0$, where $\mathscr{B}(h_0; p_1, p_2, \ldots)$ is moderate-sized,

provided none of the parameters p_l is strongly positive.[6] Analogously we shall use the symbol $\mathscr{B}(p_1, p_2, \ldots)$ for h-independent quantities that remain moderate provided none of the parameters p_l is strongly positive. Furthermore, the notation $\mathscr{B}(h, t_\nu; p_1, p_2, \ldots)$ will be used for estimates corresponding to fixed grid points $t \equiv t_\nu \equiv \nu h$.

We also introduce the following terminology: Each vector written in boldface font denotes a "supervector" in $\mathbb{R}^{sn}$, e.g.,

$$\boldsymbol{U} = \begin{pmatrix} U_1 \\ \vdots \\ U_s \end{pmatrix}, \quad U_i \in \mathbb{R}^n. \tag{3.2}$$

Furthermore, the symbol $|||\cdot|||_2$ denotes the Euclidean norm in $\mathbb{R}^{sn}$ or $\mathbb{R}^{(s+1)n}$, e.g.,

$$|||\boldsymbol{U}|||_2, \qquad \left|\left|\left| \begin{pmatrix} \boldsymbol{U} \\ U_{s+1} \end{pmatrix} \right|\right|\right|_2. \tag{3.3}$$

3.1. Review of stability concepts of the B-theory

It is convenient at this point to recall the well-known stability concepts of the B-theory, and to reformulate them using the above denotation. Consider the Runge–Kutta scheme (2.10) applied to an initial value problem $y' = f(t, y)$:

$$Y_i = \eta_{\nu-1} + h \sum_{j=1}^{s} a_{ij} f(\tau_j, Y_j), \quad i = 1, \ldots, s, \tag{3.4a}$$

$$\eta_\nu = \eta_{\nu-1} + h \sum_{j=1}^{s} b_j f(\tau_j, Y_j). \tag{3.4b}$$

Let m denote a one-sided Lipschitz constant for f.

- B-stability: A Runge–Kutta method is called *B-stable* if, for two "parallel" Runge–Kutta steps $(t_{\nu-1}, \eta_{\nu-1}) \to (t_\nu, \eta_\nu)$ and $(t_{\nu-1}, \tilde{\eta}_{\nu-1}) \to (t_\nu, \tilde{\eta}_\nu)$,

$$\|\eta_\nu - \tilde{\eta}_\nu\|_2 \leqslant [1 + \mathscr{B}(h; m)h] \, \|\eta_{\nu-1} - \tilde{\eta}_{\nu-1}\|_2. \tag{3.5}$$

[6] By definition, $\mathscr{B}(h; p_1, p_2, \ldots)$ is not allowed to become large (nor is h_0 allowed to become restrictively small) if some of the parameters p_l are strongly *negative* (recall that parameters like a spectral abscissa or a logarithmic norm may also take negative values).

The notions of BS- and BSI-stability refer to a perturbed scheme

$$\tilde{Y}_i = \eta_{\nu-1} + h \sum_{j=1}^{s} a_{ij} f\left(\tau_j, \tilde{Y}_j\right) + \Delta_i, \quad i = 1, \ldots, s, \tag{3.6a}$$

$$\tilde{\eta}_\nu = \eta_{\nu-1} + h \sum_{j=1}^{s} b_j f\left(\tau_j, \tilde{Y}_j\right) + \Delta_{s+1}. \tag{3.6b}$$

- BS-stability: A Runge–Kutta method is called *BS-stable* if, for (3.4) and (3.6),

$$\|\eta_\nu - \tilde{\eta}_\nu\|_2 \leqslant \mathscr{B}(h; m) \left|\left|\left| \begin{pmatrix} \boldsymbol{\Delta} \\ \Delta_{s+1} \end{pmatrix} \right|\right|\right|_2. \tag{3.7}$$

- BSI-stability: A Runge–Kutta method is called *BSI-stable* if, for (3.4) and (3.6),

$$|||\boldsymbol{Y} - \tilde{\boldsymbol{Y}}|||_2 \leqslant \mathscr{B}(h; m) \, |||\boldsymbol{\Delta}|||_2. \tag{3.8}$$

In the B-theory it is shown that B- and BS-stability are sufficient for B-convergence, i.e., the global error can be estimated by

$$\|\eta_\nu - y(t_\nu)\|_2 \leqslant \mathscr{B}(h, t_\nu; m) \, \|\eta_0 - y_0\|_2 + \mathscr{B}(h, t_\nu; m, M_l) h^p, \tag{3.9}$$

where the order of B-convergence p usually equals the stage order of the method (cf. for instance [10]). BSI-stability is an essential tool for the analysis of the solvability of the algebraic equations (3.4a).

B-, BS- and BSI-stability can be concluded from certain algebraic conditions on the coefficients of the Runge–Kutta scheme, namely algebraic stability and diagonal stability:

- A Runge–Kutta scheme (2.10) is called *algebraically stable* if the matrices $B := \mathrm{Diag}(b_1, \ldots, b_s)$ and $M := BA + A^{\mathrm{T}}B - bb^{\mathrm{T}}$ are nonnegative definite.
- A Runge–Kutta scheme (2.10) is called *diagonally stable* if there exists a positive diagonal matrix D such that $DA + A^{\mathrm{T}}D$ is positive definite.

