SIAM J. NUMER ANAL.
Vol. 29, No. 6, pp. 1694-1715, December 1992

010

DIFFERENTIAL/ALGEBRAIC EQUATIONS AS STIFF ORDINARY DIFFERENTIAL EQUATIONS*

MICHAEL KNORRENSCHILD†

Abstract. This paper deals with the relation between differential/algebraic equations (DAEs) and certain stiff ODEs and their respective discretizations by implicit Runge–Kutta methods. For that purpose for any DAE a singular perturbed ODE is constructed such that the DAE is its reduced problem and the solution of the ODE converges in some sense to that of the DAE. Thus the DAE can be interpreted as an infinitely stiff ODE. An analysis of the discretization error of this singular perturbed system gives insight into the relationship of order-reduction phenomena observed for stiff ODEs to that for DAEs. Analysis of a general class of singularly perturbed problems and their discretizations is not attempted; however, the technique of treating singularly perturbed problems and DAEs in a unified way is new and can possibly be applied to other systems and their discretizations as well. Since asymptotic expansions are not used, but an approach similar to the ones used in B-convergence theory is applied, one can derive error bounds that are uniform in the perturbation parameter as well as in the stepsize and do not suffer from restrictions on the ratio of these parameters. This enables one to relate the order of convergence achieved for DAEs to the order of B-convergence. This phenomenon is discussed for several classes of Runge–Kutta methods and illustrated with a numerical example.

Key words. differential/algebraic equations, singular perturbations, implicit Runge–Kutta methods, stiff ordinary differential equations, B-convergence

AMS(MOS) subject classifications. 65L05, 34A99

1. Introduction. In the past years considerable progress has been made in the development of numerical methods for differential/algebraic equations (DAEs) of the form

$$y'(t) = f(t, y(t), z(t)), \qquad f: \mathbb{R}^{m+k+1} \to \mathbb{R}^m, \tag{1.1a}$$

$$0 = g(t, y(t), z(t)), \qquad g: \mathbb{R}^{m+k+1} \to \mathbb{R}^k. \tag{1.1b}$$

One area of interest where DAEs arise is the study of singular perturbed ordinary differential equations

$$\begin{aligned} y'(t) &= f(t, y(t), z(t), \varepsilon), \\ \varepsilon z'(t) &= g(t, y(t), z(t), \varepsilon), \end{aligned} \tag{1.2}$$

where $\varepsilon > 0$ is a small parameter. Then the so-called "reduced" system is obtained by setting $\varepsilon = 0$. Under certain conditions the solution of the reduced system is a good approximation to the solution of (1.2). However, in practice not all DAEs derive from singular perturbations. In many applications (1.1b) arises as a condition that describes geometrical constraints or conservation of energy.

However, (1.1) can obviously be formally embedded in a singular perturbed problem like (1.2) defining $f(.,.,.,0) = f(.,.,.)$, $g(.,.,.,0) = g(.,.,.)$. Usually this simple approach does not give a satisfactory ODE for analytical purposes since the true

*Received by the editors May 21, 1990; accepted for publication January 23, 1992. Major parts of this work are taken from the author's Ph.D. dissertation written under supervision of Prof. R. Jeltsch at the Rheinisch-Westfälische Technische Hochschule Aachen, Germany. Portions of this work were performed at the Institute for Scientific Computing Research under the auspices of the U.S. Department of Energy at Lawrence Livermore National Laboratory under contract W-7405-ENG-48.
†Seminar für Angewandte Mathematik, ETH - Zentrum, 8092 Zürich, Switzerland.

solution of this particular ODE in general does not approach the one of (1.1). Nevertheless, this model provides the motivation for the following widely used trick to derive numerical methods for DAEs from numerical methods for ordinary differential equations.

Apply a method for ordinary differential equations to (1.2), put $\varepsilon = 0$ in the resulting recursion formula (if this is feasible), and obtain a numerical method for (1.1). E.g., in the case of the implicit Euler method that would yield

$$y_{n+1} = y_n + h \cdot f(t_{n+1}, y_{n+1}, z_{n+1})$$

$$0 = g(t_{n+1}, y_{n+1}, z_{n+1}),$$

where $h > 0$ is the stepsize and $t_{n+1} = t_n + h$, $y_n \approx y(t_n)$, $z_n \approx z(t_n)$ $(y(t), z(t)$ is the exact solution of (1.1) at time t). In a similar way a large class of implicit Runge–Kutta methods can be applied as well as multistep methods like the BDF methods.

As these methods are based on formulas suitable for stiff equations, it is natural to compare the corresponding error bounds for DAEs and for stiff ODEs. Recently some progress in this respect has been made [Pe], [BP], [HLR], [Ro], but in all cases the analysis has been based on (1.1) directly (with no reference to (1.2)).

Here we employ a different technique: we will use the discretization error of the stiff system to derive the one for the DAE. This means we want to get an error bound that allows the transition $\varepsilon \to 0$ and would therefore yield an error bound for the DAE. Furthermore, the transition $h \to 0$ independently of ε should be possible in order to allow a bound for the discretization error for the singular perturbed problem at the same time.

Some important observations have been made already in [BP]. Here we will put the considerations given there on a rigorous foundation.

In [BP] relations between the order conditions of Runge–Kutta methods applied to DAEs and to certain classes of stiff ODEs are pointed out. Basically, the order conditions for the stiff system "converge" to those of the DAE as the stiffness tends to infinity. In order to translate this property into a property of the global discretization error of stiff ODEs and DAEs we need that the true solution of the stiff system converges to the one of the DAE in a certain sense. As mentioned above, this is not naturally given, but in §2 we show that it can be achieved by the use of a special stiff system.

In §3 we look at discretizations of those systems. The discretization error for singularly perturbed systems (1.2) has been analyzed already in [HLR] under the assumption $\varepsilon <$ const $\cdot\, h$ using asymptotic expansions. We do not want to use this technique here since the restriction on the ratio ε/h does not allow the transition $h \to 0$ with fixed ε. For the special perturbed system considered here it turns out that the use of the tools developed in B-convergence theory (see [DV]) allows bounds for the global error without restricting the ratio ε/h. We will derive bounds for the DAE as well as the perturbed system in a unified proof by carefully considering the discretization errors. The error bounds for the DAE itself that our technique yields are not new and in most cases are also not optimal. Our emphasis here lies in the presentation of the new technique as opposed to the classical technique where one aims directly at the DAE. Our goal is neither the treatment of a general class of singularly perturbed problems nor the derivation of new error bounds for discretized DAEs. However, by way of our technique we get insight into the relationship between stiff ODEs and DAEs. Furthermore, it is obvious that the use of this technique is not restricted to the systems and methods considered in this paper. It could prove useful when applied to different problems and their discretizations and at the same time reveal insight into properties of other problem classes. For the class of systems considered in this paper we clarify the relation between the order of B-convergence (which is an order of convergence for special classes of stiff systems) and the order of convergence for DAEs. Furthermore, we indicate how this approach can be used to improve the stability of certain Runge–Kutta methods when applied to DAEs.

2. The regularization process. We restrict ourselves to the study of the autonomous case and to the so-called *index*-1-*case*.

DEFINITION 2.1. The DAE

$$y'(t) = f(y(t), z(t)), \qquad f \in C^{(}\mathbb{R}^{m+k}, \mathbb{R}^{m}), \tag{2.1a}$$

$$0 = g(y(t), z(t)), \qquad g \in C^{1}(\mathbb{R}^{m+k}, \mathbb{R}^{k}) \tag{2.1b}$$

is said to have *index* 1 if the Jacobian $g_z(y, z)$ is regular on $\mathbb{R}^{m+k}$.

In essence, the index-1-property guarantees local solvability of the DAE: if an initial value (y_0, z_0) with $g(y_0, z_0) = 0$ is given, then there is a mapping G defined in a neighborhood of y_0 such that

$$g(y, z) = 0 \iff z = G(y)$$

in a neighborhood of (y_0, z_0) and $z_0 = G(y_0)$.

Using this we can rewrite the initial value problem for the DAE as

$$y'(t) = f(y(t), G(y(t))), \qquad y(0) = y_0,$$

$$z(t) = G(y(t)), \qquad z(0) = z_0 = G(y_0),$$

and then we have local solvability. More precisely, we will use the following property of an index-1-DAE:

(A1) There is $S_1 \subseteq \mathbb{R}^m$, $S_2 \subseteq \mathbb{R}^k$, S_1, S_2 open, $\bar{S}_1$, $\bar{S}_2$ compact, and $G \in C^1(S_1, S_2)$ is bijective with

$$g(y, z) = 0 \iff z = G(y) \quad \text{for all } y \in S_1,\ z \in S_2.$$

While looking for a singular perturbed system that reduces for "$\varepsilon = 0$" to a given DAE, we should take care that the solutions of the singular perturbed problem converge in some sense towards the solution of the DAE as $\varepsilon \to 0$. The convergence property we can expect is described by the following definition.

DEFINITION 2.2. Let $F : \mathbb{R}^{2n+2} \to \mathbb{R}^n$ and consider the fully implicit problem

$$F(t, x(t), x'(t), \varepsilon) = 0. \tag{$*$}$$

(i) Let ε be fixed. $x_0 \in \mathbb{R}^n$ is a *consistent initial value* for $(*)$ if $(*)$ has a solution $x \in C^1([a, b], \mathbb{R}^n)$ for some $b > a$ with $x(a) = x_0$.

Now let $\bar{x}(t)$, $t \in [a, b]$ be a solution of $(*)$ for $\varepsilon = 0$; $x_a := \bar{x}(a)$, U a subspace of $\mathbb{R}^n$.

