Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Index reduction of index 1 DAE under uncertainty ⋆

H. Tidefelt * S. T. Glad **

* Department of Electrical Engineering, Linköping University, Sweden
(e-mail: tidefelt@isy.liu.se)
** Department of Electrical Engineering, Linköping University, Sweden
(e-mail: torkel@isy.liu.se)

Abstract: This paper examines an index reduction method for linear time-invariant differential algebraic equations, with uncertainty in the equation coefficients. When the bottom block of a block upper triangular leading matrix contains no elements that can be distinguished from zero, the natural action to take is to replace all numbers in the block by exact zeros, and then proceed with index reduction by differentiation. Conditions are given under which zeroing of an uncertain small block gives a small deviation in the solution.

1. INTRODUCTION

This paper is concerned with homogeneous square LTI (linear time-invariant) DAE (differential-algebraic equations), that is, equations in the form

$$E\,x'(t) + A\,x(t) = 0 \tag{1}$$

where E and A are square, possibly singular, matrices. The vector-valued function x is the unknown, which is to be solved for given initial conditions; $x(0) = x^0$. Only square equations will be considered, so *square* is dropped from the notation from here on.

Understanding sensitivity with respect to small parameters in LTI DAE is in general fundamental for the treatment of equations that are not known exactly. Such equations may, for instance, be the result of system identification, and shall not be confused with DAE which can be analyzed using only structural information (that is, the pattern of exact zeros in the matrices defining the equations). The immediate application is, as suggested by the paper title, to gain a better understanding of index reduction. However, since failing to understand index reduction would imply that we cannot make sense of the equations at hand, another prospect application is to come up with good notions of well-definedness of DAE under uncertainty. Still, there are many numerical solvers for DAE which are in use, although they do not consider these issues. As a consequence, the error bounds they deliver (if any) along with the solution do not take the index reduction process into account. It is believed that the kind of sensitivity analysis performed in this paper can fill this gap, thereby proving its practical usefullness besides its otherwise rather esoteric bearings.

It should be stressed from the beginning, that the technical results in this paper can at most be seen as a first step on a long journey — these results must be much generalized before being applicable to, for instance, quasilinear DAE of unknown index. However, the paper as a whole can also be viewed as a novel approach on how to think of the problem, namely that we shall seek assumptions that enable us to make sense of index reduction. While several sets of assumptions may enable this, they will differ in how easily they can be established in applications, and how tight error bounds they may produce in the future. Hence, we also consider this paper as a first step in a development towards more viable assumptions.

Notation: In the calculations to come, an uncertain matrix E will prevail. The set of all possible E shall be determined by context, and will not be part of our notation. For compactness, we shall write dependence on E with a subscript. For instance, writing $y_E(\epsilon)$ means the same as writing $y(\,\epsilon,\,E\,)$ or even $y(\,\epsilon,\,E(\epsilon)\,)$. We also need compact notation for limits that are uniform with respect to E, and those that are not. Writing $y_E(\epsilon) = \mathcal{O}^E(\epsilon)$ means

$$\exists\, k^0,\, \epsilon^* > 0 :\; \epsilon \in [\,0,\, \epsilon^*\,] \;\Rightarrow\; \sup_E |y_E(\epsilon)| \le k^0\,\epsilon$$

while writing $y_E(\epsilon) = \mathcal{O}_E(\epsilon)$ means

$$\forall\, E :\; \exists\, k^0,\, \epsilon^* > 0 :\; \epsilon \in [\,0,\, \epsilon^*\,] \;\Rightarrow\; |y_E(\epsilon)| \le k^0\,\epsilon$$

Think of this notation as follows: E being a subscript on $\mathcal{O}$ means that the constants of the $\mathcal{O}$ are functions of E; we could have written "$\forall\, E :\; \exists\, k^0_E,\, \epsilon^*_E > 0 :\; \ldots$" to emphasize this dependency. Also, E being a *sup*erscript can be used as a reminder of the $\sup_E$ in the definition.

New symbols are sometimes constructed by adding "over-bars" to existing symbols. For instance, this means that $\bar{E}$ does *not* denote some operation performed on E, but is just a mnemonic way of constructing a symbol that should remind of E.

The symbol I denotes the identity matrix. The norm used for vectors is the Euclidean norm. The norm used for matrices is the spectral norm, unless where a subscript "F" is used to indicate the Frobenius norm. Accordingly, the *condition number* of a matrix A refers to $\|A\|\,\left\|A^{-1}\right\|$ based on the spectral norm.

2. BACKGROUND

Performing row operations on (1) typically gives matrices of the following form, where ϵ is a small number:

$$\begin{aligned} K_0\,E &= \begin{pmatrix} E_{11} & E_{12} \\ 0 & \epsilon\,E_{22} \end{pmatrix} \\ K_0\,A &= \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \end{aligned} \tag{2}$$

The influence of the block $\epsilon\,E_{22}$ on system properties is the subject of the present paper.

⋆ With financial support from the Swedish Research Council.