For arbitrary m the following holds (cf. for instance [10]):

- An algebraically stable and diagonally stable Runge–Kutta method is B-stable.
- A diagonally stable Runge–Kutta method is BS- and BSI-stable. We also note that a further consequence of diagonal stability is that, for (3.4) and (3.6),

$$|||hf(\boldsymbol{Y}) - hf(\tilde{\boldsymbol{Y}})|||_2 \leqslant \mathscr{B}(h; m) \, |||\boldsymbol{\Delta}|||_2, \quad f(\boldsymbol{Y}) \equiv \begin{pmatrix} f(\tau_1, Y_1) \\ \vdots \\ f(\tau_s, Y_s) \end{pmatrix} \tag{3.10}$$

holds (cf. [10, Theorem 5.3.7]).

So, algebraic stability and diagonal stability entail B-convergence.

3.2. Modified stability concepts

As explained in Section 2, the concepts of the B-theory cannot be immediately applied to problems from class $\mathscr{F}_{\alpha,\sigma,\hat{\sigma},\delta,L_\psi}$, where m is usually very large. The convergence analysis must be

carried out for the transformed scheme [7] (2.15); to this end we introduce the following modified stability notions referring to problem class $\mathscr{F}_{\alpha,\sigma,\hat{\sigma},\delta,L_\Phi}$:

- B_t-stability: Let us call a Runge–Kutta method *B_t-stable* if, for two "parallel" steps $(t_{\nu-1}, \zeta_{\nu-1}) \to (t_\nu, \zeta_\nu)$ and $(t_{\nu-1}, \tilde{\zeta}_{\nu-1}) \to (t_\nu, \tilde{\zeta}_\nu)$ of the transformed scheme (2.15), it holds that

$$\| \zeta_\nu - \tilde{\zeta}_\nu \|_2 \leqslant \left[1 + \mathscr{B}(h; \alpha, \sigma, \hat{\sigma}, \delta, L_\Psi)h\right] \| \zeta_{\nu-1} - \tilde{\zeta}_{\nu-1} \|_2. \tag{3.11}$$

  Analogously as in (3.6) above we also consider the transformed scheme (2.15) with an additional perturbation:

$$\tilde{Z}_i = S_i^{-1} S_0 \zeta_{\nu-1} + h \sum_{j=1}^{s} a_{ij} S_i^{-1} S_j \left[\Gamma_j \tilde{Z}_j + \psi_j\left(\tilde{Z}_j\right)\right] + \Delta_i, \quad i = 1, \ldots, s, \tag{3.12a}$$

$$\tilde{\zeta}_\nu = S_{s+1}^{-1} S_0 \zeta_{\nu-1} + h \sum_{j=1}^{s} b_j S_{s+1}^{-1} S_j \left[\Gamma_j \tilde{Z}_j + \psi_j\left(\tilde{Z}_j\right)\right] + \Delta_{s+1}. \tag{3.12b}$$

- BS_t-stability: We call a Runge–Kutta method *BS_t-stable* if, for (2.15) and (3.12),

$$\| \zeta_\nu - \tilde{\zeta}_\nu \|_2 \leqslant \mathscr{B}(h; \alpha, \sigma, \hat{\sigma}, \delta, L_\Psi) \left\| \begin{pmatrix} \Delta \\ \Delta_{s+1} \end{pmatrix} \right\|_2. \tag{3.13}$$

- BSI_t-stability: We call a Runge–Kutta method *BSI_t-stable* if, for (2.15) and (3.12),

$$\| \mathbf{Z} - \tilde{\mathbf{Z}} \|_2 \leqslant \mathscr{B}(h; \alpha, \sigma, \hat{\sigma}, \delta, L_\Psi) \| \Delta \|_2. \tag{3.14}$$

Similarly as in the conventional B-theory it will turn out that BS_t-stability and BSI_t-stability can be concluded from the diagonal stability of (2.10). Furthermore, it will turn out that, analogously to (3.10), the estimate

$$\| h\Gamma(\mathbf{Z} - \tilde{\mathbf{Z}}) \|_2 \leqslant \mathscr{B}(h; \alpha, \sigma, \hat{\sigma}, \delta, L_\Psi) \| \Delta \|_2 \tag{3.15}$$

[7] It can be shown by means of counterexamples that the quantity C in a relation $\| \eta_\nu - \tilde{\eta}_\nu \|_2 = (1 + Ch) \cdot \| \eta_{\nu-1} - \tilde{\eta}_{\nu-1} \|_2$ may become strongly positive even if all parameters $\alpha, \sigma, \ldots$ are moderate-sized. These examples show that a "modified B-stability inequality" $\| \eta_\nu - \tilde{\eta}_\nu \|_2 \leqslant [1 + \mathscr{B}(h; \alpha, \sigma, \ldots)h] \| \eta_{\nu-1} - \tilde{\eta}_{\nu-1} \|_2$ cannot hold. It is therefore unavoidable to study stability in the transformed variables.