(ii) $(*)$ is said to be ε-*stable on* $[a, b]$ if there is an $\varepsilon_0 > 0$ such that $(*)$ has for each $0 < \varepsilon \leq \varepsilon_0$ and for each consistent initial value $x_{a\varepsilon} \in \mathbb{R}^n$ with $\|x_{a\varepsilon} - x_a\| < \varepsilon_0$ a solution $x_\varepsilon(t)$ such that $x_\varepsilon(a) = x_{a\varepsilon}$ and

$$\lim x_\varepsilon(t) = \bar{x}(t)$$

uniformly on $[a,b]$ for

$$\varepsilon + \|x_{a\varepsilon} - x_a\| \to 0.$$

(iii) $(*)$ is said to be *ε-stable on* $[a,b]$ *for* U if $(*)$ is ε-stable and there is $\delta > 0, \varepsilon_0 > 0$, such that $(*)$ has for each $0 < \varepsilon \le \varepsilon_0$ and for each $u_a \in U$, $\|u_a\| < \delta$, $x_a + u_a$ consistent for $(*)$, a solution $x_\varepsilon(t)$ with $x_\varepsilon(a) = x_a + u_a$ and

$$\lim_{\varepsilon \to 0+} x_\varepsilon(t) = \bar{x}(t) \quad \text{uniformly on } [c,d] \text{ where } a < c < d \le b.$$

In other words, small perturbations out of the subspace U added to the initial value do not destroy the uniform convergence on compact subintervals of $(a,b]$.

In general we will use the following criterion to guarantee ε-stability. For a set M we denote $B(M,\delta) := \{x \mid \|x - y\| < \delta \text{ for some } y \in M\}$.

THEOREM 2.3 [LL]. . *Assume that $(y_0(t), z_0(t))$, $t \in [a,b]$ is solution of the singular perturbed ordinary differential equation*

$$(*) \qquad \begin{aligned} y'(t) &= f(t, y(t), z(t), \varepsilon), \qquad & y(t) \in \mathbb{R}^m, \\ \varepsilon z'(t) &= g(t, y(t), z(t), \varepsilon), \qquad & z(t) \in \mathbb{R}^{n-m} \end{aligned}$$

for $\varepsilon = 0$ with the initial value $y_0(a) = y_a, z_0(a) = z_a$.

Further assume that for some $\delta_0 > 0$ and

$$R := [a,b] \times B(y_0([a,b]), \delta_0) \times B(z_0([a,b]), \delta_0) \times [0, \delta_0),$$

f, f_y, f_z, g, g_y, g_z are continuous on R; for all $t \in [a,b]$ we assume

$$\lambda \text{ eigenvalue of } g_z(t, y_0(t), z_0(t), 0) \Rightarrow \mathrm{Re}(\lambda) < 0.$$

Then $(*)$ is ε-stable on $[a,b]$ for $U := \{\mathcal{O}_m\} \times \mathbb{R}^{n-m}$.

Our goal is to give for any index-1 DAE an ε-stable singular perturbed problem $F(t, x, x', \varepsilon) = 0$, i.e., we want to regularize our system. Regularizations for linear DAEs have already been studied in [Ca], but the analysis uses some requirement on the spectrum of the involved matrices and in general does not appear to be practical. Here we will approach the regularization problem in a natural way, which we are able to do under very general assumptions.

Since we want to use Theorem 2.3, which is formulated for explicit systems, we must take care that $F(t, x, x', \varepsilon) = 0$ is solvable for the derivative x'. Systems that allow this will be called *index-0-systems*. At first, we present the basic regularization, which will be generalized later.

THEOREM 2.4. *Let the* DAE (2.1) *have index* 1, *let $f, g \in C^1$, and let* (A1) *be satisfied. We assume that for a $L_G > 0$*

$$\|G(y) - G(\tilde{y})\| \le L_G \|y - \tilde{y}\| \quad \textit{for all} \;\; y, \tilde{y} \in D_1.$$

Let $y_0 \in D_1$, $z_0 := G(y_0)$, $b > 0$, $\bar{y} : [0,b] \to D_1$, $\bar{y}(0) = y_0$, $\bar{z}(t) := G(\bar{y}(t))$ for $t \in [0,b]$ with

$$\bar{y}'(t) = f(\bar{y}(t), \bar{z}(t)) \quad \textit{for all} \;\; t \in [0,b],$$

let $V \in \mathrm{C}^1(\mathbb{R}^k, \mathbb{R}^k)$, V bijective, and, for all $z \in \mathbb{R}^k$,

$$\lambda \textit{ eigenvalue of } V'(z) \Rightarrow \mathrm{Re}(\lambda) > 0.$$

Then

$$
(2.2) \qquad \begin{aligned} y'(t) &= f(y(t), V^{-1}(V(z(t)) + \varepsilon z'(t))), \\ 0 &= g(y(t), V^{-1}(V(z(t)) + \varepsilon z'(t))) \end{aligned}
$$

has for all $\varepsilon > 0$ *index 0 and there is a* $C > 0$ *such that for all* $(\tilde{y}_0, \tilde{z}_0) \in D_1 \times \mathbb{R}^k$ *there is a* $\tilde{b} \in (0, b]$ *and a solution* $(y_\varepsilon(t), z_\varepsilon(t))$ *of* (2.2) *on* $[0, \tilde{b}]$ *with*

$$
\|y_\varepsilon(t) - \bar{y}(t)\| \leq C\|\tilde{y}_0 - y_0\| \quad \text{for all } t \in [0, \tilde{b}].
$$

Furthermore, for each $a \in (0, \tilde{b}]$ *there is a* $\omega_a : \mathbb{R} \times \mathbb{R}^3_{\geq 0} \to \mathbb{R}_{\geq 0}$ *with*

$$
\lim \omega_a(t, \varepsilon, d1, d2) = 0 \quad \text{for } \varepsilon + d1 + d2 \to 0 \text{ uniformly on } t \in [0, \tilde{b}],
$$

$$
\lim \omega_a(t, \varepsilon, d1, 0) = 0 \quad \text{for } \varepsilon \to 0 \text{ uniformly on } t \in [a, \tilde{b}],
$$

and

$$
\|z_\varepsilon(t) - \bar{z}(t)\| \leq \omega_a(t, \varepsilon, \|\tilde{z}_0 - G(\tilde{y}_0)\|, \|\tilde{y}_0 - y_0\|) + L_G C\|\tilde{y}_0 - y_0\|
$$

for all $t \in (0, \tilde{b}]$.

In particular, (2.2) *is* ε*-stable on* $[0, \tilde{b}]$ *for* $\{\mathcal{O}_m\} \times \mathbb{R}^{n-m}$.

Proof. The index-0-property is easily seen: put $\tilde{g}(y, z, z', \varepsilon) :=$ $g(y, V^{-1}(V(z) + \varepsilon z'))$. It is sufficient to prove the regularity of $\tilde{g}_{z'}$. Here we have

$$
\tilde{g}_{z'}(y, z, z', \varepsilon) = \varepsilon g_z(y, V^{-1}(V(z) + \varepsilon z'))(V^{-1})'(V(z) + \varepsilon z')
$$

and observing $(V^{-1})'(V(z) + \varepsilon z') = V'(V^{-1}(V(z) + \varepsilon z'))$ and the regularity of V' (which is guaranteed by assumption), we are done.

Let $(\tilde{y}_0, \tilde{z}_0) \in S_1 \times \mathbb{R}^k$. From the classical theory we know that there is a $\tilde{b} \in (0, b]$ and $\tilde{y}(t)$ with $\tilde{y}(0) = \tilde{y}_0$ and

$$
\tilde{y}'(t) = f(\tilde{y}(t), G(\tilde{y}(t))) \quad \text{for all } t \in [0, \tilde{b}].
$$

We now define $y_\varepsilon(t) := \tilde{y}(t)$ on $[0, \tilde{b}]$ and z_ε as the solution of the initial value problem

$$
(2.3) \qquad \begin{aligned} V^{-1}(V(z_\varepsilon(t)) + \varepsilon z_\varepsilon'(t)) &= G(\tilde{y}(t)), && z_\varepsilon(0) = \tilde{z}_0, && t \in [0, \tilde{b}], \\ \iff V(z_\varepsilon(t)) + \varepsilon z_\varepsilon'(t) &= V(G(\tilde{y}(t))), && z_\varepsilon(0) = \tilde{z}_0, && t \in [0, \tilde{b}]. \end{aligned}
$$

Obviously, y_ε and z_ε satisfy (2.2). It remains to prove the error estimate.

It is well known that the solution of $y' = f(y, G(y))$ depends Lipschitz-boundedly on the initial value. Therefore, there is a $c > 0$ with

$$
\|\tilde{y}(t) - \bar{y}(t)\| \leq C\|\tilde{y}_0 - y_0\|, \qquad t \in [0, \tilde{b}],
$$

which proves the first part already. This implies immediately

$$
(2.4) \qquad \|G(\tilde{y}(t)) - G(\bar{y}(t))\| \leq L_G C\|\tilde{y}_0 - y_0\| \quad \text{uniformly on } [0, \tilde{b}].
$$

If we invoke Theorem 2.3 on (2.3) (the assumptions of which are satisfied since the eigenvalues of $-V'(z)$ have a negative real part), we arrive for $a \in (0, \tilde{b}]$ at

$$
(2.5) \qquad \|z_\varepsilon(t) - G(\tilde{y}(t))\| = \omega_a(t, \varepsilon, \|\tilde{z}_0 - G(\tilde{y}_0)\|, \|\tilde{y}_0 - y_0\|)
$$

with a certain ω_a that has the desired properties. From (2.4) and (2.5) the claim follows immediately. □

Now we want to have a brief look at multiparameter perturbations. The simplest situation is when all perturbation parameters involved have a constant ratio. But we are interested in criteria that do not involve these ratios (since from a practical point of view knowledge about the ratios may not be available). The essential assumption here is the diagonal dominance.

