BIT Numerical Mathematics **44:** 343–361, 2004.

ON THE CONVERGENCE BEHAVIOUR OF VARIABLE STEPSIZE MULTISTEP METHODS FOR SINGULARLY PERTURBED PROBLEMS ⋆

MECHTHILD THALHAMMER[1]

[1] *Institut für Technische Mathematik, Geometrie und Bauinformatik, Universität Innsbruck, Technikerstraße 13, A-6020 Innsbruck, Austria. email: Mechthild.Thalhammer@uibk.ac.at*

Abstract.

In this note, we investigate the convergence behaviour of linear multistep discretizations for singularly perturbed systems, emphasising the features of variable stepsizes. We derive a convergence result for $A(\varphi)$-stable linear multistep methods and specify a refined error estimate for backward differentiation formulas. Important ingredients in our convergence analysis are stability bounds for non-autonomous linear problems that are obtained by perturbation techniques.

AMS subject classification (2000): 65L05, 65L06, 65L20.

Key words: singular perturbation problems, linear multistep methods, backward differentiation formulas, variable stepsizes, stability, convergence.

1 Introduction.

In this paper, we analyse the convergence and stability behaviour of variable stepsize linear multistep methods applied to singularly perturbed systems. Singular perturbation problems arise in various applications such as chemical kinetics and fluid mechanics, see for example [7, 8, 5] and references therein. Another illustration modelling oscillations in electric circuits is the well-known unforced Van der Pol equation [12, 13]

$$\ddot{z}(\tau) + \mu\big(z^2(\tau) - 1\big)\dot{z}(\tau) + z(\tau) = 0, \quad \mu \gg 1.$$

By rescaling the independent variable $\tau = \mu t$ and introducing a new function y, this nonlinear differential equation takes the usual form of a first order system

$$\begin{cases} y'(t) = -z(t), \\ \varepsilon z'(t) = y(t) + z(t) - \dfrac{1}{3} z^3(t), \end{cases} \qquad \text{where } \varepsilon = \frac{1}{\mu^2} \ll 1.$$

⋆ Received July 2003. Accepted March 2004. Communicated by Timo Eirola.

In this note, more generally, we consider a singularly perturbed system of non-linear differential equations involving a small parameter $0 < \varepsilon \leq \varepsilon_0$

$$\begin{cases} y'(t) = f\big(y(t), z(t)\big), \\ \varepsilon z'(t) = g\big(y(t), z(t)\big). \end{cases}$$

A basic assumption is that the solutions y and z are bounded and have bounded derivatives with bounds independent of the parameter ε for $\varepsilon \in (0, \varepsilon_0]$. This requirement can be achieved by choosing the initial values on the existent invariant manifold, e.g. In this situation, it is shown by Lubich [6, Theorem 3] that a strongly stable linear k-step method of order p, applied with constant time step $h > 0$ sufficiently small, satisfies the following error estimate on bounded time intervals $t_n = nh \leq T$ for some $0 < \gamma < 1$ if $h \geq \varepsilon$

$$\begin{aligned} &\|y_n - y(t_n)\| + \|z_n - z(t_n)\| \\ &\quad \leq C \max_{0 \leq i \leq k-1} \|y_i - y(t_i)\| + C\big(h + \gamma^n\big) \max_{0 \leq i \leq k-1} \|z_i - z(t_i)\| + \\ &\qquad + Ch^p \int_0^{t_n} \big\|y^{(p+1)}(\tau)\big\| \, \mathrm{d}\tau + \varepsilon \, Ch^p \max_{0 \leq \tau \leq t_n} \big\|z^{(p+1)}(\tau)\big\|. \end{aligned}$$

In consideration of practical implementations, our objective is to extend this convergence estimate to variable stepsizes. In this regard, main techniques are a linearization of the right-hand side of the singularly perturbed equation along the exact solution and a fixed-point iteration based on a discrete variation-of-constants formula. Thereto, essential tools are stability bounds for non-autonomous linear problems. For proving the needed stability estimates, we employ perturbation techniques related to [9] where variable stepsize linear multistep discretizations of parabolic equations are analysed. As in [9], following an approach used by [3] and later by [10], our stability estimates involve a stability factor which depends on the stepsize sequence. As a consequence, stability is obtained under the requirement that the considered stepsize sequence varies smoothly. Moreover, an essential ingredient is a decomposition of the companion matrix of a variable stepsize linear multistep method specified in [10] and further investigated in [9].

The contents of the present paper are as follows. In Section 2, we first introduce the problem and numerical method classes and give the precise assumptions on the singularly perturbed system, the linear multistep method, and the stepsize sequence. Besides, we collect some useful relations for the solution and the right-hand side of the singular perturbation problem. Section 3 is devoted to the derivation of the necessary stability results stated in Theorem 3.3. The main idea is to relate the original equation to a less involved problem. The desired bounds then follow from a telescopic identity and a Gronwall lemma. In Section 4, we finally prove the analogue of Lubich's convergence estimate for variable stepsizes.

2 Problem and numerical discretization.

In this section, we introduce the problem and numerical method class under consideration. We specify the general scheme of a variable stepsize linear multistep method for a singular perturbation problem and further state the precise hypotheses on the problem, the method, and the stepsize sequence. For our purposes, it is useful to write the differential equation and its discretization in compact vector notation. Auxiliary results for the solution and the right-hand side of the differential equation are given in Sections 2.3 and 2.4.

2.1 Singular perturbation problem.

We consider a singularly perturbed system of ordinary differential equations involving a small parameter $0 < \varepsilon \le \varepsilon_0$

$$\begin{cases} y'(t) = f\big(y(t), z(t)\big), & y(0) \text{ given}, \\ \varepsilon z'(t) = g\big(y(t), z(t)\big), & z(0) \text{ given}, \end{cases} \tag{2.1}$$

with solution $(y(t), z(t))^{\mathrm{T}} \in \mathbb{R}^m = \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$ defined on some finite time interval $[0, T]$. For notational simplicity, the dependence of y and z on ε is omitted.

In many cases, it is convenient to employ a compact vector notation of (2.1). For that reason, we set $u = (y, z)^{\mathrm{T}}$ and denote the function defining the right-hand side of the differential equation by $F = (F_1, F_2)^{\mathrm{T}} = (f, g)^{\mathrm{T}}$. Therewith, the above initial value problem writes as

$$I_\varepsilon u'(t) = F\big(u(t)\big), \quad u(0) \text{ given}. \tag{2.2}$$

Here, I_ε denotes a diagonal matrix of dimension $m_1 + m_2$ with entries 1 or ε, respectively. The minimum regularity assumption on the function F is as follows.

HP 1. *Assume that F is differentiable and that its first derivative DF is locally Lipschitz-continuous.*

A basic concept for our proof of the convergence estimate stated in Theorem 4.1 is a linearization of the right-hand side of (2.2) along the exact solution. This yields the equation

$$I_\varepsilon u'(t) = F\big(u(t)\big) = A(t)u(t) + G\big(t, u(t)\big), \quad u(0) \text{ given}, \tag{2.3}$$

with time-dependent matrix $A = (A_{ij})_{1 \le i,j \le 2}$ where $A_{ij}(t) = D_j F_i(u(t))$. For some $v = (v_1, v_2)^{\mathrm{T}} \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$, the nonlinear function $G = (G_1, G_2)^{\mathrm{T}}$ is given by $G(t, v) = F(v) - A(t)v$, that is, $G_i(t, v) = F_i(v) - A_{i1}(t)v_1 - A_{i2}(t)v_2$.

Clearly, $A(t)$ is uniformly bounded for $0 \le t \le T$. The following assumption on the diagonal element A_{22} is essential for our stability and convergence analysis in Sections 3 and 4.

HP 2. *Suppose that for every $0 \le t \le T$ all eigenvalues $\lambda(t)$ of $A_{22}(t)$ have negative real parts $\Re(\lambda(t)) \le a_{22} < 0$.*

By multiplying Equation (2.3) with the inverse of I_ε, we alternatively obtain

$$u'(t) = \mathscr{F}\big(u(t)\big) = \mathscr{A}(t)u(t) + \mathscr{G}\big(t, u(t)\big), \quad u(0) \text{ given}, \tag{2.4}$$

where $\mathscr{F}(v) = I_\varepsilon^{-1}F(v)$, $\mathscr{A}(t) = I_\varepsilon^{-1}A(t)$ and $\mathscr{G}(t,v) = I_\varepsilon^{-1}G(t,v)$.