To illustrate this point, let us consider the implicit midpoint rule applied to $y' = Jy$ with J from (1.2). Here, $\alpha = -1$, $\sigma = \hat{\sigma} \approx 1.62$, $\delta = L_\Psi = 0$. For $\varepsilon = 10^{-6}$ we have $\mu_2(J) \approx 2.07 * 10^5$ (cf. (1.3)). Table 1 displays the L_2-norm of the stability matrix $(I - (h/2)J)^{-1}(I + (h/2)J)$ together with the quantity C satisfying $\| (I - (h/2)J)^{-1}(I + (h/2)J) \|_2 = 1 + Ch$ for different values of h. Hence there exists no stability estimate $\| \eta_\nu - \tilde{\eta}_\nu \|_2 \leqslant (1 + Ch) \cdot \| \eta_{\nu-1} - \tilde{\eta}_{\nu-1} \|_2$ with moderate C.

Table 1

h	$\| (I - (h/2)J)^{-1}(I + (h/2)J) \|_2$	C
10^{-1}	2.30	$1.30 * 10^1$
10^{-3}	2.41	$1.40 * 10^3$
10^{-5}	2.03	$1.03 * 10^5$
10^{-7}	1.02	$2.09 * 10^5$

(where $\Gamma := \text{Blockdiag}(\Gamma_1, \ldots, \Gamma_s) \in \mathbb{R}^{sn} \times \mathbb{R}^{sn}$) also follows from diagonal stability. (See Corollary 4.2 and Proposition 4.3.) Moreover, algebraic stability and diagonal stability imply B_t-stability (cf. Proposition 4.4). All these facts will lead us to the main results of the present paper:

Theorem 3.1. *Assume that the Runge–Kutta method* (2.10) *is diagonally stable and has stage order $p \geqslant 1$. Then, for $\eta_{\nu-1}$ in a suitable* O(1)*-neighborhood of $y(t_{\nu-1})$ and under a mild stepsize restriction, the system* (2.11a) *of algebraic equations is locally uniquely solvable, i.e., in an appropriate neighbourhood of $y(t)$ the solution exists and is unique.*

Theorem 3.2. *For an algebraically stable and diagonally stable Runge–Kutta method with stage order p, the global error can be estimated by* [8]

$$\| \eta_\nu - y(t_\nu) \|_2 \leqslant \mathscr{B}(h, t_\nu; \alpha, \sigma, \hat{\sigma}, \delta, L_\Psi) \| \eta_0 - y_0 \|_2 + \mathscr{B}(h, t_\nu; \alpha, \sigma, \hat{\sigma}, \delta, L_\Psi, M_t)h^p. \tag{3.16}$$

We say that the Runge–Kutta method is B_t-*convergent* of order p.

Proofs are given in Section 4.

Thus, algebraic stability and diagonal stability are crucial for the analysis of solvability and convergence of Runge–Kutta methods—not only for problems with moderate m (conventional B-theory) but also for problem class $\mathscr{F}_{\alpha,\sigma,\hat{\sigma},\delta,L_\Phi}$. In this sense our results constitute an extension of the B-theory.

4. Proofs

4.1. BSI_t-stability. Proof of Theorem 3.1 (unique solvability)

In this subsection we show that the algebraic equations (2.11a) admit a unique solution $Y_1, \ldots, Y_s$ in a suitable neighborhood of $y(t)$, provided $\eta_{\nu-1}$ is contained in a sufficiently small O(1)-neighborhood of $y(t_{\nu-1})$. To this end we study the system (2.15a) arising after the regular transformation $Z_i = S_i^{-1} Y_i$. We rewrite (2.15a) in the form

$$\boldsymbol{M}\mathbf{Z} - h\hat{\boldsymbol{\Psi}}(\mathbf{Z}) = \boldsymbol{\Xi}(\zeta_{\nu-1}), \tag{4.1}$$

where $\boldsymbol{M}$ denotes the $sn \times sn$-matrix

$$\boldsymbol{M} := \begin{pmatrix} I - a_{11}h\Gamma_1 & a_{12}S_1^{-1}S_2 h\Gamma_2 & \ldots & a_{1s}S_1^{-1}S_s h\Gamma_s \\ a_{21}S_2^{-1}S_1 h\Gamma_1 & I - a_{22}h\Gamma_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{s1}S_s^{-1}S_1 h\Gamma_1 & \ldots & \ldots & I - a_{ss}h\Gamma_s \end{pmatrix} \tag{4.2}$$

[8] In the case of a non-equidistant grid, $[0, t_{\text{end}}]$ can be interpreted as one subinterval with constant stepsize h, and $\eta_0 - y_0$ is the accumulated global error from the preceding subintervals.

and

$$Z = \begin{pmatrix} Z_1 \\ \vdots \\ Z_s \end{pmatrix}, \quad \hat{\Psi}(Z) := \begin{pmatrix} \sum_{j=1}^{s} a_{1j} S_1^{-1} S_j \psi_j(Z_j) \\ \vdots \\ \sum_{j=1}^{s} a_{sj} S_s^{-1} S_j \psi_j(Z_j) \end{pmatrix}, \quad \Xi(\zeta) := \begin{pmatrix} S_1^{-1} S_0 \zeta \\ \vdots \\ S_s^{-1} S_0 \zeta \end{pmatrix}. \tag{4.3}$$

As an essential prerequisite we prove (3.14) and (3.15) for the linear case

$$\psi(t, z) \equiv S^{-1}(t) \varphi(t, S(t) z) \equiv 0:$$

Lemma 4.1. *A diagonally stable Runge–Kutta method is BSI$_t$-stable in the linear case, i.e.,*

$$\| \! | M^{-1} | \! \|_2 \leqslant \mathscr{B}(h; \alpha, \hat{\sigma}, \delta). \tag{4.4}$$