The issue with perturbations in DAE has been considered previously in Mattheij and Wijckmans (1998). While their setup is different to ours, we share many of their observations. However, their way of approaching the issue — even their way of formulating the problem — differs from ours. Consequently, their results are not immediately competing with ours. Their work is referred to in Kunkel and Mehrmann (2006, remark 6.7), as the latter authors remark that a perturbation analysis is still lacking in their influential framework for numerical solution of DAE. Although the current paper deals with perturbation related more to index reduction by shuffling, it is hoped that our work will inspire the development of perturbation analysis in other contexts as well.

The question of how the solution depends on the small parameter ϵ is related to the so-called *singular perturbation theory*, well developed in Kokotović et al. (1986). In that setup, $E_{11} = I$, $E_{12} = 0$, and $E_{22} = I$, so one essentially deals with an ODE with time-scale separation. The important difference to our setting is that they consider E_{22} known and of perfect condition (it is the identity matrix).

3. ANALYSIS

The analysis in this paper is limited to DAE of index at most 1. We first consider equations of index 0, before taking on the slightly more involved systems of index at most 1.

3.1 Preparation

Consider the LTI DAE

$$\bar{E}\,\bar{x}'(t) + \bar{A}\,\bar{x}(t) = 0 \tag{3}$$

with uncertain matrices. We are interested in a situation where the uncertainty can be parameterized as follows

$$0 = \begin{pmatrix} E_{11} & E_{12} \\ 0 & E_{22} \end{pmatrix} \bar{x}'(t) + \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \bar{x}(t) \tag{4}$$

with square blocks on the diagonal. The matrix E_{22} is not assumed to be known except for a small bound on $\|E_{22}\|$. As is rather natural for equations appearing in this for, E_{11} is assumed non-singular. The limiting system as $\|E_{22}\| \to 0$,

$$0 = \begin{pmatrix} E_{11} & E_{12} \\ 0 & 0 \end{pmatrix} \bar{x}'(t) + \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \bar{x}(t) \tag{5}$$

is assumed to be of index 1, that is,

$$\begin{pmatrix} E_{11} & E_{12} \\ A_{21} & A_{22} \end{pmatrix} \tag{6}$$

is non-singular.

The purpose of the paper is to find conditions that gurantee that the solutions of (4) are close to the solutions of (5) if $\|E_2\|$ is sufficiently small.

As a first step in the analysis the system is transformed into a more convenient form.

Lemma 1. The system (4) can be transformed into the form

$$\begin{aligned} x'(t) &= M_{11}(\epsilon)\,x(t) + M_{12}(\epsilon)\,z(t) \\ \epsilon\,E(\epsilon)\,z'(t) &= M_{21}(\epsilon)\,x(t) + M_{22}(\epsilon)\,z(t) \end{aligned} \tag{7}$$

where, in a neighborhood of the origin, all matrices are analytic with known bounds on their derivatives [1], the M_{ij} are known at the origin, $M_{22}(0)$ is non-singular, and $\|E(\epsilon)\| = 1$. The change of variables does not depend on the uncertain E_{22}.

[1] When speaking of bounded derivatives, we always refer to the first order derivatives.

Proof. The new form is obtained using a transformation of variables

$$\bar{x} = \begin{pmatrix} I & -E_{11}^{-1}E_{12} \\ 0 & I \end{pmatrix} \begin{pmatrix} x \\ z \end{pmatrix}$$

In the notation of (7), this yields

$$\begin{aligned} \epsilon\,E(\epsilon) &= E_{22} \\ -M_{11}(\epsilon) &= E_{11}^{-1}A_{11} \\ -M_{12}(\epsilon) &= E_{11}^{-1}A_{12} - E_{11}^{-1}A_{11}\,E_{11}^{-1}E_{12} \\ -M_{21}(\epsilon) &= A_{21} \\ -M_{22}(\epsilon) &= A_{22} - E_{11}^{-1}E_{12}\,A_{21} \end{aligned}$$

Here, ϵ is identified as $\|E_{22}\|$, and the statements regarding the matrix functions M_{ij} are trivial since they are constant (they neither depend explicitly on ϵ, nor do they depend implicitly on ϵ via E_{22}). To see that $M_{22}(0)$ is non-singular, post-multiply the non-singular (6) by the non-singular matrix defining the change of variables:

$$\begin{pmatrix} E_{11} & E_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} I & -E_{11}^{-1}E_{12} \\ 0 & I \end{pmatrix} = \begin{pmatrix} E_{11} & 0 \\ A_{21} & A_{22} - E_{11}^{-1}E_{12}\,A_{21} \end{pmatrix}$$

Non-singularity of this matrix implies that the lower right block, which equals $-M_{22}(0)$, is also non-singular. □

The reason expressions in lemma 1 which are free of ϵ are still considered parameterized by ϵ is to obtain similarity to results appearing later in the paper. Besides, since the form (7) is the starting point for the derivations to follow in the paper, this possible dependency on ϵ adds to the generality of the results.

Remark 1. Note that lemma 1 implies that we can concentrate on the form (7), since $\mathcal{O}^{E}(\,\epsilon\,)$ convergence in $(\,x,\,z\,)$ implies $\mathcal{O}^{E}(\,\epsilon\,)$ convergence in $\bar{x}$ due to the change of variables being known. Also note that a change of variables (or if there would be any row operations) will not alter the eigenvalues, which will always be the system poles.