2.2 Variable stepsize linear multistep method.

In this section, we specify our hypotheses on the linear multistep method applied to the initial value problem (2.4).

Let $(h_j)_{j\geq 0}$ be a sequence of positive time steps with ratios $\omega_j = h_j/h_{j-1}$, $j \geq 1$, and associated grid points $t_j = h_0 + h_1 + \cdots + h_{j-1}$, $j \geq 0$. For given starting values $u_0, u_1, \ldots, u_{k-1}$, the numerical approximation u_{n+k} to the value of the exact solution at time t_{n+k}, $n \geq 0$, is determined by a linear k-step method, that is, u_{n+k} is given recursively by a relation of the form

$$\begin{aligned} \sum_{i=0}^{k} \alpha_{ni} u_{n+i} &= h_{n+k-1} \sum_{i=0}^{k} \beta_{ni} \mathscr{F}(u_{n+i}) \\ &= h_{n+k-1} \sum_{i=0}^{k} \beta_{ni} \big(\mathscr{A}(t_{n+i})u_{n+i} + \mathscr{G}(t_{n+i}, u_{n+i})\big), \quad n \geq 0, \end{aligned} \tag{2.5}$$

where the coefficients of the method α_{ni} and β_{ni}, $0 \leq i \leq k$, depend on the quantities $\omega_{n+1}, \omega_{n+2}, \ldots, \omega_{n+k-1}$. Throughout, the components of the numerical solution value u_j are denoted by $u_j = (y_j, z_j)^{\mathrm{T}} \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$, $j \geq 0$.

For the analysis, we employ a compact notation of the multistep method (2.5) as a one-step method for the vector $U_n = (u_n, u_{n+1}, \ldots, u_{n+k-1})^{\mathrm{T}} \in \mathbb{R}^{k\cdot m}$ comprising k consecutive numerical approximations. We introduce complex functions $s_j(z) = (\alpha_{j-1,k} - \beta_{j-1,k} z)^{-1}$ and $c_{ji}(z, \tilde{z}) = -s_j(z)(\alpha_{j-1,i} - \beta_{j-1,i}\tilde{z})$ for $j \geq 1$ and $0 \leq i \leq k-1$. Therewith, the companion matrix of the method equals

$$r_j(z_0, z_1, \ldots, z_k) = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ \vdots & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \ldots & 0 & 1 \\ c_{j0}(z_k, z_0) & c_{j1}(z_k, z_1) & \ldots & \ldots & c_{j,k-1}(z_k, z_{k-1}) \end{pmatrix}.$$

If $z_i = z$ for all $0 \leq i \leq k$, we set $r_j(z) = r_j(z, z, \ldots, z)$. The index j indicates the dependence of r_j on the coefficients $\alpha_{j-1,i}$ and $\beta_{j-1,i}$ of the method and thus on $\omega_j, \omega_{j+1}, \ldots, \omega_{j+k-2}$. For constant stepsizes, i.e. $\omega_j = 1$ for all $j \geq 1$, we omit the index and write r for short. Further, we let $\mathscr{A}_j = \mathscr{A}(t_j)$ for $j \geq 0$ and define $\mathscr{J}_j = s_j(h_{j+k-2}\mathscr{A}_{j+k-1})$ and

$$\mathscr{R}_j = r_j(h_{j+k-2}\mathscr{A}_{j-1}, h_{j+k-2}\mathscr{A}_j, \ldots, h_{j+k-2}\mathscr{A}_{j+k-1})$$

for $j \geq 1$. Besides, with $e_k = (0, \ldots, 0, 1)^{\mathrm{T}} \in \mathbb{R}^k$, we denote

$$\mathscr{G}(U_j) = \sum_{i=0}^{k} \beta_{j-1,i} e_k \otimes \mathscr{G}(t_{j+i-1}, u_{j+i-1}).$$

We recall that for matrices $B = (b_{ij})_{ij}$ and M, the (i, j)-th component of the Kronecker product $B \otimes M$ equals $b_{ij}M$.

With the above notation, the numerical scheme (2.5) becomes

$$U_{n+1} = \mathscr{R}_{n+1}U_n + h_{n+k-1}\mathscr{J}_{n+1}\mathscr{G}(U_{n+1}), \quad n \geq 0.$$

Solving this recursion yields the following relation

$$U_n = \prod_{i=1}^{n} \mathscr{R}_i U_0 + \sum_{j=1}^{n} h_{j+k-2} \prod_{i=j+1}^{n} \mathscr{R}_i \mathscr{J}_j \mathscr{G}(U_j), \quad n \geq 0, \tag{2.6}$$

a representation of the numerical solution by means of a discrete *variation-of-constants formula*.

Our hypotheses on the stepsize sequence and the linear multistep scheme rely on [9]. As in [6], we further suppose that the stepsizes are bounded from below by the parameter ε. For the definition of the notions of order and $A(\varphi)$-stability of a variable stepsize linear multistep method, we refer to [4, 5].

The following assumption on the stepsize ratios is fulfilled by classical step size selection procedures such as the differential/algebraic system solver DASSL based on backward differentiation formulas, see [11].

HS 1. *Suppose $h_j \geq \varepsilon$ for all $j \geq 0$. Assume further that there exists $\Omega > 1$ such that the stepsize ratios $\omega_j = h_j/h_{j-1}$ fulfill $\Omega^{-1} \leq \omega_j \leq \Omega$ for $j \geq 1$.*

The stability factors of the linear multistep method are of the form

$$C_j = D_1 \prod_{i=1}^{j} \big(1 + D_2|\omega_i - 1|\big)^2$$

with positive constants D_1 and D_2, see [9, Theorem 1]. In order to obtain meaningful stability and convergence estimates, we need these quantities to be bounded by a moderate constant.

HS 2. *Assume that the stability factors C_j of the linear multistep method are uniformly bounded by a constant for all $j \geq 1$ such that $t_j \leq T$.*

Besides, we suppose that the following stability requirement is satisfied for constant stepsizes. The angle $0 < \varphi < \pi/2$ is chosen in such a way that for some $a \in \mathbb{R}$ the spectrum of $A(t)$ is contained in the interior of the sector $S_\varphi(a) = \{\lambda \in \mathbb{C} : |\arg(a - \lambda)| \leq \varphi\} \cup \{a\}$ for all $0 \leq t \leq T$.

HM 1. *Assume that the linear multistep method* (2.5) *is $A(\varphi)$-stable and strictly stable at zero and infinity. Thus, $\lambda = 1$ is the only eigenvalue of the companion matrix at zero with modulus one, and the spectral radius of the companion matrix at infinity, $\sigma = \sigma(r(\infty))$, is less than one.*

Moreover, we make use of the following hypothesis for variable stepsizes.

HM 2. *Suppose that the coefficients α_{ji} and β_{ji} of the linear multistep scheme* (2.5) *are bounded for all stepsize sequences satisfying* HS1. *Assume further that the rational functions $s_j(z)$ and $c_{ji}(z)$ remain bounded for $z \in S_\varphi(0)$.*

We close this section with an assumption concerning the order of the method.