DEFINITION 2.5. We call a $n \times n$-matrix (a_{ij}) strictly diagonally dominant if

$$a_{ii} > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \quad \text{for } i = 1, \ldots, n.$$

THEOREM 2.6 (Multiparameter version I of Theorem 2.4). Let the assumptions of Theorem 2.4 be satisfied; assume that for all $z \in \mathbb{R}^k$ the matrix $V'(z)$ is strictly diagonally dominant; for $i = 1, \ldots, k$ put $\varepsilon_i := d_i \cdot \varepsilon$, where $d_i > 0$; $D := \mathrm{diag}(d_1, \ldots, d_k)$. Then

$$(*) \qquad \begin{aligned} y'(t) &= f(y(t), V^{-1}(V(z(t)) + \varepsilon D z'(t))), \\ 0 &= g(y(t), V^{-1}(V(z(t)) + \varepsilon D z'(t))) \end{aligned}$$

has for all $\varepsilon > 0$ index 0 and the same claim as in Theorem 2.4 holds.

Proof. Put

$$\tilde{V}(z) := D^{-1} V(z) \quad \text{for } z \in \mathbb{R}^n$$

and verify the assumptions of 2.4.

From the strict diagonal dominance and the Gershgorin theorem we have for all $z \in \mathbb{R}^k$: λ eigenvalue of $V'(z) \Rightarrow \mathrm{Re}(\lambda) > 0$. The same holds for the matrix $D^{-1}V'(z) = \tilde{V}'(z)$ since it is strictly diagonally dominant too. Now we are done, since we can apply Theorem 2.4 with $\tilde{V}$ and we get the desired result. □

Under more restrictive assumptions we have the following result which guarantees convergence for $\varepsilon_1, \ldots, \varepsilon_k \to 0$ without requiring that the ratios of the ε_i be constant.

THEOREM 2.7 (Multiparameter version II of Theorem 2.4). *Let the assumptions of Theorem* 2.6 *be satisfied; in addition, assume* $V(z) = (V_1(z_1), \ldots, V_k(z_k))^T$ *for all* $z \in \mathbb{R}^k$ *and* $V_i'(z_i) > 0$ *for* $i = 1, \ldots, k$ *(the latter follows also from the strict diagonally dominance). Then*

$$(*) \qquad \begin{aligned} y'(t) &= f(y(t), V^{-1}(V(z(t)) + \mathcal{E} z'(t))), \\ 0 &= g(y(t), V^{-1}(V(z(t)) + \mathcal{E} z'(t))), \end{aligned}$$

where $\mathcal{E} = \mathrm{diag}(\varepsilon_1, \ldots, \varepsilon_k), \varepsilon_i > 0$ *has index-0 and the claim in Theorem* 2.6 *holds, if one replaces* ε *by* $\mathcal{E}$ *and*

$$\omega_a(t, \varepsilon, \|\tilde{z}_0 - G(\tilde{y}_0)\|, \|\tilde{y}_0 - y_0\|)$$

by

$$\omega_a(t, \|\mathcal{E}\|, \|\tilde{z}_0 - G(\tilde{y}_0)\|, \|\tilde{y}_0 - y_0\|).$$

Proof. The proof is analogous to Theorem 2.4; but now $z_{\mathcal{E}}$ is not defined by (2.2), but by

$$
\begin{aligned}
&V^{-1}(V(z_{\mathcal{E}}(t)) + \mathcal{E} z'_{\mathcal{E}}(t)) = G(\tilde{y}(t)), && z_{\mathcal{E}}(0) = \tilde{z}_0, && t \in [0,b],\\
\iff\ & V(z_{\mathcal{E}}(t)) + \mathcal{E} z'_{\mathcal{E}}(t) = V(G(\tilde{y}(t))), && z_{\mathcal{E}}(0) = \tilde{z}_0, && t \in [0,b],\\
\iff\ & V_i(z_{\mathcal{E}i}(t)) + \mathcal{E} z'_{\mathcal{E}i}(t) = V_i(G(\tilde{y}(t))), && z_{\mathcal{E}i}(0) = \tilde{z}_{0i}, && t \in [0,b]\\
& && && \text{for } i = 1,\dots,k.
\end{aligned}
$$

But these equations are decoupled. Therefore, we can proceed as in the proof of Theorem 2.6, by dealing with the scalar equations rather than the complete system. □

Remark. Theorems 2.4–2.7 may also be formulated, with obvious modifications, for nonautonomous DAEs.

The presented introduction of perturbation parameters is by no means some kind of artificial technique as it may seem at a first glance. Actually, these small parameters can be given a physical meaning: DAEs arise, for example, in electrical network analysis. If we consider a large class of networks, namely, that composed solely of capacitors, resistors, and current sources, then the vector of unknowns is given by the voltages in the capacitor branches (the y-component) and the voltages in the resistor branches (the z-component). Using Kirchhoff's laws we arrive at a DAE. But this model is somewhat idealized. It does not take into account so-called parasitic effects (such as lead inductance, stray capacitance, etc.). To refine the model one can introduce parasitic elements (by replacing each resistor by a resistor in parallel with a small capacity "ε").

This procedure leads exactly to our regularized system: Theorem 2.4 describes the case when all resistors have nonlinear, strictly monotone increasing characteristics and all parasitic capacitances have the same order. However, to be more exact one should allow parasitic elements of different magnitudes (i.e., introducing $\varepsilon_i,\ i = 1,\dots$), and this is exactly the situation considered in Theorems 2.6 and 2.7 (and the situation in the networks described always leads to a function V that satisfies the requirements in Theorem 2.7). Thus it is not exaggerated to claim that virtually all DAEs arising from circuit analysis are in fact reduced problems of certain singular perturbed ODEs. Our approach here gives a way to recover the singular perturbed ODE behind the DAE along physically meaningful lines. We also would like to emphasize here that the technique of introducing parasitic elements in an electrical network (and therefore regularizing the underlying DAE) is a commonly applied technique among electrical engineers when the circuit analysis software in use encounters problems during the run. A detailed discussion on regularization aspects for DAEs arising in electrical network analysis can be found in [Kn].

3. Implicit Runge–Kutta methods on regularized DAEs. In this section we study the behavior of implicit Runge–Kutta schemes when applied to our perturbed DAEs. Similar work for BDF methods has been done by Lötstedt [Lö1], [Lö2]. We derive for a large class of methods applied to (2.2) with linear V bounds for the global discretization error that are uniform in h and ε. The results extend easily to the situation considered in Theorems 2.6 and 2.7 (i.e., the multiparameter case).

In the following we denote the Kronecker product of square matrices by $A \otimes B$ (see, e.g., [DV]). Runge–Kutta methods are given by (s, A, b) where s is the number of stages, A is the $(s \times s)$-coefficient matrix and b is the vector of weights. We define $c_i := \sum_{j=1}^{s} a_{ij}$. Application of that method to an implicit differential equation $F(t,x,x') = 0$ then gives the following. Solve

$$F\left(t_n + c_i h, x_n + h\sum_{j=1}^{s} a_{ij}X_j', X_i'\right) = 0, \qquad i = 1,\ldots,s$$

and compute

$$x_{n+1} = x_n + h\sum_{i=1}^{s} b_i X_i'.$$

This gives approximations x_n to the true solution $\bar{x}(t_n)$ where $t_n = t_0 + nh$. Equivalently, if A is regular we can proceed as follows ($W := (w_{ij}) = A^{-1}$). Solve

$$F\left(t_n + c_i h, X_i, \frac{1}{h}\sum_{j=1}^{s} w_{ij}(X_j - x_n)\right) = 0, \qquad i = 1,\ldots,s$$

and compute

$$x_{n+1} = x_n + \sum_{i=1}^{s} b_i \sum_{j=1}^{s} w_{ij}(X_j - x_n).$$

Furthermore, we will use the "simplifying conditions": For $p \in \mathbb{N}$ a Runge–Kutta method (s, A, b) satisfies

$$B(p) :\Longleftrightarrow \sum_{i=1}^{s} b_i c_i^{k-1} = \frac{1}{k} \quad \text{for } k = 1,\ldots,p,$$
$$C(p) :\Longleftrightarrow \sum_{j=1}^{s} a_{ij} c_j^{k-1} = \frac{1}{k}c_i^k \quad \text{for } i = 1,\ldots,s; \quad k = 1,\ldots,p.$$

Later we will need the coefficients

$$d_i := \frac{c_i^{q+1}}{(q+1)!} - \sum_{j=1}^{s} a_{ij}\frac{c_j^q}{q!}, \qquad d := (d_1,\ldots,d_s)^T.$$

We denote the property that the local error is of order p by $A(p)$, i.e., let x_{n+1} be the approximation that the method computes within one step starting with $x_n := \bar{x}(t_n)$. Then $A(p)$ means

$$\|x_{n+1} - \bar{x}(t_{n+1})\| \le Ch^{p+1} \quad \text{for all } h \in [0, h_0]$$

for a certain $h_0 > 0$, $C > 0$.