Notation: From here on, we drop the *in a neighborhood of the origin* from our notation. Further, since element-wise bounds give norm bounds, and vice verse, we also drop the norms from our notation when speaking of bounded matrices. Hence, for example, we simply write that E_{22}' *is bounded.*

3.2 Singular perturbation in ODE

The derivation in this section follows the structure in Kokotović et al. (1986). In their analysis, results come in two flavors; one where approximations are valid on any finite time interval, and one where stability of the slow dynamics in the system make the approximations valid without restriction to finite time intervals. In the present treatment, it is from here on assumed that only finite time intervals are considered, but the other case is treated just as easily.

Lemma 2. There exists an analytic matrix function $L_E(\epsilon)$, such that the change of variables,

$$\begin{pmatrix} x \\ z \end{pmatrix} = \begin{pmatrix} I & 0 \\ L_E(\epsilon) & I \end{pmatrix} \begin{pmatrix} x \\ \eta \end{pmatrix} \tag{8}$$

transforms (7) into the system (dropping ϵ)

$$\begin{pmatrix} I & \\ & \epsilon\, E \end{pmatrix} \begin{pmatrix} x'(t) \\ \eta'(t) \end{pmatrix}$$
$$= \begin{pmatrix} M_{11} + M_{12}\, L_E & M_{12} \\ 0 & M_{22} - \epsilon\, E\, L_E\, M_{12} \end{pmatrix} \begin{pmatrix} x(t) \\ \eta(t) \end{pmatrix} \quad (9)$$

The matrix L_E satisfies

$$L_E(0) = -M_{22}^{-1}(0) M_{21}(0) \quad (10)$$

and its derivative has a known bound.

Proof. Applying the change of variables shows that x is eliminated from the η' equation provided $L_E(\epsilon)$ satisifies

$$0 = M_{21}(\epsilon) + M_{22}(\epsilon)\, L_E(\epsilon)$$
$$- \epsilon\, E(\epsilon)\, L_E(\epsilon) \left(M_{11}(\epsilon) + M_{12}(\epsilon)\, L_E(\epsilon) \right) \quad (11)$$

For $\epsilon = 0$ there is the solution

$$L_E(0) = -M_{22}^{-1}(0) M_{21}(0)$$

The derivative of the right hand side of (11) with respect to L_E at $\epsilon = 0$ is $M_{22}(0)$, which is non-singular. It follows from the analytical implicit function theorem, Hörmander (1966), that the equation can be solved to give an analytical L_E. Differentiating (11) with respect to ϵ and evaluating near $\epsilon = 0$ reveals a known bound on L_E'. □

Let the initial conditions for (7) be

$$x(0) = x^0, \quad z(0) = z^0$$

Lemma 3. If the initial conditions x^0 and z^0 are chosen to make the DAE consistent for $\epsilon = 0$, that is,

$$0 = M_{21}(0)\, x^0 + M_{22}(0)\, z^0 \quad (12)$$

then the initial condition $\eta(0) = \eta^0(\epsilon)$ for the second equation of (9) satisfies

$$\eta^0(\epsilon) = -(M_{22}(0)^{-1} M_{21}(0) + L_E(\epsilon)) x^0 \quad (13)$$

In particular η is an analytic function with

$$\eta_E^0(\epsilon) = \mathcal{O}^E(\epsilon) \quad (14)$$

Proof. From the definition of the variable change $z = L_E(\epsilon)\, x + \eta$ it follows that

$$\eta_E^0(\epsilon) = z^0 - L_E(\epsilon)\, x^0$$

Substituting z^0 from (12) gives (13) while (10) and the known bound on E' gives (14). □

Introduce the notation

$$M(\epsilon) \stackrel{\triangle}{=} M_{22}(\epsilon) - \epsilon\, E(\epsilon)\, L_E(\epsilon)\, M_{12}(\epsilon) \quad (15)$$

To emphasize the difference between uniform and pointwise convergence with respect to the uncertainty, the following lemma gives a pointwise result to be contrasted with lemma 5.

Lemma 4. Assume $E(0)$ is non-singular and that $E(0)^{-1} M(0)$ has all its eigenvalues strictly in the left half plane. Then, for $t \geq 0$,

$$|\eta(t)| = \mathcal{O}_E(\, \epsilon \,) \quad (16)$$

Proof. If $E(0)$ is invertible, then $E(\epsilon)^{-1}$ exists and is analytic. With the variable change $t = \epsilon \tau$ we get

$$\frac{\partial \eta}{\partial \tau} = E(\epsilon)^{-1} M(\epsilon)\, \eta, \quad \eta(0) = \eta^0$$

with the solution

$$\eta(\tau) = \mathrm{e}^{E(\epsilon)^{-1} M(\epsilon)\, \tau} \eta^0$$

Since $E(0)^{-1} M(0)$ has all eigenvalues in the left half plane the norm of $\mathrm{e}^{E(0)^{-1} M(0)\, \tau}$ is bounded. Since all the matrix elements are analytic in ϵ there are, for each E, positive constants C_1 and ϵ_0 such that

$$\left\| \mathrm{e}^{E(\epsilon)^{-1} M(\epsilon) \tau} \right\| \leq C_1, \quad 0 \leq \epsilon < \epsilon_0$$

Since $\eta^0 = \mathcal{O}^E(\epsilon)$ the result follows. □

If $E(0)$ is singular other estimates are possible.