Furthermore,

$$\| \! | h\Gamma M^{-1} | \! \|_2 \leqslant \mathscr{B}(h; \alpha, \hat{\sigma}, \delta). \tag{4.5}$$

Proof. The matrix M can be written in the form

$$M = M_0 + h\, \partial M h \Gamma, \tag{4.6}$$

where

$$M_0 := \begin{pmatrix} I - a_{11} h\Gamma_1 & a_{12} h\Gamma_2 & \dots & a_{1s} h\Gamma_s \\ a_{21} h\Gamma_1 & I - a_{22} h\Gamma_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{s1} h\Gamma_1 & \dots & \dots & I - a_{ss} h\Gamma_s \end{pmatrix}, \tag{4.7}$$

$$\partial M := \begin{pmatrix} 0 & a_{12}(1/h)\big(S_1^{-1} S_2 - I\big) & \dots & a_{1s}(1/h)\big(S_1^{-1} S_s - I\big) \\ a_{21}(1/h)\big(S_2^{-1} S_1 - I\big) & 0 & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{s1}(1/h)\big(S_s^{-1} S_1 - I\big) & \dots & \dots & 0 \end{pmatrix} \tag{4.8}$$

and $\Gamma = \text{Blockdiag}(\Gamma_1, \dots, \Gamma_s)$.

Estimation of $\| \! | M_0^{-1} | \! \|_2$ and $\| \! | h\Gamma M_0^{-1} | \! \|_2$: Obviously, M_0 is the coefficient matrix of the system of linear equations arising from application of the original Runge–Kutta scheme (2.10) to $z' = \Gamma(t) z$, where the logarithmic norm of $\Gamma(t)$ is bounded by α (cf. (2.5)). By assumption of diagonal stability, the Runge–Kutta scheme is BSI-stable (cf. (3.8)) and satisfies (3.10)). Thus,

$$\| \! | M_0^{-1} | \! \|_2 \leqslant \mathscr{B}(h; \alpha) \quad \text{and} \quad \| \! | h\Gamma M_0^{-1} | \! \|_2 \leqslant \mathscr{B}(h; \alpha). \tag{4.9}$$

Estimation of $\| \! | \partial M | \! \|_2$: Due to our smoothness assumptions w.r.t. $S(t)$ we have (cf. (2.7)):

$$\| (1/h)\big(S_i^{-1} S_j - I\big) \|_2 \leqslant \mathscr{B}(\hat{\sigma}, \delta), \tag{4.10}$$

hence

$$\| \! | \partial M | \! \|_2 \leqslant \mathscr{B}(\hat{\sigma}, \delta). \tag{4.11}$$

Now, (4.6) yields

$$M^{-1} = M_0^{-1} \big(I + h\, \partial M h \Gamma M_0^{-1}\big)^{-1}, \tag{4.12}$$

and the desired estimates (4.4) and (4.5) follow easily from (4.9) and (4.11). □

Proof of Theorem 3.1 (unique solvability, nonlinear case). With the denotation

$$z := \begin{pmatrix} z(\tau_1) \\ \vdots \\ z(\tau_s) \end{pmatrix} \tag{4.13}$$

we have

$$Mz = h\hat{\Psi}(z) - \Xi\big(z(t_{\nu-1})\big) = \tilde{\ell}, \tag{4.14}$$

where $\tilde{\ell}$ is the transformed local truncation error (cf. (2.17a)):

$$\tilde{\ell} = \begin{pmatrix} \tilde{\ell}_1 \\ \vdots \\ \tilde{\ell}_s \end{pmatrix} \quad \text{satisfying} \quad \| \! | \tilde{\ell} | \! \|_2 \leqslant \mathscr{B}(\hat{\sigma}, M_l) h^{p+1} \tag{4.15}$$

(p is the stage order, cf. (2.18)).

Now we rewrite the nonlinear system (4.1) in fixed point form:

$$Z = M^{-1} h\hat{\Psi}(Z) + M^{-1} \Xi(\zeta_{\nu-1}) =: F(Z) \tag{4.16}$$

and verify the assumptions of the contraction theorem. Consider the ball

$$K_\rho := \{Z \in \mathbb{R}^{sn} : \| \! | Z - z | \! \|_2 \leqslant \rho\}$$

with ρ from (2.8). Obviously, $\hat{\Psi}(Z)$ satisfies

$$\| \! | \hat{\Psi}(Z) - \hat{\Psi}(\tilde{Z}) | \! \|_2 \leqslant L_{\hat{\Psi}} \| \! | Z - \tilde{Z} | \! \|_2 \quad \text{for } Z, \tilde{Z} \in K_\rho, \tag{4.17}$$

where

$$L_{\hat{\Psi}} \leqslant \mathscr{B}\big(\sigma, \hat{\sigma}, L_\psi\big). \tag{4.18}$$

First we show that $Z \in K_\rho$ implies $F(Z) \in K_\rho$:
From (4.14) and (4.16),

$$\begin{aligned} F(Z) - z &= M^{-1}\big[h\hat{\Psi}(Z) + \Xi(\zeta_{\nu-1}) - Mz\big] \\ &= M^{-1}\big[h\hat{\Psi}(Z) - h\hat{\Psi}(z) - \tilde{\ell} + \Xi(\zeta_{\nu-1}) - \Xi(z(t_{\nu-1}))\big]. \end{aligned} \tag{4.19}$$