HM 3. *Assume that the multistep method* (2.5) *is consistent of order $p \geq 1$.*

2.3 Exact solution.

Throughout the paper, we employ the abbreviation $\hat{u}_j = u(t_j)$ for the value of the solution of (2.4) at time t_j, $j \geq 0$. Inserting the solution into the numerical scheme (2.5)

$$\begin{aligned}\sum_{i=0}^{k} \alpha_{ni}\hat{u}_{n+i} &= h_{n+k-1}\sum_{i=0}^{k}\beta_{ni}\big(\mathscr{F}(\hat{u}_{n+i}) + \delta_{n+1}\big)\\ &= h_{n+k-1}\sum_{i=0}^{k}\beta_{ni}\big(\mathscr{A}(t_{n+i})\hat{u}_{n+i} + \mathscr{G}(t_{n+i},\hat{u}_{n+i}) + \delta_{n+1}\big), \quad n \geq 0,\end{aligned}$$

defines the defect δ_{n+1} at t_{n+k}. In vector notation, we have

$$\widehat{U}_{n+1} = \mathscr{R}_{n+1}\widehat{U}_n + h_{n+k-1}\mathscr{I}_{n+1}\big(\mathscr{G}(\widehat{U}_{n+1}) + \Delta_{n+1}\big), \quad n \geq 0,$$

with $\widehat{U}_n = (\hat{u}_n, \hat{u}_{n+1}, \ldots, \hat{u}_{n+k-1})^{\mathrm{T}}$ and

$$\Delta_{n+1} = \sum_{i=0}^{k}\beta_{ni}e_k \otimes \delta_{n+1}.$$

As a consequence, we receive the analogue of (2.6) for the exact solution

$$\widehat{U}_n = \prod_{i=1}^{n}\mathscr{R}_i\,\widehat{U}_0 + \sum_{j=1}^{n}h_{j+k-2}\prod_{i=j+1}^{n}\mathscr{R}_i\mathscr{I}_j\big(\mathscr{G}(\widehat{U}_j) + \Delta_j\big), \quad n \geq 0. \tag{2.7}$$

Moreover, provided that the solution $u = (y,z)^{\mathrm{T}}$ of (2.4) is sufficiently smooth, the bounds

$$\begin{aligned}\big\|\delta_j^{(1)}\big\| &\leq Ch_{j+k-2}^{p-1}\int_{t_{j-1}}^{t_{j+k-1}}\big\|y^{(p+1)}(\tau)\big\|\,\mathrm{d}\tau,\\ \big\|\delta_j^{(2)}\big\| &\leq Ch_{j+k-2}^{p}\max_{t_{j-1}\leq\tau\leq t_{j+k-1}}\big\|z^{(p+1)}(\tau)\big\|,\end{aligned} \tag{2.8}$$

for the components $\delta_j^{(i)} \in \mathbb{R}^{m_i}$ of δ_j, $i = 1,2$, follow by means of a Taylor series expansion. Here, $\|\cdot\|$ denotes an arbitrary norm on $\mathbb{R}^{m_i}$. For notational simplicity, we do not consider different norms on $\mathbb{R}^{m_1}$ and $\mathbb{R}^{m_2}$.

2.4 Nonlinearity.

In the following, we state an auxiliary estimate for the nonlinear function G defined in (2.3). For simplicity, as indicated above, we endow $\mathbb{R}^{m_1}$ and $\mathbb{R}^{m_2}$ with the same norm $\|\cdot\|$ and define the norm on the product space $\mathbb{R}^m = \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$ through $\|v\| = \|v_1\| + \|v_2\|$ for $v = (v_1, v_2)^{\mathrm{T}} \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$.

LEMMA 2.1. *Under hypothesis* HP1*, there exists a constant $L > 0$ such that*

$$\|G_i(t,v) - G_i(t,w)\| \le L\varrho\|v - w\|, \quad i = 1,2,$$

for all $v, w \in \mathbb{R}^m$ satisfying $\|v - u(t)\| \le \varrho$ and $\|w - u(t)\| \le \varrho$.

PROOF. Fix $t \in [0,T]$ and consider a ball of radius $\varrho > 0$ around the value of the solution $u(t) = (y(t), z(t))^{\mathrm{T}}$. Due to the fact that DF is locally Lipschitz continuous by HP1, there exists $C > 0$ such that

$$\big\|D_jF_i(v) - D_jF_i(w)\big\| \le C\|v - w\|, \quad i,j = 1,2,$$

for all $v, w \in \mathbb{R}^m$ with $\|v - u(t)\| \le \varrho$ and $\|w - u(t)\| \le \varrho$. From the identity

$$\begin{aligned}
G_i(t,v) - G_i(t,w) &= F_i(v) - F_i(w) - A_{i1}(t)(v_1 - w_1) - A_{i2}(t)(v_2 - w_2)\\
&= \int_0^1 \big(D_1F_i(\sigma v_1 + (1-\sigma)w_1, v_2) - D_1F_i(y(t), z(t))\big)(v_1 - w_1)\,\mathrm{d}\sigma +\\
&\quad + \int_0^1 \big(D_2F_i(w_1, \sigma v_2 + (1-\sigma)w_2) - D_2F_i(y(t), z(t))\big)(v_2 - w_2)\,\mathrm{d}\sigma
\end{aligned}$$

the desired estimate follows with $L = 2C$. □

3 Stability estimates.

Throughout this section, we make use of the hypotheses and notation introduced in Section 2.

We next derive stability bounds for the linear multistep discretization (2.5) of (2.4). Hence, it suffices to consider the associated linear equation

$$u'(t) = \mathscr{A}(t)u(t), \tag{3.1}$$

where the linear multistep approximation simplifies to $U_n = \mathscr{R}_n\mathscr{R}_{n-1}\cdots\mathscr{R}_1U_0$. So, we study $\mathscr{R}_n\mathscr{R}_{n-1}\cdots\mathscr{R}_\ell X$ for arbitrary $X \in \mathbb{R}^{k\cdot m}$ and $1 \le \ell \le n$, and, in view of formula (2.6), also $h_{j+k-2}\mathscr{R}_n\mathscr{R}_{n-1}\cdots\mathscr{R}_{j+1}e_k \otimes \mathscr{J}_j(I_\varepsilon^{-1}x)$ for some $x \in \mathbb{R}^m$ and $1 \le j \le n$. Our basic idea is to compare $\mathscr{R}_i$ with the companion matrix

$$\mathscr{T}_i = r_i(h_{i+k-2}\mathscr{L}_{i-1}, h_{i+k-2}\mathscr{L}_i, \ldots, h_{i+k-2}\mathscr{L}_{i+k-1})$$

that corresponds to the lower triangular matrix $\mathscr{L}(t)$ resulting from $\mathscr{A}(t)$. In other words, we relate (3.1) to the partly coupled problem

$$u'(t) = \mathscr{L}(t)u(t), \quad \text{where } \mathscr{L}(t) = \begin{pmatrix} A_{11}(t) & 0 \\ \frac{1}{\varepsilon}A_{21}(t) & \frac{1}{\varepsilon}A_{22}(t) \end{pmatrix}. \tag{3.2}$$

In order to prove the necessary stability results for (3.2), as a first step, we consider in Section 3.1 the fully decoupled system

$$u'(t) = \mathscr{D}(t)u(t) \quad \text{with } \mathscr{D}(t) = \begin{pmatrix} A_{11}(t) & 0 \\ 0 & \frac{1}{\varepsilon} A_{22}(t) \end{pmatrix}. \tag{3.3}$$

In this special case, the desired stability estimates for the associated companion matrix

$$\mathscr{S}_i = r_i(h_{i+k-2}\mathscr{D}_{i-1}, h_{i+k-2}\mathscr{D}_i, \dots, h_{i+k-2}\mathscr{D}_{i+k-1})$$

are a consequence of the results given in [9].

3.1 The decoupled problem.

For studying the stability behaviour of the linear multistep method (2.5) applied to the decoupled equation (3.3), it is useful to consider each component of the numerical solution $u_n = (y_n, z_n)^{\mathrm{T}}$ separately, that is, we henceforth identify $U_n = (u_n, u_{n+1}, \dots, u_{n+k-1})^{\mathrm{T}}$ with the reordered vector $U_n = (Y_n, Z_n)^{\mathrm{T}}$ where $Y_n = (y_n, y_{n+1}, \dots, y_{n+k-1})^{\mathrm{T}}$ and $Z_n = (z_n, z_{n+1}, \dots, z_{n+k-1})^{\mathrm{T}}$. Hence, in this new order of components, for $X = (X_1, X_2)^{\mathrm{T}} \in \mathbb{R}^{k\cdot m_1} \times \mathbb{R}^{k\cdot m_2}$, we receive

$$\prod_{i=\ell}^{n} \mathscr{S}_i X = \prod_{i=\ell}^{n} \begin{pmatrix} R_i & 0 \\ 0 & S_i \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} \prod_{i=\ell}^{n} R_i X_1 & 0 \\ 0 & \prod_{i=\ell}^{n} S_i X_2 \end{pmatrix}. \tag{3.4}$$

Here, $R_i = r_i\big(h_{i+k-2}A_{11}(t_{i-1}), h_{i+k-2}A_{11}(t_i), \dots, h_{i+k-2}A_{11}(t_{i+k-1})\big)$ and

$$S_i = r_i\bigg(\frac{h_{i+k-2}}{\varepsilon}A_{22}(t_{i-1}), \frac{h_{i+k-2}}{\varepsilon}A_{22}(t_i), \dots, \frac{h_{i+k-2}}{\varepsilon}A_{22}(t_{i+k-1})\bigg)$$

denote the companion matrix associated with the first and the second component, respectively. The following result provides an estimate for (3.4). For later use, we further introduce $J_j = s_j(h_{j+k-2}A_{11}(t_{j+k-1}))$ and

$$K_{\varepsilon,j} = \frac{h_{j+k-2}}{\varepsilon} K_j = \frac{h_{j+k-2}}{\varepsilon} s_j\bigg(\frac{h_{j+k-2}}{\varepsilon} A_{22}(t_{j+k-1})\bigg).$$

An application of the integral formula of Cauchy as indicated in the proof of Lemma 3.1 shows the boundedness of J_j and $K_{\varepsilon,j}$, see also Remark 3.1 below.