Since the location of the poles of a system conveys interesting information that has engineering interpretation, we give a result in terms of the locations.

Lemma 5. In addition to the assumptions of lemma 3, assume $E(\epsilon)$ is known to be non-singular and that there exist $R_0 > 0$ and $\phi_0 < \pi/2$ such that for λ being a pole of (7),

$$|\lambda| > R_0 \implies |\arg(\, -\lambda\,)| < \phi_0$$

Also assume that the DAE is not close to index 1 in the sense that there exists a bound κ_0 on the condition number of $E(\epsilon)$.

Then, for any fixed $t_1 \geq t_0$, for all $t \in [\, t_0,\, t_1\,]$,

$$|\eta_E(\, t,\, \epsilon\,)| = \mathcal{O}^E(\epsilon)$$

Proof. The isolated system in η has the state-feedback matrix

$$M_\eta(\epsilon) \stackrel{\triangle}{=} \frac{1}{\epsilon}\, E(\epsilon)^{-1} M_{22}(\epsilon) - L_E(\epsilon)\, M_{12}(\epsilon)$$

Recall that $\|E(\epsilon)\| = 1$. Hence the condition number bound gives $\left\| E(\epsilon)^{-1} \right\| \leq \kappa_0$, and hence $\|M_\eta(\epsilon)\| < \frac{\kappa_0 + 1}{\epsilon}\, \|M_{22}(0)\|$ for sufficiently small ϵ. By lemmas 13 (see section A) and 3, it only remains to show that $\epsilon\, \alpha(\, M_\eta(\epsilon)\,)$ can be bounded by a negative constant. By showing that the there exists a constant $k^1 > 0$ such that any eigenvalue λ of $\epsilon\, M_\eta(\epsilon)$ is larger in magnitude than k^1 as $\epsilon \to 0$, it follows that all eigenvalues of $M_\eta(\epsilon)$ approach infinity like $\frac{k_1}{\epsilon}$, as $\epsilon \to 0$. It then follows that they will all satisfy the argument condition for sufficiently small ϵ, and that $\alpha(\, M_\eta(\epsilon)\,) < -\frac{k_1}{\epsilon}\, \cos(\, \phi_0\,)$.

This is shown by using that all eigenvalues of $\epsilon\, M_\eta(\epsilon)$ are greater than $\left\| \left(\epsilon\, M_\eta(\epsilon) \right)^{-1} \right\|^{-1}$, where (dropping ϵ)

$$\begin{aligned} \left\| \left(\epsilon\, M_\eta \right)^{-1} \right\|^{-1} &= \left\| \left(E^{-1} \left(M_{22} - \epsilon\, E\, L_E\, M_{12} \right) \right)^{-1} \right\|^{-1} \\ &\geq \left\| \left(M_{22} - \epsilon\, E\, L_E\, M_{12} \right)^{-1} \right\|^{-1} \\ &\geq \left\| \left(M_{22} - \epsilon\, E\, L_E\, M_{12} \right)^{-1} \right\|_{\mathrm{F}}^{-1} \end{aligned}$$

Here, it is clear that the limit is positive since $M_{22}(0)$ is non-singular, but to ensure that there is an $\epsilon^* > 0$ such that $\left\| \left(\epsilon\, M_\eta(\epsilon) \right)^{-1} \right\|^{-1}$ is greater than some positive constant for all $\epsilon \in [\, 0,\, \epsilon^*\,]$, we must also show that the derivative with respect to ϵ is finitely bounded independently of E. By differentiability of the matrix inverse and Frobenius norm, this follows if the derivative of the inverted matrix is bounded independently of E, which is readily seen. □

Continuing on the result of lemma 2, the following lemma shows that the influence of η on x is small.

Lemma 6. There exists a change of variables,

$$\begin{pmatrix} x \\ \eta \end{pmatrix} = \begin{pmatrix} I & \epsilon\, H_E(\epsilon)\, E(\epsilon) \\ 0 & I \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} \quad (17)$$

such that the implicit ODE (9) can be written (dropping ϵ)

$$\begin{pmatrix} I & \\ & \epsilon\, E \end{pmatrix} \begin{pmatrix} \xi'(t) \\ \eta'(t) \end{pmatrix} = \begin{pmatrix} M_{11} + M_{12}\, L_E & 0 \\ 0 & M_{22} - \epsilon\, E\, L_E\, M_{12} \end{pmatrix} \begin{pmatrix} \xi(t) \\ \eta(t) \end{pmatrix} \quad (18)$$

and for sufficiently small ϵ, $\| H_E(\epsilon) \|$ is bounded by a constant independently of E.