Using (4.4), (4.15), (4.17) and (4.18) we obtain

$$|||\, F(Z)-z\,|||_2 \leqslant \mathscr{B}(h;\alpha,\hat{\sigma},\delta)\big[h\mathscr{B}(\sigma,\hat{\sigma},L_\psi)\,|||\, Z-z\,|||_2 + \mathscr{B}(h;\hat{\sigma},M_l)h^{p+1} + |||\Xi(\zeta_{\nu-1})-\Xi(z(t_{\nu-1}))\,|||_2\big]. \tag{4.20}$$

By definition of $\Xi(\zeta)$ (cf. (4.3)) it obviously follows that, for $Z\in K_\rho$,

$$|||\, F(Z)-z\,|||_2 \leqslant \mathscr{B}(h;\alpha,\hat{\sigma},\delta)\big[h\mathscr{B}(\sigma,\hat{\sigma},L_\psi)\,|||\, Z-z\,|||_2 + \mathscr{B}(h;\hat{\sigma},M_l)h^{p+1} + \mathscr{B}(\sigma,\hat{\sigma})\,\|\zeta_{\nu-1}-z(t_{\nu-1})\,\|_2\big] \tag{4.21}$$

hence, for sufficiently small $\zeta_{\nu-1}-z(t_{\nu-1})$ (at O(1)-level), $F(Z)\in K_\rho$ holds under a mild stepsize restriction.

Norm contractivity of F in K_ρ:

By definition of F and due to (4.4) we obtain for $Z,\tilde{Z}\in K_\rho$,

$$\begin{aligned}|||\, F(Z)-F(\tilde{Z})\,|||_2 &\leqslant h\,|||\, M^{-1}\,|||_2\,|||\,\hat{\Psi}(Z)-\hat{\Psi}(\tilde{Z})\,|||_2\\ &\leqslant h\mathscr{B}(h;\alpha,\hat{\sigma},\delta)L_{\hat{\Psi}}\,|||\, Z-\tilde{Z}\,|||_2\\ &\leqslant h\mathscr{B}(h;\alpha,\sigma,\hat{\sigma},\delta,L_\psi)\,|||\, Z-\tilde{Z}\,|||_2,\end{aligned} \tag{4.22}$$

and thus, F is a contractive operator under the mild stepsize restriction $h\mathscr{B}(h;\alpha,\sigma,\hat{\sigma},\delta,L_\psi)\leqslant k<1$.

Thus the assumptions of the contraction theorem are satisfied, and the unique existence of the Z_i and consequently of the original $Y_i=S_iZ_i$ has been established. □

Corollary 4.2. *A diagonally stable Runge–Kutta method is BSI$_t$-stable (i.e., (3.14) holds) and satisfies (3.15).*

Proof. Consider the system of equations (4.1) and a perturbed system

$$M\tilde{Z}-h\hat{\Psi}(\tilde{Z})=\Xi(\zeta_{\nu-1})+\Delta. \tag{4.23}$$

For the difference $Z-\tilde{Z}$ we easily obtain

$$|||\, Z-\tilde{Z}\,|||_2 \leqslant h\,|||\, M^{-1}\,|||_2 L_{\hat{\Psi}}\,|||\, Z-\tilde{Z}\,|||_2 + |||\, M^{-1}\,|||_2\,|||\,\Delta\,|||_2, \tag{4.24}$$

hence (3.14) holds due to (4.4) and (4.18), i.e., the Runge–Kutta method is BSI$_t$-stable.

Multiplication of $Z-\tilde{Z}$ by $h\Gamma$ leads to

$$|||\, h\Gamma(Z-\tilde{Z})\,|||_2 \leqslant |||\, h\Gamma M^{-1}\,|||_2\big[hL_{\hat{\Psi}}\,|||\, Z-\tilde{Z}\,|||_2 + |||\,\Delta\,|||_2\big]. \tag{4.25}$$

Now, since (3.14) has already been verified, the desired estimate (3.15) follows easily with the help of (4.5). □

4.2. BS$_t$- and B$_t$-Stability. Proof of Theorem 3.2 (convergence)

In the following we write the transformed Runge–Kutta scheme (2.15) in the form (cf. (4.1))

$$MZ-h\hat{\Psi}(Z)=\Xi(\zeta_{\nu-1}), \tag{4.26a}$$

$$\zeta_\nu=S_{s+1}^{-1}S_0\zeta_{\nu-1}+\hat{B}\big[h\Gamma Z+h\Psi(Z)\big]. \tag{4.26b}$$

Here we have introduced

$$\hat{B}:=\big(b_1S_{s+1}^{-1}S_1,\ldots,b_sS_{s+1}^{-1}S_s\big),\qquad \Psi(Z):=\begin{pmatrix}\psi_1(Z_1)\\ \vdots\\ \psi_s(Z_s)\end{pmatrix}. \tag{4.27}$$