LEMMA 3.1. *Under* HP2 *assume that the linear multistep discretization* (2.5) *of Equation* (3.3) *satisfies* HS1–2, HM1–3, *and further* $\sigma\Omega^2 < 1$. *Then, there exist* $H > 0$ *and* $C > 0$ *such that for any stepsize sequence* $(h_j)_{j\geq 0}$ *with* $0 < h_j \leq H$ *the following estimate holds with some* $0 < \gamma < 1$ *for* $n \geq 1$ *as long as* $t_{n+k-1} \leq T$

$$\bigg\| \prod_{i=\ell}^{n} \mathscr{S}_i X \bigg\| \leq C\|X_1\| + C\gamma^{n-\ell+1}\|X_2\|, \quad 1 \leq \ell \leq n.$$

In particular, the constant C *does not depend on* n, h_j *and* ε.

PROOF. Our proof is substantially based on the results and techniques from [9]. Owing to (3.4), it is sufficient to treat each component separately. For the first one, the bound $\|R_n R_{n-1} \cdots R_\ell\| \le C$ is a simple special case of [9, Theorem 3]. In order to estimate the second component, we compare $S_n S_{n-1} \cdots S_\ell$ with the frozen product $S_n^* S_{n-1}^* \cdots S_\ell^*$ where

$$S_i^* = r_i\Big(\frac{h_{i+k-2}}{\varepsilon} A_{22}^*\Big)$$

with fixed $A_{22}^* = A_{22}(t^*)$ for some $0 \le t^* \le T$. A telescopic identity for the difference $\Delta S_n = S_n S_{n-1} \cdots S_\ell - S_n^* S_{n-1}^* \cdots S_\ell^*$ yields

$$\Delta S_n = \sum_{j=\ell}^{n} \prod_{i=j+1}^{n} S_i^* \,(S_j - S_j^*)\, \Delta S_{j-1} + \sum_{j=\ell}^{n} \prod_{i=j+1}^{n} S_i^* \,(S_j - S_j^*) \prod_{i=\ell}^{j-1} S_i^* . \tag{3.5}$$

We recall that all eigenvalues of A_{22}^* are strictly negative by hypothesis HP2. With the help of Cauchy's integral formula, we thus receive the representation

$$\prod_{i=\ell}^{n} S_i^* = \frac{1}{2\pi\iota} \int_\Gamma \prod_{i=\ell}^{n} r_i\Big(\frac{h_{i+k-2}\lambda}{\varepsilon}\Big) (\lambda I - A_{22}^*)^{-1} \,\mathrm{d}\lambda \tag{3.6}$$

with a finite path $\Gamma \subset \mathbb{C}_{<0}$ contained in the negative complex plane that encircles the eigenvalues of A_{22}^*. By [9, Lemma 2], for some $0 < \gamma < 1$, it holds

$$\Big\| \prod_{i=\ell}^{n} r_i\Big(\frac{h_{i+k-2}\lambda}{\varepsilon}\Big) \Big\| \le C\gamma^{n-\ell+1}.$$

Using that $(\lambda I - A_{22}^*)^{-1}$ is bounded, we therefore obtain from (3.6)

$$\Big\| \prod_{i=\ell}^{n} S_i^* \Big\| \le C\gamma^{n-\ell+1}. \tag{3.7}$$

In addition, a comparison of S_i with S_i^* shows the boundedness of S_i. Consequently, by estimating (3.5), after slightly increasing $\gamma < 1$, we have

$$\|\Delta S_n\| \le C \sum_{j=\ell}^{n} \gamma^{n-j+1} \|\Delta S_{j-1}\| + C\gamma^{n-\ell+1}.$$

Now, a Gronwall inequality yields the bound $\|\Delta S_n\| \le C\gamma^{n-\ell+1}$, and, finally, another application of (3.7) gives the desired result. □

We next summarize some useful relations for the quantities J_j, $R_n R_{n-1} \cdots R_\ell$, $K_{\varepsilon,j}$, and $S_n S_{n-1} \cdots S_\ell$, see (3.4) and below. Note that the specified estimates for $R_n R_{n-1} \cdots R_\ell$, and $S_n S_{n-1} \cdots S_\ell$ are a direct consequence of Lemma 3.1.

Further, the boundedness of J_j follows in an easy way from Cauchy's integral formula. As $\frac{h_{j+k-2}}{\varepsilon}(1+\frac{h_{j+k-2}}{\varepsilon})^{-1} \le 1$, an estimation of

$$K_{\varepsilon,j} = \frac{1}{2\pi\iota}\int_\Gamma \frac{h_{j+k-2}}{\varepsilon}\left(\alpha_{j-1,k} - \beta_{j-1,k}\frac{h_{j+k-2}\lambda}{\varepsilon}\right)^{-1}\big(\lambda I - A_{22}(t_{j+k-1})\big)^{-1}\mathrm{d}\lambda$$

shows that $K_{\varepsilon,j}$ is bounded.

Remark 3.1. In the situation of Lemma 3.1, for any $1 \le \ell, j \le n$, the bounds

$$\|J_j\| \le C, \quad \left\|\prod_{i=\ell}^{n} R_i\right\| \le C, \quad \|K_{\varepsilon,j}\| \le C, \quad \text{and} \quad \left\|\prod_{i=\ell}^{n} S_i\right\| \le C\gamma^{n-\ell+1}$$

are valid with constants $C > 0$ and $0 < \gamma < 1$.

We close this section with a remark on BDF-methods where $\beta_{ji} = 0$ for all $0 \le i \le k-1$ and $j \ge 0$ and thus $\sigma = 0$. In particular, the condition $\sigma\Omega^2 < 1$ of Lemma 3.1 is fulfilled for any $\Omega > 1$. Here, a further investigation of $S_n S_{n-1}\cdots S_1$ shows that the sharper estimate

$$\left\|\prod_{i=1}^{n} S_i\right\| \le C\varepsilon\frac{\gamma^n}{h_{k-1}} \tag{3.8}$$

is valid for $n \ge k$.