We know (see, e.g., [DV]) that for all $\bar{x} \in C^{p+1}([t_0, t_0+h])$, property $C(p)$ implies

$$\bar{x}(t_0 + c_i h) = \bar{x}(t_0) + h\sum_{j=1}^{s} a_{ij}\bar{x}'(t_0 + c_j h) + O(h^{p+1}), \qquad i = 1,\ldots,s$$

and $B(p)$ implies

$$\bar{x}(t_0 + h) = \bar{x}(t_0) + h\sum_{i=1}^{s} b_i\bar{x}'(t_0 + c_i h) + O(h^{p+1}).$$

While applying our method to the perturbed system we expect the numerical solution $y_{\nu\varepsilon} \approx y_\varepsilon(t_\nu)$, $z_{\nu\varepsilon} \approx z_\varepsilon(t_\nu)$ (notation as in Theorem 2.4) to be dependent on two small parameters, namely, the perturbation parameter $\varepsilon > 0$ and the stepsize $h > 0$. If ε is considered to be fixed we can apply the classical theory and get bounds like

$$\left\| \begin{pmatrix} y_{\nu\varepsilon} \\ \\ z_{\nu\varepsilon} \end{pmatrix} - \begin{pmatrix} y_\varepsilon(t_\nu) \\ \\ z_\varepsilon(t_\nu) \end{pmatrix} \right\| \le Ch^p.$$

Since we want to have approximations for the true solution of the DAE $(\bar{y}(t), \bar{z}(t))^T$ we can derive from this (by using the triangle inequality)

$$\left\| \begin{pmatrix} y_{\nu\varepsilon} \\ z_{\nu\varepsilon} \end{pmatrix} - \begin{pmatrix} \bar{y}(t_\nu) \\ \bar{z}(t_\nu) \end{pmatrix} \right\| \le Ch^p + \left\| \begin{pmatrix} y_\varepsilon(t_\nu) \\ z_\varepsilon(t_\nu) \end{pmatrix} - \begin{pmatrix} \bar{y}(t_\nu) \\ \bar{z}(t_\nu) \end{pmatrix} \right\| = Ch^p + o(\varepsilon).$$

But the constant C always involves higher derivatives of the true solution of the perturbed system $(y_\varepsilon(t), z_\varepsilon(t))^T$ and these are in general unbounded in ε (since essentially z_ε' is of the form $z_\varepsilon'(t) = \varepsilon^{-1}(\bar{z}(t) - z_\varepsilon(t))$).

If under these circumstances we want to keep the discretization error Ch^p small we have to choose h *very* small compared to ε. Obviously, this is not desirable in real computation. But since we know that $(y_\varepsilon(t), z_\varepsilon(t))^T$ is close to $(\bar{y}(t), \bar{z}(t))^T$ we expect that appropriate methods will yield numerical solutions $(y_{\nu\varepsilon}, z_{\nu\varepsilon})^T$ quite close to $(\bar{y}(t), \bar{z}(t))^T$ even if ε is small compared to h. Results of this type have been derived already in [HLR] under the assumption $\varepsilon < \text{const} \cdot h$. Our goal here is to derive estimates like

$$\left\| \begin{pmatrix} y_{\nu\varepsilon} \\ z_{\nu\varepsilon} \end{pmatrix} - \begin{pmatrix} \bar{y}(t_\nu) \\ \bar{z}(t_\nu) \end{pmatrix} \right\| \le C_1\varepsilon + C_2h^p,$$

without restricting the ratio ε/h. This way we would get

$$\left\| \begin{pmatrix} y_{\nu\varepsilon} \\ z_{\nu\varepsilon} \end{pmatrix} - \begin{pmatrix} y_\varepsilon(t_\nu) \\ z_\varepsilon(t_\nu) \end{pmatrix} \right\| \le C_1\varepsilon + C_2h^p + \left\| \begin{pmatrix} y_\varepsilon(t_\nu) \\ z_\varepsilon(t_\nu) \end{pmatrix} - \begin{pmatrix} \bar{y}(t_\nu) \\ \bar{z}(t_\nu) \end{pmatrix} \right\| = C_1\varepsilon + C_2h^p + o(\varepsilon)$$

uniformly in $\varepsilon \in [0, \varepsilon_0]$ and $h \in [0, h_0]$ where ε_0 and h_0 depend only on the method and on the underlying DAE. The above estimate would on the one hand allow $\varepsilon \to 0$ (keeping h fixed) and on the other hand $h \to 0$ (keeping ε fixed) thus relating the orders of convergence of the unperturbed DAE to the one of the stiff (regularized) system.

The appropriate tool to achieve our goal of keeping the perturbation parameter ε independent of the discretization parameter h and vice versa is the B-*convergence theory*. There one deals with systems of ordinary differential equations $y' = f(y)$ where f satisfies $(f(y) - f(z), y - z) \le \mu\|y - z\|^2$ (a so-called one-sided Lipschitz condition), where $(.,.)$ is an inner product. Observe that μ may be negative. The goal of the B-convergence approach is to derive global error bounds Ch^p where C is not allowed to depend on the usual Lipschitz constant but only on the one-sided μ. This is important for stiff systems where it is characteristic that the usual Lipschitz constant is very large, whereas the one-sided μ may be of moderate size. In our perturbed problem we deal basically with systems $\varepsilon z' = -z + \ldots$ and that means that the usual Lipschitz constant is essentially $1/\varepsilon$, whereas the one-sided Lipschitz constant could be taken as $-1/\varepsilon$ or also $\mu = 0$. Setting $\mu = 0$ represents a uniform one-sided Lipschitz constant suitable for every $\varepsilon > 0$ and promises to lead us in the direction of uniform error bounds.

Since the requirement not to make use of the classical Lipschitz constant imposes restrictions on the kind of error bounds obtained, it is not surprising that the order of B-convergence of a method may be lower than its classical order (it may happen that a method is not B-convergent at all, but convergent in the classical sense). In many cases the order of B-convergence reflects the practical behavior of a method applied to a stiff system better than the classical order. For a detailed discussion on these aspects see the excellent book by Dekker and Verwer [DV].

Unfortunately, we cannot apply the results of the B-convergence theory immediately to our perturbed DAEs. Although by using these results we would get rid of the (classical) Lipschitz constant we would still have the higher derivatives of the analytical solution involved in the error bound. Also we can show that in general our perturbed system does not satisfy a one-sided Lipschitz condition with μ independent of ε. But it turns out that we can use very similar techniques if at first we use the true solution of the DAE (which, of course, does not depend on ε) in the estimates. So our technique is as follows.

We consider the discretized regularized DAE and derive bounds for the difference between the numerical solution of the regularized system and the exact solution of the DAE. The tools of the B-convergence theory enable us to handle all bounds involved uniformly for $\varepsilon \in [0, \varepsilon_0]$ and $h \in [0, h_0]$ without relating ε to h (or vice versa). Particularly, we do not need to restrict the ratio ε/h. Since we are working with the higher derivatives of the solution of the DAE (which are independent of ε), we do not have to worry about singularities in $\varepsilon = 0$. The ε-dependent solution of the perturbed system is introduced only in the final stage (Corollary 3.3) via triangle inequality, as mentioned earlier.

We will denote in what follows:

$$e := (1, \dots, 1)^T \in \mathbb{R}^s, \qquad v \otimes I_k := \begin{pmatrix} v_1 \cdot I_k \\ \vdots \\ v_s \cdot I_k \end{pmatrix} \in \mathbb{R}^{ks \times k},$$

$$v^T \otimes I_k := (v_1 \cdot I_k, \dots, v_s \cdot I_k) \in \mathbb{R}^{k \times ks} \quad \text{for } v \in \mathbb{R}^s, \qquad E := e \otimes I_k.$$

Before we present our main theorem we will describe some technical assumptions that are needed. Assuming (A1) is satisfied we will make use of the following.

(A2) The underlying DAE (2.1) has a solution $(\bar{y}, \bar{z})^T$ on an interval $[t_0, t]$ for an initial value (y_0, z_0) such that $(\bar{y}(t), \bar{z}(t))^T \in S_1 \times S_2$. (Note that this implies $g(y_0, z_0) = 0$.) $\bar{y}$, $\bar{z}$, f, g are sufficiently smooth.

(A3) The numerical approximations computed by the underlying method and all stage values arising in this process remain in $S_1 \times S_2$ for the initial value $(y_0, \tilde{z}_0)$, where $\tilde{z}_0$ is sufficiently close to z_0.

For (A2) we refer to the remark after Definition 2.1; (A3) is always satisfied when h is sufficiently small.

The main result of this paper is the following theorem.

THEOREM 3.1. *Let (s, A, b) be an A-stable Runge–Kutta method such that A has eigenvalues with positive real part only and $C(q)$, $B(p)$, and $A(p)$, where $p \geq q + 1$, are satisfied. Let* (A1), (A2), *and* (A3) *be satisfied for the index*-1 DAE (2.1) *and the method, respectively, and consider the method applied to the regularized system*

$$y' = f(y, M^{-1}(Mz + \varepsilon z')),$$
$$0 = g(y, M^{-1}(Mz + \varepsilon z')),$$

where $M \in \mathbb{R}^{k\times k}$ is regular with $(Mx, x) \geq 0$ for all $x \in \mathbb{R}^k$ for some inner product $(.,.)$ on $\mathbb{R}^k$.