Proof. Applying the change of variables and then performing row operations on the equations to eliminate η' from the first group of equations, lead to the condition defining $H_E(\epsilon)$:

$$0 = \big(M_{11}(\epsilon) + M_{12}(\epsilon)\, L_E(\epsilon) \big) \epsilon\, H_E(\epsilon)\, E(\epsilon) + M_{12}(\epsilon) - H_E(\epsilon) \big(M_{22}(\epsilon) - \epsilon\, E(\epsilon)\, L_E(\epsilon)\, M_{12}(\epsilon) \big) \quad (19)$$

It follows that

$$H_E(0) = M_{12}(0)\, M_{22}(0)^{-1}$$

which is clearly bounded independently of E. The equation is linear in $H_E(\epsilon)$ and the coefficients depend smoothly on ϵ, so the solution is differentiable at $\epsilon = 0$. It thus remains to show that the derivative of $H_E(\epsilon)$ with respect to ϵ at 0 can be bounded independently of E. As with L_E, the bound on the derivative of H'_E can be revealed by differentiating (19) with respect to ϵ. □

Recall remark 1 and consider (7). Let the solution at time t be denoted $x_E(\, t,\, \epsilon\,)$, and let $x(\, t,\, 0\,) \overset{\triangle}{=} x_E(\, t,\, 0\,)$ to emphasize that E does not matter if $\epsilon = 0$. Let $z(\, t,\, 0\,) \overset{\triangle}{=} -M_{22}^{-1}(0)\, M_{21}(0)\, x(\, t,\, 0\,)$.

Theorem 7. Consider the form (7). Assume that the initial conditions are consistent with $\epsilon = 0$, and that there exists a bound κ_0 on the condition number of $E(\epsilon)$. Assume there exist $R_0 > 0$ and $\phi_0 < \pi/2$ such that for λ being a system pole,

$$|\lambda| > R_0 \implies |\arg(\, -\lambda\,)| < \phi_0$$

Then

$$|x_E(\, t,\, \epsilon\,) - x(\, t,\, 0\,)| = \mathcal{O}^E(\epsilon) \quad (20)$$

$$|z_E(\, t,\, \epsilon\,) - z(\, t,\, 0\,)| = \mathcal{O}^E(\epsilon) \quad (21)$$

Proof. Define $L_E(\epsilon)$ and $H_E(\epsilon)$ as above, and consider the solution expressed in the variables ξ and η. Lemma 5 shows how η is bounded uniformly over time and with respect to E. Note that $x(\, t,\, 0\,)$ coincides with $\xi(\, t,\, 0\,)$, so the left hand side of (20) can be bounded as

$$\begin{aligned} &|x_E(\, t,\, \epsilon\,) - x(\, t,\, 0\,)| \\ &= |\xi_E(\, t,\, \epsilon\,) + \epsilon\, H_E(\epsilon)\, E(\epsilon)\, \eta_E(\, t,\, \epsilon\,) - \xi(\, t,\, 0\,)| \\ &\le |\xi_E(\, t,\, \epsilon\,) - \xi(\, t,\, 0\,)| + \mathcal{O}^E(\, \epsilon^2\,) \end{aligned}$$

To see that the first of these terms is $\mathcal{O}^E(\epsilon)$, note first that lemmas 3 and 6 give that the initial conditions for ξ are only $\mathcal{O}^E(\, \epsilon^2\,)$ away from x^0. Hence, the restriction to a finite time interval gives that the contribution from initial conditions is negligible. The contribution from perturbation of the state-feedback matrix for ξ depends on the perturbed matrix in a non-trivial manner, but useful bounds exist.(Van Loan, 1977) Since lemma 6 shows that the size of the perturbation is $\mathcal{O}^E(\epsilon)$, it follows that the contribution is $\mathcal{O}^E(\epsilon)$ at any fixed time t. The constants of the $\mathcal{O}^E(\epsilon)$ bounds will of course be a continuous function of time, and since the time interval of interest is compact, it follows that a dominating constant exists.

Concerning z (recall the definition of $z(\, t,\, 0\,)$),

$$\begin{aligned} &\big| z_E(\, t,\, \epsilon\,) + M_{22}^{-1}(0) M_{21}(0)\, x(\, t,\, 0\,) \big| \\ &\le \big| z_E(\, t,\, \epsilon\,) + M_{22}(0)^{-1} M_{21}(0)\, x_E(\, t,\, \epsilon\,) \big| \\ &\quad + \big| M_{22}(0)^{-1} M_{21}(0)\ (\, x(\, t,\, 0\,) - x_E(\, t,\, \epsilon\,)\,) \big| \\ &\le |z_E(\, t,\, \epsilon\,) + L_E(\epsilon)\, x_E(\, t,\, \epsilon\,)| + \mathcal{O}^E(\epsilon)\ |x_E(\, t,\, \epsilon\,)| \\ &\quad + \big\| M_{22}(0)^{-1} M_{21}(0) \big\|\ \mathcal{O}^E(\epsilon) \\ &= |\eta_E(\, t,\, \epsilon\,)| + \mathcal{O}^E(\epsilon)\ |x_E(\, t,\, \epsilon\,)| \\ &\quad + \big\| M_{22}(0)^{-1} M_{21}(0) \big\|\ \mathcal{O}^E(\epsilon) \\ &= \mathcal{O}^E(\epsilon) \end{aligned}$$

since $|x_E(\, t,\, \epsilon\,)|$ can be bounded over any finite time interval. □

3.3 Singular perturbation in index 1 DAE

With the exceptions of lemmas 2, 3, and 6, the theorems so far require, via lemma 5, that E (or E_{22}) have bounded condition number. However, it is possible to proceed also when some singular values are exactly zero, if assuming that the DAE is not close to index 2. Next, the results of the previous section will be extended to this situation by revisiting the relevant proofs.