Proposition 4.3. *A diagonally stable Runge–Kutta method is BS$_t$-stable.*

Proof. Consider the transformed Runge–Kutta scheme (4.26) and a perturbed scheme

$$M\tilde{Z}-h\hat{\Psi}(\tilde{Z})=\Xi(\zeta_{\nu-1})+\Delta, \tag{4.28a}$$

$$\tilde{\zeta}_\nu=S_{s+1}^{-1}S_0\zeta_{\nu-1}+\hat{B}\big[h\Gamma\tilde{Z}+h\Psi(\tilde{Z})\big]+\Delta_{s+1}. \tag{4.28b}$$

The difference of (4.26b) and (4.28b) reads

$$\zeta_\nu-\tilde{\zeta}_\nu=\hat{B}\big[h\Gamma(Z-\tilde{Z})+h\big(\Psi(Z)-\Psi(\tilde{Z})\big)\big]-\Delta_{s+1}, \tag{4.29}$$

and the desired BS$_t$-stability estimate (3.13) easily follows with the help of (3.15) (which is proved by Corollary 4.2) and the Lipschitz continuity of $\psi(t,z)$. □

Proposition 4.4. *An algebraically stable and diagonally stable Runge–Kutta method is B$_t$-stable.*

Proof. We use the denotations (4.2), (4.3), (4.6)–(4.8), (4.27) and introduce

$$\hat{\Psi}_0(Z):=\begin{pmatrix}\sum_{j=1}^s a_{1j}\psi_j(Z_j)\\ \vdots\\ \sum_{j=1}^s a_{sj}\psi_j(Z_j)\end{pmatrix},\qquad \Xi_0(\zeta):=\begin{pmatrix}\zeta\\ \vdots\\ \zeta\end{pmatrix},\qquad \hat{B}_0:=(b_1I,\ldots,b_sI) \tag{4.30}$$

and

$$\partial\hat{\Psi}(Z):=\frac{1}{h}\big(\hat{\Psi}(Z)-\hat{\Psi}_0(Z)\big)=\begin{pmatrix}\sum_{j=1}^s a_{1j}(1/h)\big(S_1^{-1}S_j-I\big)\psi_j(Z_j)\\ \vdots\\ \sum_{j=1}^s a_{sj}(1/h)\big(S_s^{-1}S_j-I\big)\psi_j(Z_j)\end{pmatrix},$$

$$\partial\Xi(\zeta) := \frac{1}{h}(\Xi(\zeta) - \Xi_0(\zeta)) = \begin{pmatrix} (1/h)(S_1^{-1}S_0 - I)\zeta \\ \vdots \\ (1/h)(S_s^{-1}S_0 - I)\zeta \end{pmatrix}, \tag{4.31}$$

$$\partial\hat{B} := (1/h)(\hat{B} - \hat{B}_0) = (b_1(1/h)(S_{s+1}^{-1}S_1 - I), \ldots, b_s(1/h)(S_{s+1}^{-1}S_s - I)).$$

In the following we make use of the well-known stability properties (in the sense of the traditional B-theory) of the Runge–Kutta scheme (2.10). To this end we rewrite the transformed Runge–Kutta scheme (4.26) as

$$M_0 Z - h\hat{\Psi}_0(Z) = \Xi_0(\zeta_{\nu-1}) \underbrace{-h\partial M h\Gamma Z + h^2\partial\hat{\Psi}(Z) + h\,\partial\Xi(\zeta_{\nu-1})}_{:=\Delta}, \tag{4.32a}$$

$$\zeta_\nu = \zeta_{\nu-1} + \hat{B}_0[h\Gamma Z + h\Psi(Z)] \underbrace{+h(1/h)(S_{s+1}^{-1}S_0 - I)\zeta_{\nu-1} + h\,\partial\hat{B}[h\Gamma Z + h\Psi(Z)]}_{:=\Delta_{s+1}}. \tag{4.32b}$$

(4.32) can be interpreted as one step $(t_{\nu-1}, \zeta_{\nu-1}) \to (t_\nu, \zeta_\nu)$ of the original Runge–Kutta scheme (2.10), with the indicated perturbation Δ resp. Δ_{s+1}, applied to $z' = \Gamma(t)z + \psi(t, z)$. We also consider a step $(t_{\nu-1}, \tilde{\zeta}_{\nu-1}) \to (t_\nu, \tilde{\zeta}_\nu)$ of the transformed scheme (4.26) with a perturbed initial value $\tilde{\zeta}_{\nu-1}$. Again we interpret this as one step of the original scheme (2.10) applied to $z' = \Gamma(t)z + \psi(t, z)$, with an appropriately defined perturbation $\tilde{\Delta}$, $\tilde{\Delta}_{s+1}$:

$$M_0\tilde{Z} - h\hat{\Psi}_0(\tilde{Z}) = \Xi_0(\tilde{\zeta}_{\nu-1}) \underbrace{-h\,\partial M h\Gamma\tilde{Z} + h^2\partial\hat{\Psi}(\tilde{Z}) + h\,\partial\Xi(\tilde{\zeta}_{\nu-1})}_{:=\tilde{\Delta}}, \tag{4.33a}$$

$$\tilde{\zeta}_\nu = \tilde{\zeta}_{\nu-1} + \hat{B}_0[h\Gamma\tilde{Z} + h\Psi(\tilde{Z})] \underbrace{+h(1/h)(S_{s+1}^{-1}S_0 - I)\tilde{\zeta}_{\nu-1} + h\,\partial\hat{B}[h\Gamma\tilde{Z} + h\Psi(\tilde{Z})]}_{:=\tilde{\Delta}_{s+1}}. \tag{4.33b}$$

Furthermore, we consider an auxiliary step $(t_{\nu-1}, \tilde{\zeta}_{\nu-1}) \to (t_\nu, \mathring{\zeta}_\nu)$ of the form (4.32), with the same perturbation Δ, Δ_{s+1} as in (4.32) but with initial value $\tilde{\zeta}_{\nu-1}$ from (4.33) (cf. Fig. 1).