Let $\varepsilon_0 > 0$, and let $\|.\|$ be the norm corresponding to the inner product above; denote $e_0 := \tilde{z}_0 - \bar{z}(t_0)$. Then there is $h_0 > 0$ that depends only on ε_0, the method, and the DAE *above, such that*

$$e_{y\nu} := y_\nu - \bar{y}(t) = O(h^p),$$
$$e_\nu := z_\nu - \bar{z}(t) = O(\|e_0\|) + O(\varepsilon) + O(h^q).$$

Furthermore, if $\psi(z) := [b^T(I_s - Az)^{-1}e]^{-1}b^T(I_s - Az)^{-1}d$ is uniformly bounded on $\mathbb{C}^-$ then we have

$$e_\nu = O(\|e_0\|) + O(\varepsilon) + O(h^{q+1}).$$

The constants involved in the $O(.)$ terms depend only on the method, the DAE, *and the constants ε_0 and h_0.*

Proof. Without restriction we assume that the h_0 mentioned in the assumption is small enough to meet all the requirements that will be made in the following proof.

We want to compute one step with our method applied to (2.1) and starting at $(y_\mu, z_\mu) \in S_1 \times S_2$, i.e., the $\varepsilon = 0$ case. The stage values Y_i, Z_i, $i = 1, \ldots, s$ in that case are given as the solution of

$$\begin{aligned} Y_i &= y_\mu + h\sum_{j=1}^{s} a_{ij} f(Y_j, Z_j), \qquad i = 1, \ldots, s, \\ 0 &= g(Y_i, Z_i). \end{aligned} \tag{3.1}$$

We are just interested in solutions $Y_i \in S_1$, $Z_i \in S_2$ $(i = 1, \ldots, s)$, so this is equivalent with

$$\begin{aligned} Y_i &= y_\mu + h\sum_{j=1}^{s} a_{ij} f(Y_j, G(Y_j)), \qquad i = 1, \ldots, s, \\ Z_i &= G(Y_i). \end{aligned} \tag{3.2}$$

We want to apply the implicit function theorem to this system in order to derive the existence of solutions $Y_i(h)$ and therefore also $Z_i(h)$ that are continuous in h.

The nonlinear system of equations to solve is

$$F(h, \vec{Y}) = 0, \qquad \vec{Y} := \begin{pmatrix} Y_1 \\ \vdots \\ Y_s \end{pmatrix},$$

where

$$F(h,\vec{Y}) := (F_1(h,\vec{Y}),\dots,F_s(h,\vec{Y}))^T,$$
$$F_i(h,\vec{Y}) := Y_i - y_\mu - h\sum_{j=1}^{s} a_{ij} f(Y_j, G(Y_j)) \quad \text{for } i=1,\dots,s.$$

We have

$$F(0,(y_0,\dots,y_0)^T) = 0, \qquad F_{\vec{Y}}(0,\vec{Y}) = I_{ms} \quad \text{for all } \vec{Y}\in \mathbb{R}^{ms},$$

which allows us to solve for $Y(h)$ as desired. In particular if h_0 is small enough there exists $\vec{Y}\in C([0,h_0],\mathbb{R}^{ms})$ with

$$F(h,\vec{Y}(h)) = 0 \quad \text{on } [0,h_0].$$

Since

$$y_{\mu+1} = y_\mu + h\sum_{i=1}^{s} b_i f(Y_i(h), G(Y_i(h))),$$

it is clear that for h small enough $y_{\mu+1}$ will be in S_1. Also it is obvious that we have done nothing more than apply our method to the ODE $y' = f(y,G(y))$. Using the classical techniques the assertion can be derived for the y-component.

It remains to prove the assertion for the z-component. Define $M_\varepsilon := \frac{1}{\varepsilon}M$. One step with our method for the regularized system requires stage values $Y_{i\varepsilon}, Z_{i\varepsilon}$ satisfying

$$Y_{i\varepsilon} = y_\mu + h\sum_{j=1}^{s} a_{ij} f(Y_{j\varepsilon}, Z_{j\varepsilon} + M_\varepsilon^{-1} Z'_{j\varepsilon}), \qquad i=1,\dots,s,$$
$$0 = g(Y_{i\varepsilon}, Z_{i\varepsilon} + M_\varepsilon^{-1} Z'_{i\varepsilon}),$$

where as usual $Z'_{i\varepsilon}$ is implicitly defined by

$$Z_{i\varepsilon} = z_\mu + h\sum_{j=1}^{s} a_{ij} Z'_{j\varepsilon}, \qquad i=1,\dots,s. \tag{3.3}$$

Therefore, it is obvious that

$$Y_{i\varepsilon} := Y_i, \qquad i=1,\dots,s$$

and $Z_{i\varepsilon}$ satisfies

$$\begin{aligned} Z_{i\varepsilon} + M_\varepsilon^{-1} Z'_{i\varepsilon} &= Z_i, & i&=1,\dots,s\\ \Longleftrightarrow \qquad Z'_{i\varepsilon} &= M_\varepsilon(Z_i - Z_{i\varepsilon}), & i&=1,\dots,s. \end{aligned} \tag{3.4}$$

If we denote

$$\vec{Z}_\varepsilon := \begin{pmatrix} Z_{1\varepsilon}\\ \vdots \\ Z_{s\varepsilon}\end{pmatrix} \in \mathbb{R}^{ks}, \qquad \vec{Z}_0 := \begin{pmatrix} Z_1\\ \vdots \\ Z_s\end{pmatrix} \in \mathbb{R}^{ks},$$

(3.3) and (3.4) lead us to

$$\vec{Z}_\varepsilon = Ez_\mu + h(A \otimes I_k)(I_s \otimes M_\varepsilon)(\vec{Z}_0 - \vec{Z}_\varepsilon). \tag{3.5}$$

Using Lemma 2.4.6 in [Hu] we see that (3.5) has a solution $\vec{Z}_\varepsilon$ for every $h \geq 0$, $\varepsilon \geq 0$, since

$$(I_{ks} + h(A \otimes M_\varepsilon))^{-1} = (I + hAz)|_{z=M_\varepsilon}^{-1} = (I + hzA)^{-1}|_{z=M_\varepsilon}.$$

The approximation $z_{\mu+1}$ is given by

$$z_{\mu+1} = z_\mu + h(b^T \otimes I_k)(I_s \otimes M_\varepsilon)(\vec{Z}_0 - \vec{Z}_\varepsilon). \tag{3.6}$$

If we denote in addition

$$\bar{Z}_\mu := \begin{pmatrix} \bar{z}(t_\mu + c_1 h) \\ \vdots \\ \bar{z}(t_\mu + c_s h) \end{pmatrix}, \qquad \bar{Z}'_\mu := \begin{pmatrix} \bar{z}'(t_\mu + c_1 h) \\ \vdots \\ \bar{z}'(t_\mu + c_s h) \end{pmatrix},$$

we have (using Taylor expansion)

$$\begin{aligned} \bar{Z}_\mu &= E\bar{z}(t_\mu) + h(A \otimes I_k)\bar{Z}'_\mu + r_\mu, \\ \bar{z}(t_\mu + h) &= \bar{z}(t_\mu) + h(b^T \otimes I_k)\bar{Z}'_\mu + O(h^{p+1}), \end{aligned} \tag{3.7}$$

where

$$r_\mu = (d \otimes I_k)\bar{z}^{(q+1)}(t_\mu)h^{q+1} + O(h^{q+2}).$$

For the errors $e_{\mu+1} := z_{\mu+1} - \bar{z}(t_\mu + h)$ and $e_\mu := z_\mu - \bar{z}(t_\mu)$ we derive, using (3.6) and (3.7),

$$e_{\mu+1} = e_\mu + h(b^T \otimes I_k)((I_s \otimes M_\varepsilon)(\vec{Z}_0 - \vec{Z}_\varepsilon) - \bar{Z}'_\mu) + O(h^{p+1}). \tag{3.8}$$

Furthermore, (3.5) together with (3.7) gives

$$-\vec{Z}_\varepsilon - \bar{Z}_\mu = Ee_\mu + h(A \otimes I_k)((I_s \otimes M_\varepsilon)(\vec{Z}_0 - \vec{Z}_\varepsilon) - \bar{Z}'_\mu) - r_\mu. \tag{3.9}$$

To make use of that in (3.8) we need a relation between $\bar{Z}_\mu$ and $\vec{Z}_0$. To get this we return temporarily to the events in the y-component.

Remembering the beginning of the proof and defining $\bar{f} := f(.,G(.))$, and

$$\begin{aligned} D_i &:= h \int_0^1 \overline{f}'(Y_i + \theta(\bar{y}(t_\mu + c_i h) - Y_i))\, d\theta \qquad (i = 1, \ldots, s), \\ D &:= \operatorname{diag}(D_1, \ldots, D_s) \in \mathbb{R}^{ks \times ks} \end{aligned}$$

we get (as in [BHV]), with $\bar{Y}_\mu$ is defined analogous to $\bar{Z}_\mu$,

$$\bar{Y}_\mu - \vec{Y} = (I_{ks} - (A \otimes I_k)D)^{-1}(E(y_\mu - \bar{y}(t_\mu)) + R_\mu), \tag{3.10}$$

where

$$R_\mu = (d \otimes I_k)\bar{y}^{(q+1)}(t_\mu)h^{q+1} + O(h^{q+2}).$$

Observe that $I_{ks} - (A \otimes I_k)D$ is regular for h small enough (since $D = O(h)$). Also $(I_{ks} - (A \otimes I_k)D)^{-1} = I_{ks} + O(h)$ and therefore (since $y_\mu - \bar{y}(t_\mu) = O(h^p)$ with $p \geq q+1$) we have $\bar{Y}_\mu - \vec{Y} = O(h^{q+1})$.