Common to the proofs in this section is the observation that there is a non-empty interval including 0 of positive ϵ values in which the perturbation has constant rank. Since there are only finitely many possible values for the rank to take, proving an $\mathcal{O}^E(\epsilon)$ result for the case when the rank is known immediately leads to the corresponding $\mathcal{O}^E(\epsilon)$ for the case of unknown rank.

Lemma 8. (Compare lemma 5.)

In addition to the assumptions of lemma 3, assume the perturbed DAE is known to have index no more than 1, and that there exist $R_0 > 0$ and $\phi_0 < \pi/2$ like in lemma 3. Also assume that the ratio between the largest and smallest non-zero singular value of E is bounded by some constant κ_0. Then, for any fixed $t_1 \ge t_0$, for all $t \in [\, t_0,\, t_1\,]$,

$$|E(\epsilon)\, \eta_E(\, t,\, \epsilon\,)| = \mathcal{O}^E(\epsilon)$$

Proof. The case of index 0, when E is full-rank, was treated in lemma 5, so it remains to consider the case of index 1. When the rank is zero, $E = 0$ and it is immediately seen from (9) that η must be identically zero and the conclusion follows trivially. Hence, assume that the rank is neither full nor zero and let

$$E(\epsilon) = (U_1(\epsilon)\ U_2(\epsilon)) \begin{pmatrix} \Sigma(\epsilon) & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1(\epsilon)^{\mathrm{T}} \\ V_2(\epsilon)^{\mathrm{T}} \end{pmatrix}$$

be an SVD of $E(\epsilon)$ where $\Sigma(\epsilon)$ is of known dimensions and has condition number less than κ_0. It must be ensured that the components of the SVD have bounded derivatives, but the existence of such a factorization follows by modifying Steinbrecher (2006, theorem 2.4.1) to suit our needs. Applying the unknown change of variables $\eta = V(\epsilon) \begin{pmatrix} \eta'_1 \\ \eta'_2 \end{pmatrix}$ and the row operations represented by $U(\epsilon)^{\mathrm{T}}$, (9) turns into (dropping ϵ)

$$\begin{pmatrix} I & 0 & 0 \\ & \epsilon\, \Sigma & 0 \\ & 0 & 0 \end{pmatrix} \begin{pmatrix} \bar{\xi}(t) \\ \bar{\eta}'_1(t) \\ \bar{\eta}'_2(t) \end{pmatrix} = \begin{pmatrix} M_{11} + M_{12}\, L_E & M_{12}\, V_1 & M_{12}\, V_2 \\ 0 & A_{22} & A_{23} \\ 0 & A_{32} & A_{33} \end{pmatrix} \begin{pmatrix} \xi(t) \\ \bar{\eta}_1(t) \\ \bar{\eta}_2(t) \end{pmatrix}$$

where, for instance and in particular, (dropping ϵ)

$$A_{33} \stackrel{\triangle}{=} U_2^{\mathrm{T}} M_{22} V_2 - \epsilon U_2^{\mathrm{T}} E L_E M_{12} V_2 = U_2^{\mathrm{T}} M_{22} V_2$$

Since the DAE is known to be index 1, differentiation of the last group of equations shows that $A_{33}(\epsilon)$ is non-singular, and hence the change of variables

$$\begin{pmatrix} \bar{\eta}_1(t) \\ \bar{\eta}_2(t) \end{pmatrix} = \begin{pmatrix} I & 0 \\ -A_{33}(\epsilon)^{-1} A_{32}(\epsilon) & I \end{pmatrix} \begin{pmatrix} \bar{\bar{\eta}}_1(t) \\ \bar{\bar{\eta}}_2(t) \end{pmatrix} \tag{22}$$

leads to the DAE in $(\xi, \bar{\bar{\eta}}_1, \bar{\bar{\eta}}_2)$ with matrices (dropping ϵ)

$$\begin{pmatrix} I & 0 & 0 \\ 0 & \epsilon \Sigma & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$-\begin{pmatrix} M_{11} + M_{12} L_E & M_{12} V_1 - M_{12} V_2 A_{33}^{-1} A_{32} & M_{12} V_2 \\ 0 & A_{22} - A_{23} A_{33}^{-1} A_{32} & A_{23} \\ 0 & 0 & A_{33} \end{pmatrix}$$