Now we make use of the following facts:

- Our assumptions imply that the original Runge–Kutta scheme (2.10) is B-stable; therefore the difference [9] $\zeta_\nu - \mathring{\zeta}_\nu$ can be estimated by a B-stability inequality (3.5) with $m = \alpha + L_\psi$ which is a one-sided Lipschitz constant for $\Gamma(t)z + \psi(t, z)$.

[9] The additional perturbation Δ, Δ_{s+1} plays no role in this argument because it is the same in the equations defining ζ_ν and $\mathring{\zeta}_\nu$.

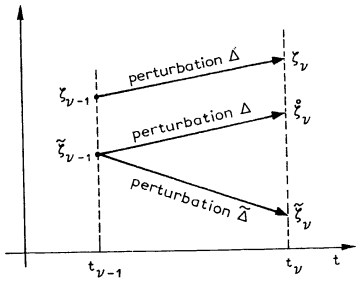

Fig. 1. Proof of B_t-stability.

- Similarly, the difference $\mathring{\zeta}_\nu - \tilde{\zeta}_\nu$ can be estimated on the basis of BS-stability (cf. (3.7)), which follows from diagonal stability.

This leads us to

$$\begin{aligned} \|\zeta_\nu - \tilde{\zeta}_\nu\|_2 &\leq \|\zeta_\nu - \mathring{\zeta}_\nu\|_2 + \|\mathring{\zeta}_\nu - \tilde{\zeta}_\nu\|_2 \\ &\leq [1 + \mathscr{B}(h; \alpha, L_\psi)h]\,\|\zeta_{\nu-1} - \tilde{\zeta}_{\nu-1}\|_2 \\ &\quad + \mathscr{B}(h; \alpha, L_\psi)\left|\!\left|\!\left| \begin{pmatrix} \Delta - \tilde{\Delta} \\ \Delta_{s+1} - \tilde{\Delta}_{s+1} \end{pmatrix} \right|\!\right|\!\right|_2 . \end{aligned} \tag{4.34}$$

Furthermore, the difference $\Delta - \tilde{\Delta}$ (cf. (4.32), (4.33)) can be estimated by

$$\begin{aligned} |||\,\Delta - \tilde{\Delta}\,|||_2 \leq h\big[&\mathscr{B}(\hat{\sigma}, \delta)\,|||\,h\Gamma(Z - \tilde{Z})\,|||_2 + h\mathscr{B}(\hat{\sigma}, \delta, L_\psi)\,|||\,Z - \tilde{Z}\,|||_2 \\ &+ \mathscr{B}(\hat{\sigma}, \delta)\,\|\zeta_{\nu-1} - \tilde{\zeta}_{\nu-1}\|_2\big] \end{aligned} \tag{4.35}$$

(cf. (4.11), (4.10)).

Up to now we have interpreted Z and $\tilde{Z}$ as solutions of perturbed original Runge–Kutta steps (2.10) applied to $z' = \Gamma(t)z + \psi(t, z)$. For the rest of the proof we consider Z and $\tilde{Z}$ as solutions of our transformed scheme (4.26a) starting from $\zeta_{\nu-1}$; the step defining Z is considered as unperturbed and that one defining $\tilde{Z}$ is affected with a perturbation $\Xi(\tilde{\zeta}_{\nu-1} - \zeta_{\nu-1})$. From

$$|||\,\Xi(\tilde{\zeta}_{\nu-1} - \zeta_{\nu-1})\,|||_2 \leq \mathscr{B}(\sigma, \hat{\sigma})\,\|\tilde{\zeta}_{\nu-1} - \zeta_{\nu-1}\|_2,$$

from (3.14) and (3.15) (cf. Corollary 4.2), and from (4.35) we conclude

$$|||\,\Delta - \tilde{\Delta}\,|||_2 \leq \mathscr{B}(h; \alpha, \sigma, \hat{\sigma}, \delta, L_\psi)h\,\|\zeta_{\nu-1} - \tilde{\zeta}_{\nu-1}\|_2. \tag{4.36}$$

In a similar way,

$$\|\Delta_{s+1} - \tilde{\Delta}_{s+1}\|_2 \leq \mathscr{B}(h; \alpha, \sigma, \hat{\sigma}, \delta, L_\psi)h\,\|\zeta_{\nu-1} - \tilde{\zeta}_{\nu-1}\|_2 \tag{4.37}$$

can be shown with the help of BS_t-stability (cf. Proposition 4.3 above).

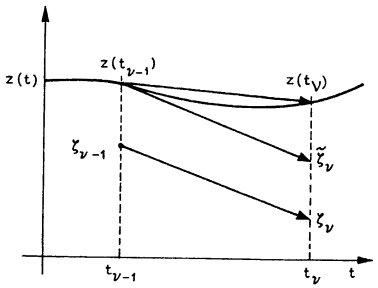

Fig. 2. Proof of B_t-convergence.