Using the mean value theorem for G and the s components in (3.10) we get

$$\begin{aligned}
\bar{Z}_\mu - \vec{Z}_0 &= \begin{pmatrix} G'(\bar{y}(t_\mu + c_1 h)) & & \\ & \ddots & \\ & & G'(\bar{y}(t_\mu + c_s h)) \end{pmatrix} (\bar{Y}_\mu - \vec{Y}) + O(h^{2q+2}) \\
&= (I_s \otimes G'(\bar{y}(t_\mu))(\bar{Y}_\mu - \vec{Y}) + O(h^{q+2}) \\
&= (I_s \otimes G'(\bar{y}(t_\mu))(EO(h^p) + (d \otimes I_k)O(h^{q+1})) + O(h^{q+2}). \\
&= EG'(\bar{y}(t_\mu))O(h^p) + (d \otimes I_k)G'(\bar{y}(t_\mu))O(h^{q+1}) + O(h^{q+2}) \\
&= EO(h^p) + (d \otimes I_k)O(h^{q+1}) + O(h^{q+2}) =: \tilde{R}_\mu.
\end{aligned}$$

Now insert this into (3.9) and we have

$$\vec{Z}_\varepsilon - \vec{Z}_0 = Ee_\mu + h(A \otimes I_k)((I_s \otimes M_\varepsilon)(\vec{Z}_0 - \vec{Z}_\varepsilon) - \bar{Z}'_\mu) + \tilde{R}_\mu - r_\mu.$$

But this means nothing other than

$$\begin{aligned}
&(I_{ks} + h(A \otimes M_\varepsilon))(\vec{Z}_\varepsilon - \vec{Z}_0) = Ee_\mu - h(A \otimes I_k)\bar{Z}'_\mu + \tilde{R}_\mu - r_\mu \\
\iff &(I_s \otimes M_\varepsilon)(\vec{Z}_0 - \vec{Z}_\varepsilon) - \bar{Z}'_\mu = (I_s \otimes M_\varepsilon)(I_{ks} + h(A \otimes M_\varepsilon))^{-1}[h(A \otimes I_k)\bar{Z}'_\mu \\
&\qquad - Ee_\mu - \tilde{R}_\mu + r_\mu - (I_{ks} + h(A \otimes M_\varepsilon))(I_s \otimes M_\varepsilon)^{-1}\bar{Z}'_\mu] \\
&\qquad = (I_s \otimes M_\varepsilon)(I_{ks} + h(A \otimes M_\varepsilon))^{-1} \\
&\qquad \cdot \left[-Ee_\mu - (I_s \otimes M_\varepsilon)^{-1}\bar{Z}'_\mu - \tilde{R}_\mu + r_\mu\right].
\end{aligned}$$

Inserting into (3.8) leads us to

$$\begin{aligned}
e_{\mu+1} = e_\mu &+ h(b^T \otimes I_k)(I_s \otimes M_\varepsilon)(I_{ks} + h(A \otimes M_\varepsilon))^{-1} \cdot \\
&\cdot [-Ee_\mu - (I_s \otimes M_\varepsilon)^{-1}\bar{Z}'_\mu + r_\mu - \tilde{R}_\mu] + O(h^{p+1}).
\end{aligned} \tag{3.11}$$

If we denote

$$N(h, \varepsilon) := h(b^T \otimes I_k)(I_s \otimes M_\varepsilon)(I_{ks} + h(A \otimes M_\varepsilon))^{-1} \in \mathbb{R}^{k \times ks},$$

we finally arrive at

$$e_{\mu+1} = [I_k - N(h,\varepsilon)E]e_\mu - N(h,\varepsilon)[(I_s \otimes M_\varepsilon)^{-1}\bar{Z}'_\mu + r_\mu - \tilde{R}_\mu] + O(h^{p+1}). \tag{3.12}$$

Since we have

$$\bar{Z}'_\mu = E\bar{z}'(t_\mu) + O(h)$$

it is clear that

$$(I_s \otimes M_\varepsilon)^{-1}\bar{Z}'_\mu = (I_s \otimes M_\varepsilon^{-1})[E\bar{z}'(t_\mu) + O(h)] = E \cdot M_\varepsilon^{-1}\bar{z}'(t_\mu) + \varepsilon(I_s \otimes M^{-1})O(h)$$
$$= \varepsilon E \cdot M^{-1}\bar{z}'(t_\mu) + O(\varepsilon h).$$

Now we can write (3.12) as

$$e_{\mu+1} = [I_k - N(h,\varepsilon)E]e_\mu - N(h,\varepsilon)[\varepsilon EO(1) + O(\varepsilon h) + r_\mu - \tilde{R}_\mu] + O(h^{p+1}).$$

Using an induction argument yields

$$e_\nu = (I_k - N(h,\varepsilon)E)^\nu e_0$$
$$+ \sum_{j=0}^{\nu-1}(I_k - N(h,\varepsilon)E)^j[-N(h,\varepsilon)EO(\varepsilon) - N(h,\varepsilon)(O(\varepsilon h) + r_j - \tilde{R}_j) + O(h^{p+1})]$$
$$= (I_k - N(h,\varepsilon)E)^\nu e_0 + ((I_k - N(h,\varepsilon)E)^\nu - I_k)O(\varepsilon)$$
$$+ \sum_{j=0}^{\nu-1}(I_k - N(h,\varepsilon)E)^j[-N(h,\varepsilon)(O(\varepsilon h) + r_j - \tilde{R}_j) + O(h^{p+1})].$$

Furthermore, we have

$$N(h,\varepsilon) = h(b^T \otimes I_k)(I_s \otimes M_\varepsilon)(I_{ks} + h(A \otimes M_\varepsilon))^{-1}$$
$$= \left[hb^T z(I + hzA)^{-1}\right]_{z=M_\varepsilon}.$$

If we denote the *stability function* of our method by

$$R(z) := 1 + zb^T(I_s - zA)^{-1}e,$$

we then have

$$I_k - N(h,\varepsilon)E = R(-hM_\varepsilon).$$

If we further write the vector of polynomials $p_1(z) := -zb^T(I - zA)^{-1}$ (then $N(h,\varepsilon) = p_1(-hM_\varepsilon)$), we finally arrive at

$$e_\nu = R(-hM_\varepsilon)^\nu e_0 + (R(-hM_\varepsilon)^\nu - 1)O(\varepsilon)$$
$$+ \sum_{j=0}^{\nu-1} R(-hM_\varepsilon)^j \left[p_1(-hM_\varepsilon)(O(\varepsilon h) + r_j - \tilde{R}_j) + O(h^{p+1})\right]. \tag{3.13}$$

Because of the A-stability we have $|R(-hM_\varepsilon)| \leq 1$ for all h, $\varepsilon > 0$. This, and the uniform boundedness of p_1 (according to Lemma 4.2 in [BHV], using the assumption on the eigenvalues of A) and $\nu h = \text{const.}$ gives

$$e_\nu = R(-hM_\varepsilon)^\nu e_0 + O(\varepsilon) + \sum_{j=0}^{\nu-1} R(-hM_\varepsilon)^j p_1(-hM_\varepsilon)(r_j - \tilde{R}_j) + O(h^p).$$

But

$$\begin{aligned}
&\sum_{j=0}^{\nu-1} R(-hM_\varepsilon)^j p_1(-hM_\varepsilon)(r_j - \tilde{R}_j) \\
&\quad = (1 - R(-hM_\varepsilon)^\nu)(1 - R(-hM_\varepsilon))^{-1} p_1(-hM_\varepsilon)[EO(h^p) + O(h^{q+2})] \\
&\qquad + \sum_{j=0}^{\nu-1} R(-hM_\varepsilon)^j p_1(-hM_\varepsilon)(d \otimes I_k) O(h^{q+1}) \\
&\quad = (1 - R(-hM_\varepsilon)^\nu) O(h^p) \\
&\qquad + \sum_{j=0}^{\nu-1} R(-hM_\varepsilon)^j p_1(-hM_\varepsilon)(d \otimes I_k) O(h^{q+1}) + O(h^{q+1})
\end{aligned}$$

and from this the first part of the assertion follows immediately. To prove the second part one uses that

$$\sum_{j=0}^{\nu-1} R(-hM_\varepsilon)^j p_1(-hM_\varepsilon)(d \otimes I_k) = (1 - R(-hM_\varepsilon)^\nu)\psi(-hM_\varepsilon). \qquad \square$$

Having done that work already we can easily derive a multiparameter version.