3.2 The partly coupled problem.

We are now ready to estimate the linear multistep approximation of the partly coupled equation (3.2). Following the lines of the previous section, we employ an alternative representation of $U_n = \mathscr{T}_n\mathscr{T}_{n-1}\cdots\mathscr{T}_1 U_0$ by determining successively Y_n and Z_n. For the y-component, it clearly holds $Y_n = R_n R_{n-1}\cdots R_1 Y_0$. Thus, by applying the discrete variation-of-constants formula to the z-component and inserting the above representation for Y_j, Z_n writes as

$$Z_n = \prod_{i=1}^{n} S_i Z_0 + \sum_{j=1}^{n}\prod_{i=j+1}^{n} S_i e_k \otimes K_{\varepsilon,j}\big(B_j + \widetilde{B}_j R_j\big)\prod_{i=1}^{j-1} R_i Y_0$$

with $B_j = (\beta_{j-1,0}A_{21}(t_{j-1}), \frac{1}{2}\beta_{j-1,1}A_{21}(t_j), \ldots, \frac{1}{2}\beta_{j-1,k-1}A_{21}(t_{j+k-2}))$ and also $\widetilde{B}_j = (\frac{1}{2}\beta_{j-1,1}\,A_{21}(t_j), \ldots, \frac{1}{2}\beta_{j-1,k-1}A_{21}(t_{j+k-2}), \beta_{j-1,k}A_{21}(t_{j+k-1}))$ denoting a bounded matrix of dimension $m_2 \times k \cdot m_1$. In particular, for $k = 1$, let $B_j = \beta_{j-1,0}A_{21}(t_{j-1})$ and $\widetilde{B}_j = \beta_{j-1,1}A_{21}(t_j)$. Henceforth, we identify the transfer operator $\mathscr{T}_n\mathscr{T}_{n-1}\cdots\mathscr{T}_\ell$ with

$$\prod_{i=\ell}^{n}\mathscr{T}_i = \begin{pmatrix} \prod_{i=\ell}^{n} R_i & 0 \\ P_{n\ell} & \prod_{i=\ell}^{n} S_i \end{pmatrix},$$

where the quantity $P_{n\ell}$ is defined through

$$P_{n\ell} = \sum_{j=\ell}^{n} \prod_{i=j+1}^{n} S_i \, e_k \otimes K_{\varepsilon,j} \left(B_j + \widetilde{B}_j R_j\right) \prod_{i=\ell}^{j-1} R_i.$$

With the help of Remark 3.1, it is easy to see that $P_{n\ell}$ is bounded by a constant, and, consequently, we obtain

$$\begin{aligned} \left\| \prod_{i=\ell}^{n} \mathscr{T}_i X \right\| &\le \left(\left\| \prod_{i=\ell}^{n} R_i \right\| + \|P_{n\ell}\| \right) \|X_1\| + \left\| \prod_{i=\ell}^{n} S_i \right\| \|X_2\| \\ &\le C\|X_1\| + C\gamma^{n-\ell+1}\|X_2\|. \end{aligned}$$

This proves the following result.

LEMMA 3.2. *In the situation of Lemma* 3.1, *the linear multistep discretization* (2.5) *of Equation* (3.2) *fulfills the estimate*

$$\left\| \prod_{i=\ell}^{n} \mathscr{T}_i X \right\| \le C\|X_1\| + C\gamma^{n-\ell+1}\|X_2\|, \quad 1 \le \ell \le n,$$

with constant C independent of n, h_j and ε.

We note for later use that for BDF-methods the refined estimate

$$\left\| \prod_{i=1}^{n} \mathscr{T}_i X \right\| \le C\|X_1\| + C\varepsilon \frac{\gamma^n}{h_{k-1}} \|X_2\|, \quad n \ge k, \tag{3.9}$$

follows from (3.8).

3.3 The coupled problem.

In the following, we derive stability estimates for the linear multistep approximation $U_n = \mathscr{R}_n \mathscr{R}_{n-1} \cdots \mathscr{R}_1 U_0$ of the original coupled equation (3.1). As in the preceding sections, we henceforth identify $U_n = (u_n, u_{n+1}, \ldots, u_{n+k-1})^{\mathrm{T}}$ comprising the solution values $u_n = (y_n, z_n)^{\mathrm{T}}$ with the reordered vector $U_n = (Y_n, Z_n)^{\mathrm{T}}$ where $Y_n = (y_n, y_{n+1}, \ldots, y_{n+k-1})^{\mathrm{T}}$ and $Z_n = (z_n, z_{n+1}, \ldots, z_{n+k-1})^{\mathrm{T}}$. Accordingly to that new order, we interprete elements $X = (X_1, X_2)^{\mathrm{T}} \in \mathbb{R}^{k \cdot m_1} \times \mathbb{R}^{k \cdot m_2}$. Further, let $x = (x_1, x_2)^{\mathrm{T}} \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$.

In order to prove stability results for the transfer operator $\mathscr{R}_n \mathscr{R}_{n-1} \cdots \mathscr{R}_\ell$ of (3.1), we make use of the fact that bounds for $\mathscr{T}_n \mathscr{T}_{n-1} \cdots \mathscr{T}_\ell$ are provided by Lemma 3.2. Thus, it suffices to study $\Delta\mathscr{R}_n = \mathscr{R}_n \mathscr{R}_{n-1} \cdots \mathscr{R}_\ell - \mathscr{T}_n \mathscr{T}_{n-1} \cdots \mathscr{T}_\ell$. For estimating this difference, our basic tool is the telescopic identity

$$\Delta\mathscr{R}_n = \sum_{j=\ell}^{n} \prod_{i=j+1}^{n} \mathscr{T}_i (\mathscr{R}_j - \mathscr{T}_j) \, \Delta\mathscr{R}_{j-1} + \sum_{j=\ell}^{n} \prod_{i=j+1}^{n} \mathscr{T}_i \, (\mathscr{R}_j - \mathscr{T}_j) \prod_{i=\ell}^{j-1} \mathscr{T}_i \tag{3.10}$$

combined with a Gronwall inequality. Therewith, we are able to establish the following stability bounds.

Theorem 3.3. *In the situation of Lemma* 3.1, *the linear multistep discretization* (2.5) *of Equation* (3.1) *satisfies the relations*

$$\begin{aligned} \Big\| \prod_{i=\ell}^{n} \mathscr{R}_i X \Big\| &\le C\|X_1\| + C\big(h_{\ell+k-2} + \gamma^{n-\ell+1}\big)\|X_2\|, \\ h_{j+k-2} \Big\| \prod_{i=j+1}^{n} \mathscr{R}_i e_k \otimes \mathscr{J}_j\big(I_\varepsilon^{-1}x\big) \Big\| &\le C h_{j+k-2}\|x_1\| + C\big(h_{j+k-2} + \gamma^{n-j}\big)\|x_2\|, \end{aligned}$$

for $1 \le \ell, j \le n$ *with constants* C *independent of* n, h_j *and* ε.

Proof. In order to estimate (3.10), we first indicate the derivation of a needful relation for the difference $\mathscr{R}_j - \mathscr{T}_j$ of the companion matrices associated with the fully and partly coupled equations (3.1) and (3.2). By definition, it holds

$$\mathscr{R}_j - \mathscr{T}_j = \begin{pmatrix} 0 & 0 & \dots & 0 \\ \Delta c_{j0} & \Delta c_{j1} & \dots & \Delta c_{j,k-1} \end{pmatrix}$$

where the entries Δc_{ji} are given by

$$\begin{aligned} \Delta c_{ji} &= c_{ji}(h_{j+k-2}\mathscr{A}_{j+k-1}, h_{j+k-2}\mathscr{A}_{j+i-1}) - \\ &\quad - c_{ji}(h_{j+k-2}\mathscr{L}_{j+k-1}, h_{j+k-2}\mathscr{L}_{j+i-1}), \quad 0 \le i \le k-1, \end{aligned}$$

see beginning of Section 2.2. With the help of the quantity

$$\mathscr{K}_j = s_j(h_{j+k-2}\mathscr{L}_{j+k-1}) = \begin{pmatrix} J_j & 0 \\ \beta_{j-1,k} K_{\varepsilon,j} A_{21}(t_{j+k-1}) J_j & K_j \end{pmatrix}$$

which is bounded according to Remark 3.1, Δc_{ji} also writes as

$$\begin{aligned} \Delta c_{ji} &= -\alpha_{j-1,i}\big(\mathscr{J}_j - \mathscr{K}_j\big) + h_{j+k-2}\beta_{j-1,i}\big(\mathscr{J}_j - \mathscr{K}_j\big)\mathscr{A}_{j+i-1} + \\ &\quad + h_{j+k-2}\beta_{j-1,i}\mathscr{K}_j\big(\mathscr{A}_{j+i-1} - \mathscr{L}_{j+i-1}\big). \end{aligned}$$