It is seen that $\bar{\bar{\eta}}_2 = 0$ and that $\bar{\bar{\eta}}_1$ is given by an ODE with state-feedback matrix

$$M_{\bar{\bar{\eta}}_1}(\epsilon) \stackrel{\triangle}{=} \frac{1}{\epsilon} \Sigma(\epsilon)^{-1} \left(A_{22}(\epsilon) - A_{23}(\epsilon) A_{33}(\epsilon)^{-1} A_{32}(\epsilon) \right)$$

Just like in lemma 5 it needs to be shown that the eigenvalues of this matrix tend to infinity as $\epsilon \to 0$, independently of E, but here we need to recall that E is not only present in $\Sigma(\epsilon)$, but also in the unknown unitary matrices $U(\epsilon)$ and $V(\epsilon)$. Again, we do this by showing

$$\liminf_{\epsilon \to 0} \sup_{E} \left\| \left(\epsilon M_{\bar{\bar{\eta}}_1}(\epsilon) \right)^{-1} \right\|^{-1} > 0$$

Using $\|\Sigma(\epsilon)\| = \|E(\epsilon)\| \leq 1$, and that

$$\begin{pmatrix} A_{22}(\epsilon) & A_{23}(\epsilon) \\ A_{32}(\epsilon) & A_{33}(\epsilon) \end{pmatrix}^{-1} = \begin{pmatrix} \left(A_{22}(\epsilon) - A_{23}(\epsilon) A_{33}(\epsilon)^{-1} A_{32}(\epsilon) \right)^{-1} & ? \\ ? & ? \end{pmatrix}$$

implies

$$\left\| \begin{pmatrix} A_{22}(\epsilon) & A_{23}(\epsilon) \\ A_{32}(\epsilon) & A_{33}(\epsilon) \end{pmatrix}^{-1} \right\| \geq \left\| \left(A_{22}(\epsilon) - A_{23}(\epsilon) A_{33}(\epsilon)^{-1} A_{32}(\epsilon) \right)^{-1} \right\|$$

we find

$$\begin{aligned} &\left\| \left(\epsilon M_{\bar{\bar{\eta}}_1}(\epsilon) \right)^{-1} \right\|^{-1} \\ &= \left\| \left(A_{22}(\epsilon) - A_{23}(\epsilon) A_{33}(\epsilon)^{-1} A_{32}(\epsilon) \right)^{-1} \Sigma(\epsilon) \right\|^{-1} \\ &\geq \left\| \left(A_{22}(\epsilon) - A_{23}(\epsilon) A_{33}(\epsilon)^{-1} A_{32}(\epsilon) \right)^{-1} \right\|^{-1} \\ &\geq \left\| \left(U(\epsilon)^{\mathrm{T}} \left(M_{22}(\epsilon) - \epsilon E(\epsilon) L_E(\epsilon) M_{12}(\epsilon) \right) V(\epsilon) \right)^{-1} \right\|^{-1} \\ &= \left\| V(\epsilon)^{\mathrm{T}} \left(M_{22}(\epsilon) - \epsilon E(\epsilon) L_E(\epsilon) M_{12}(\epsilon) \right)^{-1} U(\epsilon) \right\|^{-1} \\ &= \left\| \left(M_{22}(\epsilon) - \epsilon E(\epsilon) L_E(\epsilon) M_{12}(\epsilon) \right)^{-1} \right\|^{-1} \end{aligned}$$

and just like in lemma 5 the expression gives that the eigenvalues tend to infinity uniformly with respect to E, and hence that ϵ can be chosen sufficiently small to make $|\bar{\bar{\eta}}_1|$ bounded by some factor times $|\bar{\bar{\eta}}_1(0)|$. Further,

$$|\bar{\bar{\eta}}_1(0)| = \left| \begin{pmatrix} \bar{\bar{\eta}}_1(0) \\ \bar{\bar{\eta}}_2(0) \end{pmatrix} \right| = \left| \begin{pmatrix} \bar{\eta}_1(0) \\ 0 \end{pmatrix} \right| \leq \left| \begin{pmatrix} \bar{\eta}_1(0) \\ \bar{\eta}_2(0) \end{pmatrix} \right| = |\eta_E^0(\epsilon)| = \mathcal{O}^E(\epsilon)$$

Using this, the conclusion finally follows by taking such a small ϵ: (dropping ϵ)

$$\begin{aligned} |E \eta_E(t, \epsilon)| &= \left| E V \begin{pmatrix} I & 0 \\ -A_{33}^{-1} A_{32} & I \end{pmatrix} \begin{pmatrix} \bar{\bar{\eta}}_1(t) \\ 0 \end{pmatrix} \right| \\ &\leq \left\| U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V^{\mathrm{T}} V \begin{pmatrix} I & 0 \\ -A_{33}^{-1} A_{32} & I \end{pmatrix} \right\| \mathcal{O}^E(\epsilon) \\ &= \left\| \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} \right\| \mathcal{O}^E(\epsilon) = \mathcal{O}^E(\epsilon) \end{aligned}$$