THEOREM 3.2. *Let the assumptions of Theorem* 3.1 *be satisfied, but replace this time in the* DAE ε *by* $\mathcal{E}$ *where* $\mathcal{E} := \mathrm{diag}(\varepsilon_1, \ldots, \varepsilon_k)$ *with* $\varepsilon_i > 0$ *for* $i = 1, \ldots, k$; M *is assumed to be a diagonal matrix and* $(.,.)$ *denotes the Euclidean inner product. Then*

$$\begin{aligned}
e_{y\nu} &:= y_\nu - \bar{y}(t) = O(h^p), \\
e_\nu &:= z_\nu - \bar{z}(t) = O(\|e_0\|) + O(\|\mathcal{E}\|) + O(h^q).
\end{aligned}$$

Furthermore, if $\psi(z) := [b^T(I_s - Az)^{-1}e]^{-1} b^T (I_s - Az)^{-1} d$ *is uniformly bounded on* $\mathbb{C}^-$, *then we have*

$$e_\nu = O(\|e_0\|) + O(\|\mathcal{E}\|) + O(h^{q+1}).$$

Proof. Denote $M_\varepsilon := \mathcal{E}^{-1} \cdot M = M \cdot \mathcal{E}^{-1}$ and observe that for all $x \in \mathbb{R}^k$

$$(M_\varepsilon x, x) = \sum_{i=1}^{k} \varepsilon_i^{-1} m_{ii} x_i^2 \geq 0,$$

since we have for all $y \in \mathbb{R}^k$ $0 \leq (My, y) = \sum_{i=1}^k m_{ii} y_i^2$. Hence $m_{ii} \geq 0$ for $i = 1, \ldots, k$. Using this M_ε we can simply use the proof for Theorem 3.1; note that

$$\begin{aligned}
(I_s \otimes (\mathcal{E}^{-1} \cdot M))^{-1} \bar{Z}'_\mu &= E \cdot M^{-1} \cdot \mathcal{E} \bar{z}'(t_\mu) + (I_s \otimes M^{-1})(I_s \otimes \mathcal{E}) O(h) \\
&= EO(\|\mathcal{E}\|) + O(\|\mathcal{E}\| h).
\end{aligned}$$

The remainder is obvious. $\square$

COROLLARY 3.3. *Let the assumptions of Theorem* 3.1, *respectively, Theorem* 3.2, *be satisfied. Then*

$$e_{y\nu} := y_\nu - y_\varepsilon(t) = O(h^p),$$
$$e_\nu := z_\nu - z_\varepsilon(t) = O(\|e_0\|) + o(\varepsilon) + O(h^q).$$

Furthermore, if $\psi(z) := [b^T(I_s - Az)^{-1}e]^{-1}b^T(I_s - Az)^{-1}d$ *is uniformly bounded on* $\mathbb{C}^-$, *then we have*

$$e_\nu = O(\|e_0\|) + o(\varepsilon) + O(h^{q+1}),$$

or, respectively, the corresponding assertion for $\mathcal{E}$.

Remarks. (1) *Discussion of some methods.* The question of which methods ψ is uniform bounded on $\mathbb{C}^-$ has already been discussed in some detail in [BHV]. For example, this is the case for the Radau-IA, Radau-IIA, and Lobatto-IIIC methods. Remember that for the s-stage methods we have (see e.g., [DV]):

for the Gauss methods: $A(2s),\ B(2s),\ C(s)$,

for the Radau-IA methods: $A(2s-1),\ B(2s-1),\ C(s-1)$,

for the Radau-IIA methods: $A(2s-1),\ B(2s-1),\ C(s)$,

for the Lobatto-IIIC methods: $A(2s-2),\ B(2s-2),\ C(s-1)$.

In the y-component the order of convergence is the same as for ODEs, i.e., l, if we have $A(l)$. Now we are able to state the final results.

The following orders of convergence for s-stage Runge–Kutta methods applied to (2.2) with a linear V hold uniformly in parameters $(\varepsilon, h) \in [0, \varepsilon_0] \times [0, h_0]$:

for the Gauss methods: s if $s \geq 2$, 2 if $s = 1$,

for the Radau-IA methods: s,

for the Radau-IIA methods: $s+1$ if $s \geq 2$, 1 if $s = 1$,

for the Lobatto-IIIC methods: s.

The same orders of convergence are stated in [BHV] as orders of B-convergence for the class of stiff ODEs

$$U'(t) = QU(t) + g(t, U(t)),$$

where

$$(Qu, u) \leq \beta\|u\|^2 \quad \forall u \in \mathbb{R}^n,$$
$$\|g(t,u) - g(t,\tilde{u})\| \leq \alpha\|u - \tilde{u}\| \quad \forall u, \tilde{u} \in \mathbb{R}^n, \quad t \in \mathbb{R}.$$

This is no surprise since to a great extent our proof followed the lines of the proof in [BHV], and consequently the assumptions in their theorem are exactly identical with ours. Let us point out again that it is not possible to use their theorem directly since our perturbed system does not fit in the above-mentioned form of a stiff ODE.

The same orders of convergence have been derived already for fully implicit index-1 DAEs in [BP]. Our results above prove the same for (unperturbed) semi-explicit index-1 DAEs since the results remain valid for $\varepsilon = 0$. However, in the semi-explicit case these orders are not always optimal as becomes evident in [Ro]. For the Radau-IA methods they are indeed optimal, but it is known that the Gauss methods, for example, yield order $s+1$ if s is odd (and order s if s is even). We do not intend to investigate this case here. Furthermore, note that if the numerical approximations satisfy the constraint equation (as is the case for the Radau-IIA and Lobatto-IIIC methods), then because of the index-1 property the order in the z-component is the same as in the y-component. For more details see [Ro].

(2) *Comparison of Theorem* 3.1 *with a previous result from Griepentrog* [Gri]. He considers

$$(*)\qquad \begin{aligned} y'(t) &= f(t, y(t), z(t), \varepsilon), \qquad && y(t) \in \mathbb{R}^m, \\ z'(t) &= g(t, y(t), z(t), \varepsilon) + \frac{1}{\varepsilon} M z, \qquad && z(t) \in \mathbb{R}^k, \end{aligned}$$

where the eigenvalues of the constant matrix M have a negative real part. His result states uniform (in ε) order of convergence for this system.

In contrast to our regularization, the one in $(*)$ is linear in z'. On the other hand, the idea behind our approach is an even simpler and also linear regularization ("replace z by $z + \varepsilon M^{-1} z'$"), so that in principle we solve a very simple linear but *discrete* differential equation. The continuous dependency of the right-hand side in $(*)$ presents no essential difference since our results could be easily modified to include this case also.

Griepentrog proves uniform convergence for all L-stable Runge–Kutta methods L-stable means A-stable and

$$\lim_{z \to -\infty} R(z) = 0$$

holds, where R is the stability function of the method. But there is no discussion on the rate of convergence. For our regularization we were able to give an error representation and order conditions. Furthermore, our assumptions are less restrictive since there are Runge–Kutta methods that have a coefficient matrix with eigenvalue with positive real part but are not L-stable, e.g., the Gauss methods. Another example is the following [Pe], [HW]:

Define a semi-explicit method by

$$A := \begin{pmatrix} \frac{3+\sqrt{3}}{6} & 0 \\ \frac{-\sqrt{3}}{3} & \frac{3+\sqrt{3}}{6} \end{pmatrix}, \qquad b := \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}.$$

Obviously the real parts of the eigenvalues of A are positive.

But this method is not L-stable since

$$\lim_{z \to -\infty} R(z) = \frac{1+\sqrt{3}}{2+\sqrt{3}}.$$

So our theorem is applicable, whereas Griepentrog's is not. This method satisfies $B(4)$, $C(1)$, and $A(3)$, and it can be shown that the corresponding ψ is uniformly bounded on $\mathbb{C}^-$. Therefore, Theorem 3.1 shows that the uniform order of convergence is 3 in the y-component and 2 in the z-component.

(3) *Comparison of Theorem* 3.1 *and Corollary* 3.3 *with a previous result from Hairer, Lubich, and Roche* [HLR]: In [HLR, Thm. 1] the following result is proved: Let the assumptions of our Theorem 3.3 be satisfied. Furthermore, let $(Mx, x) \geq \|x\|^2$ for all $x \in \mathbb{R}^k$ and assume that consistent initial values are used. Then for $\varepsilon \leq$ const$\cdot h$ the following conditions hold. If $b_i = a_{si}$, $i = 1, \ldots, s$ then

$$y_\nu - y_\varepsilon(t) = O(h^p) + O(\varepsilon h^{q+1}),\ z_\nu - z_\varepsilon(t) = O(h^p) + O(\varepsilon h^q),$$

and if

$$\lim_{z \to -\infty} |R(z)| < 1 \text{ then } y_\nu - y_\varepsilon(t) = O(h^p) + O(\varepsilon h^{q+1}), \qquad z_\nu - z_\varepsilon(t) = O(h^{q+1}).$$

In this result the estimate for the y-component is ε-dependent since in [HLR] a more general system (namely, $y' = f(y, z)$, $\varepsilon z' = g(y, z)$) is considered. Therefore, let us focus our attention on the z-components.

As in Corollary 3.3 this result gives order q or $q + 1$ depending on properties of the method. Our assumption in Corollary 3.3, that ψ is uniformly bounded on $\mathbb{C}^-$, appears to be more restrictive than $R(-\infty) < 1$, but this difference is of no importance if one considers the classical Runge–Kutta methods (see Remark (1)). However, note that the result from [HLR] only holds for $\varepsilon \leq$ const $\cdot\, h$, whereas our result allows arbitrary h/ε and ε/h and is independent on these ratios. This is the crucial point. Because of that restriction the result in [HLR] does not give the order of convergence of a method applied to singular perturbed system in the usual sense—the limiting process $h \to 0$ is not permitted since the condition $\varepsilon <$ const $\cdot\, h$ enforces $\varepsilon \to 0$ at the same time.

Also, we may later want to make use of the regularization idea by solving the regularized system numerically rather than the original DAE. In some situations numerical experiments suggest the use of a rather large ε (e.g., 10^{-2}) and then the restriction $\varepsilon <$ const $\cdot\, h$ would not allow us to choose the stepsize small enough to get a reasonable accuracy.

Another difference in the results is that ours covers methods with $|R(-\infty)| = 1$ as well. For example, the result in [HLR] does not apply to Gauss methods whereas ours does (see Remark (1)). Furthermore, our approach aims at inconsistent initial values, a case that is not included in the analysis in [HLR]. See the following remark for more details.

(4) *Influence of the initial error* e_0. The approach presented is of advantage in the situation where there is an initial error e_0 present. In this case we may use the numerical solution of the regularized system which should model the fast initial transition of the solution of the singular perturbed problem. That way we get close to the solution of the DAE again.