A straightforward calculation shows the identity $\mathscr{J}_j = (I + h_{j+k-2}\mathscr{B}_j)\mathscr{K}_j$ where $D_j = (I - h_{j+k-2}\beta_{j-1,k}^2 K_{\varepsilon,j} A_{21}(t_{j+k-1}) J_j A_{12}(t_{j+k-1}))^{-1}$ and thus

$$\mathscr{B}_j = \begin{pmatrix} 0 & B_{j1} \\ 0 & B_{j2} \end{pmatrix} = \begin{pmatrix} 0 & \beta_{j-1,k} J_j A_{12}(t_{j+k-1}) D_j \\ 0 & \beta_{j-1,k}^2 K_{\varepsilon,j} A_{21}(t_{j+k-1}) J_j A_{12}(t_{j+k-1}) D_j \end{pmatrix}$$

is bounded for $h_{j+k-2} \le H$ sufficiently small, see again Remark 3.1. Thus, we obtain the relation $\Delta c_{ji} = h_{j+k-2}(\mathscr{B}_{ji}\mathscr{K}_j + \widetilde{\mathscr{B}}_{ji})$ with $\mathscr{B}_{ji} = -\alpha_{j-1,i}\mathscr{B}_j$ and $\widetilde{\mathscr{B}}_{ji} = \beta_{j-1,i}(h_{j+k-2}\mathscr{B}_j\mathscr{K}_j\mathscr{A}_{j+i-1} + \mathscr{K}_j(\mathscr{A}_{j+i-1} - \mathscr{L}_{j+i-1}))$ bounded. Finally, this yields the identity

$$\mathscr{R}_j - \mathscr{T}_j = h_{j+k-2}\mathscr{C}_j, \quad \text{where } \mathscr{C}_j = \begin{pmatrix} 0 & 0 & \dots & 0 \\ \mathscr{C}_{j0} & \mathscr{C}_{j1} & \dots & \mathscr{C}_{j,k-1} \end{pmatrix} \tag{3.11}$$

comprises the bounded entries $\mathscr{C}_{ji} = \mathscr{B}_{ji}\mathscr{K}_j + \widetilde{\mathscr{B}}_{ji}$.

Now, a thorough investigation of the second term in (3.10) and a further application of Remark 3.1 shows that

$$\left\| \sum_{j=\ell}^{n} \prod_{i=j+1}^{n} \mathscr{T}_i \left(\mathscr{R}_j - \mathscr{T}_j\right) \prod_{i=\ell}^{j-1} \mathscr{T}_i X \right\| \leq C\|X_1\| + C\big(h_{\ell+k-2} + \gamma^{n-\ell+1}\big)\|X_2\|,$$

and, altogether, we obtain

$$\begin{aligned} \|\Delta\mathscr{R}_n X\| \ &\leq \ C \sum_{j=\ell}^{n} h_{j+k-2}(1+\gamma^{n-j})\|\Delta\mathscr{R}_{j-1} X\| + \\ &\quad + C\|X_1\| + C\big(h_{\ell+k-2} + \gamma^{n-\ell+1}\big)\|X_2\|. \end{aligned} \tag{3.12}$$

For estimating $\Delta\mathscr{R}_n X$, we next split $X = (X_1, 0)^{\mathrm{T}} + (0, X_2)^{\mathrm{T}}$ and replace (3.12) with the following two inequalities. On the one hand, it holds

$$\big\|\Delta\mathscr{R}_n (X_1, 0)^{\mathrm{T}}\big\| \leq C \sum_{j=\ell}^{n} h_{j+k-2} \big\|\Delta\mathscr{R}_{j-1}(X_1, 0)^{\mathrm{T}}\big\| + C\|X_1\| \tag{3.13a}$$

and, on the other hand, we have

$$\begin{aligned} \big\|\Delta\mathscr{R}_n (0, X_2)^{\mathrm{T}}\big\| \ &\leq \ C \sum_{j=\ell}^{n} h_{j+k-2}(1+\gamma^{n-j}) \big\|\Delta\mathscr{R}_{j-1}(0, X_2)^{\mathrm{T}}\big\| + \\ &\quad + C\big(h_{\ell+k-2} + \gamma^{n-\ell+1}\big)\|X_2\|. \end{aligned} \tag{3.13b}$$

Now, at each time, the desired bound results from a discrete Gronwall-type inequality, see [1, 2], e.g. In fact, for (3.13a), the estimate

$$\big\|\Delta\mathscr{R}_n (X_1, 0)^{\mathrm{T}}\big\| \leq C\|X_1\| \tag{3.14a}$$

follows at once from a standard Gronwall inequality. In view of (3.13b), we consider a sequence $(\xi_j)_{j\geq \ell-1}$ of positive numbers satisfying a relation of the following form involving constants $a, b > 0$

$$\xi_n = a \sum_{j=\ell}^{n} h_{j+k-2}(1+\gamma^{n-j})\xi_{j-1} + b\big(h_{\ell+k-2} + \gamma^{n-\ell+1}\big).$$

Due to the fact that this identity is reducible to the recursion

$$\xi_{n+1} = (2ah_{n+k-1} + \gamma)\xi_n + a(1-\gamma) \sum_{j=\ell-1}^{n-1} h_{j+k-1}\xi_j + b(1-\gamma)h_{\ell+k-2},$$

we further obtain $\xi_n \leq Cb(h_{\ell+k-2} + \gamma^{n-\ell+1})$ which proves

$$\big\|\Delta\mathscr{R}_n (0, X_2)^{\mathrm{T}}\big\| \leq C\big(h_{\ell+k-2} + \gamma^{n-\ell+1}\big)\|X_2\|. \tag{3.14b}$$

Therefore, by combining (3.14) and Lemma 3.2, the first bound of Theorem 3.3 follows.

It remains to derive the second bound of the theorem. An easy calculation shows the identity

$$\begin{aligned} &h_{j+k-2}\mathscr{J}_j\big(I_\varepsilon^{-1}x\big) \\ &\quad= \begin{pmatrix} h_{j+k-2}J_jx_1+ \\ h_{j+k-2}\beta_{j-1,k}K_{\varepsilon,j}A_{21}(t_{j+k-1})J_jx_1+K_{\varepsilon,j}x_2 \end{pmatrix} + \\ &\quad\quad+h_{j+k-2}\begin{pmatrix} h_{j+k-2}B_{j1}\beta_{j-1,k}K_{\varepsilon,j}A_{21}(t_{j+k-1})J_jx_1+B_{j1}K_{\varepsilon,j}x_2 \\ h_{j+k-2}B_{j2}\beta_{j-1,k}K_{\varepsilon,j}A_{21}(t_{j+k-1})J_jx_1+B_{j2}K_{\varepsilon,j}x_2 \end{pmatrix}. \end{aligned}$$

Thus, the first bound applied with $X = h_{j+k-2}\mathscr{J}_j(I_\varepsilon^{-1}x)$ and $\ell = j+1$ proves the desired result. □

In view of our convergence estimate for BDF-methods, the first relation of Theorem 3.3 with $\ell = 1$ is replaced by

$$\Big\|\prod_{i=1}^{n}\mathscr{R}_iX\Big\| \le C\|X_1\| + C\varepsilon\Big(1+\frac{\gamma^n}{h_{k-1}}\Big)\|X_2\|, \quad n \ge k. \tag{3.15}$$

This sharper bound is obtained by modifying slightly the proof of Theorem 3.3. In the present situation, formula (3.11) holds for $\mathscr{C}_{ji} = \mathscr{B}_{ji}\mathscr{K}_j$. As a consequence, $\mathscr{C}_{ji}$ is of the form

$$\mathscr{C}_{ji} = \begin{pmatrix} C_{11} & C_{12}K_j \\ C_{21} & C_{22}K_j \end{pmatrix} = \begin{pmatrix} C_{11} & \frac{\varepsilon}{h_{j+k-2}}C_{12}K_{\varepsilon,j} \\ C_{21} & \frac{\varepsilon}{h_{j+k-2}}C_{22}K_{\varepsilon,j} \end{pmatrix}$$

with bounded matrices C_{11}, C_{12}, C_{21}, and C_{22}. Following the above proof of Theorem 3.3 and tracing the z-component together with (3.9) then yields the refined stability bound for BDF-methods.

4 Convergence result.

In this section, we state our convergence estimate for variable stepsize linear multistep methods applied to singular perturbation problems of the form (2.4). For some function ϕ denote

$$\|\phi\|_{1,[\tau_1,\tau_2]} = \int_{\tau_1}^{\tau_2}\|\phi(\tau)\|\,\mathrm{d}\tau \quad\text{and}\quad \|\phi\|_{\infty,[\tau_1,\tau_2]} = \max_{\tau_1\le\tau\le\tau_2}\|\phi(\tau)\|.$$

Then, the following result holds, provided that the $(p+1)$-st order derivatives of the solution $u = (y,z)^{\mathrm{T}}$ of (2.4) remain bounded, precisely, if the bounds

$$\big\|y^{(p+1)}\big\|_{1,[t_{j-k},t_j]} \le C \quad\text{and}\quad \big\|z^{(p+1)}\big\|_{\infty,[t_{j-k},t_j]} \le C \tag{4.1}$$

are valid for every $t_j \le T$ with constants $C > 0$ not depending on the parameter ε for $\varepsilon \in (0,\varepsilon_0]$.