Summarizing (4.34)–(4.37) we end up with the desired B_t-stability inequality

$$\| \zeta_\nu - \tilde{\zeta}_\nu \|_2 \leqslant \left[1 + \mathscr{B}(h; \alpha, \sigma, \hat{\sigma}, \delta, L_\psi)h\right] \| \zeta_{\nu-1} - \tilde{\zeta}_{\nu-1} \|_2. \quad \square \tag{4.38}$$

Proof of Theorem 3.2 (B_t-convergence). The convergence argument is now standard: Consider a step $(t_{\nu-1}, \zeta_{\nu-1}) \to (t_\nu, \zeta_\nu)$ of the transformed scheme (4.26) and a perturbed step with the perturbation $\Delta = \bar{\ell}$, $\Delta_{s+1} = \ell_{s+1}$ (i.e., the transformed local truncation error (2.17)) starting from the exact solution value $z(t_{\nu-1})$. By definition of $\bar{\ell}$, ℓ_{s+1} this perturbed step exactly results in $z(t_\nu)$. We also consider an auxiliary step $z(t_{\nu-1}) \to \tilde{\zeta}_\nu$ of the unperturbed scheme (see Fig. 2). With the help of B_t-stability, BS_t-stability and the local error estimate (2.18) we then obtain

$$\begin{aligned} \| \zeta_\nu - z(t_\nu) \|_2 &\leqslant \| \zeta_\nu - \tilde{\zeta}_\nu \|_2 + \| \tilde{\zeta}_\nu - z(t_\nu) \|_2 \\ &\leqslant \left[1 + \mathscr{B}(h; \alpha, \sigma, \hat{\sigma}, \delta, L_\psi)h\right] \| \zeta_{\nu-1} - z(t_{\nu-1}) \|_2 \\ &\quad + \mathscr{B}(h; \alpha, \sigma, \hat{\sigma}, \delta, L_\psi, M_l)h^{p+1}, \end{aligned} \tag{4.39}$$

which enables us to estimate the transformed global error $\zeta_\nu - z(t_\nu)$ by means of the usual recursion. Finally, the desired global error estimate (3.16) immediately follows due to $\eta_\nu - y(t_\nu) = S(t_\nu)(\zeta_\nu - z(t_\nu))$. $\square$

References

[1] W. Auzinger, R. Frank and F. Macsek, Asymptotic error expansions for stiff equations: the implicit Euler scheme, *SIAM J. Numer. Anal.* 27 (1990) 67–104.
[2] W. Auzinger and R. Frank, Asymptotic error expansions for stiff equations: an analysis for the implicit midpoint and trapezoidal rules in the strongly stiff case, *Numer. Math.* 56 (1989) 469–499.
[3] W. Auzinger and R. Frank, Asymptotic expansions of the global discretization error for stiff problems, *SIAM J. Sci. Statist. Comput.* 10 (1989) 950–963.
[4] W. Auzinger, R. Frank and G. Kirlinger, A note on convergence concepts for stiff problems, *Computing* 44 (1990) 197–208.

[5] K. Burrage and J.C. Butcher, Stability criteria for implicit Runge–Kutta methods, *SIAM J. Numer. Anal.* 16 (1979) 46–57.
[6] J.C. Butcher, A stability property of implicit Runge–Kutta methods, *BIT* 15 (1975) 358–361.
[7] M. Crouzeix, Sur la B-stabilité des méthodes de Runge–Kutta, *Numer. Math.* 32 (1979) 75–82.
[8] G. Dahlquist, Error analysis for a class of methods for stiff nonlinear initial value problems, in: G.A. Watson, ed., Lecture Notes in Mathematics 506 (Springer, Berlin, 1976).
[9] K. Dekker, Error bounds for the solution to the algebraic equations in Runge–Kutta methods, *BIT* 24 (1984) 347–356.
[10] K. Dekker and J.G. Verwer, Stability of Runge–Kutta Methods for Stiff Nonlinear Differential Equations (North-Holland, Amsterdam, 1984).
[11] R. Frank, J. Schneid and C.W. Ueberhuber, Einseitige Lipschitzbedingungen für gewöhnliche Differentialgleichungen, Report Nr. 33/78, Institut für Numerische Mathematik, TU Wien (1978).
[12] R. Frank, J. Schneid and C.W. Ueberhuber, The concept of B-convergence, *SIAM J. Numer. Anal.* 18 (1981) 753–780.
[13] R. Frank, J. Schneid and C.W. Ueberhuber, Stability properties of implicit Runge–Kutta methods, *SIAM J. Numer. Anal.* 22 (1985) 497–515.
[14] 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) 515–534.
[15] E. Hairer, C. Lubich and M. Roche, Error of Runge–Kutta methods for stiff problems studied via differential algebraic equations, *BIT* 28 (1988) 678–700.
[16] J.F.B.M. Kraaijevanger and J. Schneid, On the unique solvability of the Runge–Kutta equations, *Numer. Math.* 59 (1991) 129–157.
[17] C. Lubich, On the convergence of multistep methods for nonlinear stiff differential equations, *Numer. Math.* 58 (1991) 839–853.
[18] M. van Veldhuizen, D-Stability, *SIAM J. Numer. Anal.* 18 (1981) 45–64.