It should be noted, however, that in the case of semi-explicit index-1 DAEs it is not too hard to compute consistent initial values (i.e., those with $e_0 = 0$) directly: if we know the y-components of an initial value the perhaps missing z-components can be found by solving the constraint equation $g(y, z) = 0$ numerically for z. In view of the index-1 assumption this is possible using Newton iteration. We may then use special Runge–Kutta methods that ensure that the numerical approximations satisfy the constraint in each step (for example, Radau-IIA or Lobatto-IIIC methods). Another alternative is to use a general Runge–Kutta method and correct the numerical z-values after each step with a Newton iteration (as mentioned above). That way we

TABLE 1

$\downarrow \varepsilon \setminus h \rightarrow$	0.005	0.01	0.02	order
0	−6.83657937E−01	−6.83657937E−01	−6.83657945E−01	4.00
	7.29802594E−01	7.29802593E−01	7.29802575E−01	4.00
	−1.70524200E+00	−1.70524200E+00	−1.70524198E+00	4.00
	−1.59742133E+00	−1.59742133E+00	−1.59742137E+00	4.00
	1.21893956E+01	1.21893591E+01	1.21892134E+01	2.00
0.0025	6.20132189E+00	6.20132223E+00	6.20132752E+00	3.97
0.005	6.21310034E+00	6.21310052E+00	6.21310327E+00	3.98
0.01	6.23623565E+00	6.23623574E+00	6.23623720E+00	3.99

would get y- and z-components numerically with the same order of convergence as for ordinary differential equations. So from a practical point of view it is not necessary to apply regularization in this situation.

In order to illustrate the flavor of the results obtained in this paper let us consider the following index-1 DAE (pendulum DAE):

$$\begin{aligned} y' &= f(y,z) := \begin{pmatrix} y_3 \\ y_4 \\ -y_1 z \\ -y_2 z + 1 \end{pmatrix}, \\ 0 &= y_3^2 - y_1^2 z + y_4^2 - y_2^2 z + y_2, \end{aligned}$$

where $y = (y_1, y_2, y_3, y_4)^T$. A consistent initial value would be $(y(0), z(0))^T = (1, 0, 0, 2, 4)^T$, in which case the solution at $t = 1$ is determined to be

$$(y(1), z(1))^T = (-0.6836579, 0.729802, -1.70524, -1.597421, 6.189408)^T.$$

In order to demonstrate the effects of the initial error we incorporate quite a large perturbation in the z-component by choosing $(1, 0, 0, 2, 10)^T$ as the numerical initial value. Note that the DAE has no solution to this initial value. Using the two-stage Gauss formula,

$$A = \begin{pmatrix} 0.25 & 0.25 - 0.5/\sqrt{3} \\ 0.25 + 0.5/\sqrt{3} & 0.25 \end{pmatrix}, \qquad b = (0.5, 0.5)^T$$

and constant stepsizes on $[0, 1]$, we obtain at $t = 1$ the results given in Table 1. Here the order (of convergence in h) has been estimated from the numerical results. We did not include the values for the y-components in the cases where $\varepsilon > 0$; since those components are not affected by the regularization they are the same as in the $\varepsilon = 0$ case. Note that in the case $\varepsilon = 0$ the initial error stays unreduced in the z-components all the time. Also we observe an order reduction down to 2; in the (nonstiff) y-components we observe the classical order 4 as expected. If $\varepsilon > 0$ is introduced we observe immediately that we get the correct z-component numerically—the initial error has been damped and is no longer present. Together with it comes an increase in the order up to about the classical order 4 which indicates rapid convergence. Here we expect the classical order rather than the order of B-convergence (which is 2), since the magnitude of ε is not really "small." If ε would be chosen smaller than the values given in Table 1 we would expect the order to decrease slowly down to 2 (which is the order for the $\varepsilon = 0$ case).

This result is in agreement with the analysis presented earlier. From the error representation in Theorem 3.1 we see that e_0 is propagated after ν steps with stepsize h as $R(-hM_\varepsilon)^\nu e_0$. Since we assume A-stability this term is always bounded by e_0. For $h > 0$, $\varepsilon > 0$ we have $R(-hM_\varepsilon) < 1$ so that the influence of e_0 decreases with an increasing number of steps. Keeping in mind that we are interested in an approximation for the $\varepsilon = 0$ case this means that the numerical solution of the perturbed system tends to the true solution of the reduced system within $O(\varepsilon)$ even if we use an incorrect initial value in the z-component. (Note that perturbations in the y-component influence the numerical solution in the same way as they do for ordinary differential equations since in that component we simply solve an ordinary differential equation.) However, in case $\varepsilon = 0$ we use a L-stable method (i.e., $R(-\infty) = 0$), then after one step the error e_0 no longer influences the numerical solution.

If the method satisfies $|R(-\infty)| = 1$ (as is the case for the Gauss methods) then the influence of e_0 on the numerical solution of the reduced system will not decrease; it will remain in the amount of $\|e_0\|$ and therefore the numerical solution will not come close to the true solution but will remain in a tube with diameter $\|e_0\| + O(\varepsilon)$ around it. But if we compute the numerical solution of the perturbed problem the e_0-term will die away (since $R(-hM_\varepsilon) < 1$ even if the method is not L-stable) and therefore the numerical solution will tend to the true solution of the reduced problem within $O(\varepsilon)$. Consequently, in this situation it is of significant advantage to solve the perturbed system rather than the unperturbed one.

Note that in the table above we cannot expect convergence (for $h \to 0$) in the z-component to the true solution of the DAE. Keep in mind that we are solving the singular perturbed problem numerically which has a different solution than the DAE. The difference between the true solutions however disappears as ε decreases (and as t increases). Thus, if we want to improve the approximation to the true solution of the DAE we can of course decrease the magnitude of ε. For example, the choice $\varepsilon = 10^{-5}$ gives (using $h = 0.001$ and the specifications in the table) an approximation at $t = 1$ of 6.189456 that gives the solution of the DAE accurately within 5 digits.

The optimal choice of the magnitude of ε, however, is a topic of its own. Usually one may determine it by numerical experiments. One has to take care that ε is large enough to avoid being too close to the DAE. If ε is too small one would run into the same computational problems that one encounters with DAEs. On the other hand, if ε is too large the numerical solution will not approximate the true solution of the underlying DAE very well.

Finally, we remark that the technique presented here can also be used to tackle DAEs with perturbed initial values numerically where the index-1 property does not hold. This will be the subject of future research.

Acknowledgment. I thank L. Petzold for a careful reading of a first version of this paper.

REFERENCES

[BuHu] K. BURRAGE AND W. H. HUNDSDORFER, *The order of* B*-convergence of algebraically stable Runge–Kutta methods*, BIT, 27 (1987), pp. 62–71.

[BHV] K. BURRAGE, W. H. HUNDSDORFER, AND J. G. VERWER, *A study of* B*-convergence of Runge–Kutta methods*, Computing, 36 (1986), pp. 17–34.

[BP] K. BURRAGE AND L. PETZOLD, *On order reduction for Runge–Kutta Methods applied to differential/algebraic systems and to stiff systems of* ODE*s*, SIAM J. Numer. Anal., 27 (1990), pp. 447–456.

[Ca] S. L. CAMPBELL, *Regularizations for linear time varying singular systems*, Automatica, 20 (1984), pp. 365–370.

[DV] K. DEKKER AND J. G. VERWER, *Stability of Runge–Kutta Methods for Stiff Nonlinear Differential Equations*, North-Holland, Amsterdam, 1984.

[Gri] E. GRIEPENTROG, *Numerische Integration von steifen Differentialgleichungssystemen mit Einschrittverfahren*, Beitr. Numer. Math., 8 (1980), pp. 59–74.

[HLR] E. HAIRER, CH. LUBICH, AND M. ROCHE, *Error of Runge–Kutta methods for stiff problems studied via differential algebraic equations*, BIT, 28 (1988), pp. 678–700.

[Hu] W. H. HUNDSDORFER, *The numerical solution of nonlinear stiff initial value problems—an analysis of one-step methods*, CWI Tract 12, Amsterdam, 1985.

[HW] G. HALL AND J. M. WATT, *Modern Numerical Methods for Ordinary Differential Equations*, Oxford University Press, London, 1976.

[Kn] M. KNORRENSCHILD, *Regularisierungen von Differentiell-Algebraischen Systemen—theoretische und numerische Aspekte*, Ph.D. thesis, Aachen University of Technology, Aachen, Germany, 1988.

[LeLe] J. J. LEVIN AND N. LEVINSON, *Singular perturbations of nonlinear systems of differential equations and an associated boundary layer equation*, J. Rational Mech. Anal., 3 (1954), pp. 247–270.

[Lö1] P. LÖTSTEDT, *Discretizations of singular perturbation problems by* BDF*-methods*, Tech-Report 99, Uppsala University, Uppsala, Sweden, 1985.

[Lö2] ———, *On the relation between singular perturbation problems and differential-algebraic equations*, Tech.-Report 100, Uppsala University, Uppsala, Sweden, 1985.

[Pe] L. R. PETZOLD, *Order results for implicit Runge–Kutta methods applied to differential/algebraic systems*, SIAM J. Numer. Anal., 23 (1986), pp. 837–852.

[Ro] M. ROCHE, *Runge–Kutta and Rosenbrock methods for differential-algebraic equations and stiff* ODE*s*, Ph.D. thesis, Université de Genève, Genève, Switzerland, 1988.