THEOREM 4.1. *Under* HP2 *assume that the solution* $u = (y, z)^{\mathrm{T}}$ *of the initial value problem* (2.4) *is sufficiently often differentiable and that its derivatives fulfill* (4.1). *Assume further that the linear multistep discretization* (2.5) *of* (2.4) *satisfies* HS1–2, HM1–3, *and* $\sigma\Omega^2 < 1$. *Then, there exist* $H > 0$, $d > 0$ *and* $C > 0$ *such that for any stepsize sequence* $(h_j)_{j\geq 0}$ *with* $0 < h_j \leq H$ *and for initial values satisfying* $\|y_i - y(t_i)\| + \|z_i - z(t_i)\| \leq d$ *for each* $0 \leq i \leq k-1$ *the estimate*

$$\begin{aligned}\|u_n - u(t_n)\| \ \leq\ & C \max_{0\leq i\leq k-1} \|y_i - y(t_i)\| + C\big(h_{k-1} + \gamma^n\big) \max_{0\leq i\leq k-1} \|z_i - z(t_i)\| + \\ & + C \sum_{j=k}^{n} h_{j-1}^p \big\|y^{(p+1)}\big\|_{1,[t_{j-k},t_j]} + \\ & + \varepsilon C \sum_{j=k}^{n} \big(h_{j-1} + \gamma^{n-j}\big) h_{j-1}^p \big\|z^{(p+1)}\big\|_{\infty,[t_{j-k},t_j]}\end{aligned}$$

is valid with some $0 < \gamma < 1$ *for all* $n \geq k$ *as long as* $t_n \leq T$. *Especially, the constant* C *does not depend on* n, $(h_j)_{j\geq 0}$ *and* ε.

REMARK 4.1. Note that for BDF-methods, due to $\sigma = 0$, the requirement $\sigma\Omega^2 < 1$ is satisfied for any $\Omega > 1$. Here, a refined convergence estimate is valid, namely, the factor $h_{k-1} + \gamma^n$ multiplying the z-component of the errors in the starting values, is replaced with $\varepsilon(1 + \frac{\gamma^n}{h_{k-1}})$. This result follows at once from the proof of Theorem 4.1 by estimating (4.3) with the help of relation (3.15).

REMARK 4.2. In particular, for constant or bounded stepsizes $h_j \leq h$, $j \geq 0$, we receive the convergence estimate

$$\begin{aligned}\|u_n - u(t_n)\| \ \leq\ & C \max_{0\leq i\leq k-1} \|y_i - y(t_i)\| + C\big(h + \gamma^n\big) \max_{0\leq i\leq k-1} \|z_i - z(t_i)\| + \\ & + Ch^p \big\|y^{(p+1)}\big\|_{1,[0,t_n]} + \varepsilon\, Ch^p \big\|z^{(p+1)}\big\|_{\infty,[0,t_n]}, \quad t_n \leq T,\end{aligned}$$

which is in accordance with the bound from [6, Theorem 3] for a constant stepsize linear multistep method.

PROOF OF THEOREM 4.1. For constructing the linear multistep solution (2.5) of (2.4), we carry out a fixed-point iteration based on the discrete variation-of-constants formula (2.6). Thereto, we first introduce some useful notation.

For $N \in \mathbb{N}$ such that $t_{N+k-1} \leq T$, consider a sequence $\mathbf{V} = (V_n)_{n=0}^N$ comprising the vectors $V_n = (v_n, v_{n+1}, \ldots, v_{n+k-1})^{\mathrm{T}} \in \mathbb{R}^{k\cdot m}$ with entries

$$v_j = \big(v_j^{(1)}, v_j^{(2)}\big)^{\mathrm{T}} \in \mathbb{R}^m = \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}, \quad j \geq 0.$$

In particular, we denote by $\mathbf{U} = (U_n)_{n=0}^N$ and $\widehat{\mathbf{U}} = (\widehat{U}_n)_{n=0}^N$ the sequences that comprise the numerical and exact solution values. We recall the notation $u_n = (y_n, z_n)^{\mathrm{T}}$ for the components of the numerical solution. As in Section 3, we henceforth identify the associated vector $U_n = (u_n, u_{n+1}, \ldots, u_{n+k-1})^{\mathrm{T}}$ with

$U_n = (Y_n, Z_n)^{\mathrm{T}} = (y_n, \ldots, y_{n+k-1}, z_n, \ldots, z_{n+k-1})^{\mathrm{T}}$. Likewise, we employ the notation $\hat{u}_n = (\hat{y}_n, \hat{z}_n)^{\mathrm{T}}$ for the values of the exact solution and define $\widehat{U}_n = (\hat{u}_n, \hat{u}_{n+1}, \ldots, \hat{u}_{n+k-1})^{\mathrm{T}} = (\widehat{Y}_n, \widehat{Z}_n)^{\mathrm{T}}$. In accordance with the preceding sections, we further set

$$\|v_n\| = \|v_n^{(1)}\| + \|v_n^{(2)}\| \quad \text{for } v_n = \big(v_n^{(1)}, v_n^{(2)}\big)^{\mathrm{T}} \in \mathbb{R}^m = \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$$

and define the vector norm through

$$\|V_n\| = \max_{0 \le i \le k-1} \|v_{n+i}\|.$$

In order to guarantee the contraction property of the fixed-point iteration for a reasonable time, we introduce additional weights in the sequence norm and set

$$\|\mathbf{V}\|_{\infty,\mu} = \max_{0 \le n \le N} \|V_n\|_\mu, \quad \text{where } \|V_n\|_\mu = \mathrm{e}^{-\mu t_n} \|V_n\|$$

for some exponent $\mu > 0$ sufficiently large.

With the help of these abbreviations, we are in the position to define the fixed point iteration Φ on a ball around $\widehat{\mathbf{U}}$ by means of formula (2.6)

$$\begin{aligned} \Phi\colon \mathscr{V} &= \big\{\mathbf{V} = (V_n)_{n=0}^N : \|\mathbf{V} - \widehat{\mathbf{U}}\|_{\infty,\mu} \le \varrho\big\} \longrightarrow \mathscr{V} : \mathbf{V} \longmapsto \Phi(\mathbf{V}), \\ \Phi(\mathbf{V})_n &= \prod_{i=1}^n \mathscr{R}_i U_0 + \sum_{j=1}^n h_{j+k-2} \prod_{i=j+1}^n \mathscr{R}_i \mathscr{J}_j \mathscr{G}(V_j). \end{aligned}$$

It remains to verify that the function Φ defining the iteration is a contraction on $\mathscr{V}$. On the one hand, we have for sequences $\mathbf{V}, \mathbf{W} \in \mathscr{V}$

$$\begin{aligned} \big(\Phi(\mathbf{V}) - \Phi(\mathbf{W})\big)_n &= \sum_{j=1}^n h_{j+k-2} \prod_{i=j+1}^n \mathscr{R}_i\, e_k \otimes \mathscr{J}_j\big(I_\varepsilon^{-1} x\big), \quad \text{where} \\ x &= \sum_{i=0}^k \beta_{j-1,i}\big(G(t_{j+i-1}, v_{j+i-1}) - G(t_{j+i-1}, w_{j+i-1})\big). \end{aligned}$$

An application of Lemma 2.1 shows

$$\big\|G_\ell(t_{j+i-1}, v_{j+i-1}) - G_\ell(t_{j+i-1}, w_{j+i-1})\big\| \le C \mathrm{e}^{\mu t_j} \varrho \, \|\mathbf{V} - \mathbf{W}\|_{\infty,\mu}, \quad \ell = 1, 2.$$

Together with the second estimate from Theorem 3.3, this yields

$$\big\|\big(\Phi(\mathbf{V}) - \Phi(\mathbf{W})\big)_n\big\|_\mu \le C\varrho \sum_{j=1}^n \mathrm{e}^{-\mu(t_n - t_j)} \big(h_{j+k-2} + \gamma^{n-j}\big) \|\mathbf{V} - \mathbf{W}\|_{\infty,\mu},$$

and, furthermore,

$$\|\Phi(\mathbf{V}) - \Phi(\mathbf{W})\|_{\infty,\mu} \le \kappa \|\mathbf{V} - \mathbf{W}\|_{\infty,\mu},$$

that is, Φ is contractive with contraction factor

$$\kappa = C\varrho \max_{0\le n\le N} \sum_{j=1}^{n} \mathrm{e}^{-\mu(t_n-t_j)} \big(h_{j+k-2} + \gamma^{n-j}\big) < 1$$

for ϱ sufficiently small. We note that the size of $C > 0$ is moderate for μ large, whereas for $\mu = 0$ the above relation becomes

$$\kappa = C\varrho\left(T + \frac{1}{1-\gamma}\right) < 1.$$

As a consequence, this condition considerably restricts the size of T.

We next prove that Φ maps $\mathscr{V}$ to $\mathscr{V}$, that is,

$$\big\|\Phi(\mathbf{V}) - \widehat{\mathbf{U}}\big\|_{\infty,\mu} \le \varrho \quad \text{whenever} \quad \|\mathbf{V} - \widehat{\mathbf{U}}\|_{\infty,\mu} \le \varrho.$$

By means of the contraction property of Φ on $\mathscr{V}$, we obtain

$$\begin{aligned}\big\|\Phi(\mathbf{V}) - \widehat{\mathbf{U}}\big\|_{\infty,\mu} &\le \big\|\Phi(\mathbf{V}) - \Phi(\widehat{\mathbf{U}})\big\|_{\infty,\mu} + \big\|\Phi(\widehat{\mathbf{U}}) - \widehat{\mathbf{U}}\big\|_{\infty,\mu} \\ &\le \kappa\varrho + \big\|\Phi(\widehat{\mathbf{U}}) - \widehat{\mathbf{U}}\big\|_{\infty,\mu}.\end{aligned}$$

Thus, it suffices to show that the quantity

$$\begin{aligned}\big(\Phi(\widehat{\mathbf{U}}) - \widehat{\mathbf{U}}\big)_n &= \prod_{i=1}^{n} \mathscr{R}_i\big(U_0 - \widehat{U}_0\big) - \sum_{j=1}^{n} h_{j+k-2} \prod_{i=j+1}^{n} \mathscr{R}_i\, e_k \otimes \mathscr{J}_j I_\varepsilon^{-1} x \\ &\qquad \text{with } x = \sum_{i=0}^{k} \beta_{j-1,i} I_\varepsilon \delta_j\end{aligned}$$

is small enough, see (2.7). With the help of Theorem 3.3 and the bound (2.8) for the defects, it follows

$$\begin{aligned}\big\|\Phi(\widehat{\mathbf{U}})_n - \widehat{\mathbf{U}}_n\big\|_{\mu} \le\ & C\mathrm{e}^{-\mu t_n}\Bigg(\big\|Y_0 - \widehat{Y}_0\big\| + \big(h_{k-1} + \gamma^n\big)\big\|Z_0 - \widehat{Z}_0\big\| + \\ & + \sum_{j=1}^{n} h_{j+k-2}^p \big\|y^{(p+1)}\big\|_{1,[t_{j-1},t_{j+k-1}]} + \\ & + \varepsilon \sum_{j=1}^{n} h_{j+k-2}^p \big(h_{j+k-2} + \gamma^{n-j}\big) \times \\ & \times \big\|z^{(p+1)}\big\|_{\infty,[t_{j-1},t_{j+k-1}]}\Bigg).\end{aligned} \tag{4.2}$$

Taking the maximum over $0 \le n \le N$ finally gives

$$\big\|\Phi(\widehat{\mathbf{U}}) - \widehat{\mathbf{U}}\big\|_{\infty,\mu} \le (1-\kappa)\varrho$$

if the errors of the initial values and $h_j \le H$ are sufficiently small. Altogether, an application of Banach's Fixed Point Theorem yields the existence of the numerical solution $\mathbf{U}$ as unique fixed point of Φ.

In order to estimate $E_n = \|U_n - \widehat{U}_n\|$, we employ the following representation obtained by formulas (2.6) and (2.7)

$$U_n - \widehat{U}_n = \prod_{i=1}^{n} \mathscr{R}_i \big(U_0 - \widehat{U}_0\big) + + \sum_{j=1}^{n} h_{j+k-2} \prod_{i=j+1}^{n} \mathscr{R}_i \mathscr{J}_j \big(\mathscr{G}(U_j) - \mathscr{G}(\widehat{U}_j) - \Delta_j\big). \tag{4.3}$$

Now, the above considerations together with the bounds from Theorem 3.3, Lemma 2.1, and (2.8) show that the error satisfies

$$\begin{aligned} E_n \le\ & C\|Y_0 - \widehat{Y}_0\| + C\big(h_{k-1} + \gamma^n\big)\|Z_0 - \widehat{Z}_0\| + \\ & + C \sum_{j=1}^{n-1} \big(h_{j+k-2} + \gamma^{n-j}\big) E_j + C \sum_{j=1}^{n} h_{j+k-2}^{p} \big\|y^{(p+1)}\big\|_{1,[t_{j-1},t_{j+k-1}]} + \\ & + \varepsilon\, C \sum_{j=1}^{n} h_{j+k-2}^{p} \big(h_{j+k-2} + \gamma^{n-j}\big) \big\|z^{(p+1)}\big\|_{\infty,[t_{j-1},t_{j+k-1}]}. \end{aligned}$$

Hence, the desired convergence estimate for $\|u_n - \hat{u}_n\| \le E_{n-k+1}$ follows from a Gronwall lemma, see proof of Theorem 3.3. □

Acknowledgement.

I thank Gabriela Schranz-Kirlinger for her invitation and many fruitful discussions. Also, I am grateful to Alexander Ostermann for valuable suggestions and comments relating to the present work.

REFERENCES

1. R. Agarwal, *Difference Equations and Inequalities. Theory, Methods and Applications*, Dekker, New York, 1992.
2. H. Brunnner and P. J. van der Houwen, *The numerical solution of Volterra equations*, CWI Monographs 3, North-Holland, Amsterdam, 1986.
3. C. W. Gear and K. W. Tu, *The effect of variable mesh size on the stability of multistep methods*, SIAM J. Numer. Anal. 11 (1974), pp. 1025–1043.
4. E. Hairer, S. P. Nørsett and G. Wanner, *Solving Ordinary Differential Equations I. Nonstiff Problems*, 2nd rev. edn, Springer, Berlin, 1993.
5. E. Hairer and G. Wanner, *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems*, 2nd rev. edn, Springer, Berlin, 1996.
6. Ch. Lubich, *On the convergence of multistep methods for nonlinear stiff differential equations*, Numer. Math. 58 (1991), pp. 839–853.
7. W. L. Miranker, *Numerical Methods for Stiff Equations and Singular Perturbation Problems*, Mathematics and Its Applications 5, Reidel, Dordrecht, 1981.
8. R. E. O'Malley, Jr., *Singular Perturbation Methods for Ordinary Differential Equations*, Applied Mathematical Sciences 89, Springer, Berlin, 1991.

9. A. Ostermann, M. Thalhammer, and G. Kirlinger, *Stability of linear multistep methods and applications to nonlinear parabolic problems*, Appl. Numer. Math. 48 (2004), pp. 389–407.
10. C. Palencia and B. García-Archilla, *Stability of linear multistep methods for sectorial operators in Banach spaces*, Appl. Numer. Math. 445 (1993), pp. 503–520.
11. L. R. Petzold, *A description of DASSL: A differential/algebraic system solver*, in Scientific Computing, R. S. Stepleman et al., eds., North-Holland, Amsterdam, 1983, pp. 65–68.
12. B. van der Pol, *A theory of the amplitude of free and forced triode vibrations*, Radio Review 1 (1920), pp. 701–710/754–762.
13. B. van der Pol, *On relaxation oscillations I*, Phil. Mag. 2 (1926), pp. 978–992.