□

Corollary 9. Lemma 8 can be strengthened when z has only two components. Then, just like in lemma 5, the conclusion is

$$|\eta_E(t, \epsilon)| = \mathcal{O}^E(\epsilon)$$

Proof. The only rank of E that needs to be considered is 1, and then $A_{33}(\epsilon)^{-1} A_{32}(\epsilon)$ will be a scalar. From (22) it follows that $A_{33}(\epsilon)^{-1} A_{32}(\epsilon) \bar{\bar{\eta}}_1(0) = \mathcal{O}^E(\epsilon)$, which is then extended to all later times t, and hence

$$\left| \begin{pmatrix} \bar{\eta}_1(t) \\ \bar{\eta}_2(t) \end{pmatrix} \right| = \left| \begin{pmatrix} \bar{\bar{\eta}}_1(t) \\ -A_{33}(\epsilon)^{-1} A_{32}(\epsilon) \bar{\bar{\eta}}_1(t) \end{pmatrix} \right| = \mathcal{O}^E(\epsilon)$$

□

Theorem 7 can be extended as follows.

Theorem 10. Consider the setup (7), but rather than assuming that E be of bounded condition, it is assumed that E is a matrix with $\|E\| \leq 1$, bounded ratio between the non-zero singular values, and that the perturbed equation has index no more than 1. Except regarding E, the same assumptions that were made in theorem 7 are made here. Then

$$|x_E(t, \epsilon) - x(t, 0)| = \mathcal{O}^E(\epsilon) \tag{23}$$

$$|z_E(t, \epsilon) - z(t, 0)| = \mathcal{O}_E(\epsilon) \tag{24}$$

where the rather useless second equation is included for comparison with theorem 7.

Proof. Define $L_E(\epsilon)$ and $H_E(\epsilon)$ as above, and consider the solution expressed in the variables ξ and η. Lemma 8 shows how $E(\epsilon)\,\eta$ is bounded uniformly over time. Note that $x(t, 0)$ coincides with $\xi(t, 0)$, so the left hand side of (23) can be bounded as

$$\begin{aligned} &|x_E(t, \epsilon) - x(t, 0)| \\ &= |\xi_E(t, \epsilon) + \epsilon H_E(\epsilon) E(\epsilon) \eta_E(t, \epsilon) - \xi(t, 0)| \\ &\leq |\xi_E(t, \epsilon) - \xi(t, 0)| + \mathcal{O}^E(\epsilon^2) \end{aligned}$$

The conclusion concerning x then follows by an identical argument to that found in the proof of theorem 7. The weak conclusion regarding z follows by noting that, in lemma 8, given $E(\epsilon)$, $\left\|A_{33}(\epsilon)^{-1} A_{32}(\epsilon)\right\|$ approaches some finite value as $\epsilon \to 0$, since $A_{33}(\epsilon)$ must approach a non-singular matrix. □

Corollary 11. (Main theorem). Theorem 10 can be strengthened in case z has only two components. Then (24) can be written with $\mathcal{O}^E(\epsilon)$ on the right hand side.

Proof. Follows by using corollary 9 in the proof of theorem 10. □

4. DISCUSSION

To conclude, we make some remarks on the scope of the results and the assumptions used, and include an example that indicate a direction for future research.

4.1 Scope

We believe that the results presented have given insight into the properties of index reduction under uncertainty. However, there are obviously many desirable extensions. We note the following ones:

- More quantitative results.
- Replacement of singular value and condition number conditions with more intuitive or application oriented ones.
- Less conservative bounds in lemma 13.

Although more precise bounds in lemma 13 can readily be extracted from the proof, easily obtained bounds will not be good enough. Having excluded the possibility of bounding η by looking at the matrix exponential alone, it remains to explore the fact that we are actually not interested in knowing the *maximum* gain from initial conditions to later states of the trajectory of η, but the initial conditions are a function of E, and hence it might be sufficient to maximize over a subset of initial conditions.

4.2 Example

In this section we follow up the discussion on the condition number in the previous section by providing an example which should shed some more light on — and stimulate future research on — the problem of singular perturbation in DAE.

In this example, the bounding of η over time is considered in case η has two components. For simplicity, we shall assume that η is given by

$$\eta'(t) = \frac{1}{\epsilon} E^{-1} M_{22} \eta(t)$$

where $M_{22} = I$, and we set $\epsilon = 1$. By selecting E as

$$E = \begin{pmatrix} -\delta & 1-\delta \\ 0 & -\delta \end{pmatrix}$$

where $\delta > 0$ is a small parameter we ensure $\|E\| \leq 1$, and since

$$E^{-1} = \begin{pmatrix} -1/\delta & 1/\delta^2 - 1/\delta \\ 0 & -1/\delta \end{pmatrix}$$

we see that both eigenvalues are perfectly stable and far into the left half plane, while the off-diagonal element is at the same time arbitrarily big. It is easy to verify using software that the maximum norm of the matrix exponential grows without bound as δ tends to zero. This shows that using only the norm of the initial conditions is not enough if we would like to find a bound on $|\eta(t)|$ which does not depend on the condition number of E.

ACKNOWLEDGEMENTS

With financial support from the Swedish Research Council.

REFERENCES

Lars Hörmander. *An introduction to complex analysis in several variables*. The University Series in Higher Mathematics. D. Van Nostrand, Princeton, New Jersey, 1966.

Peter V. Kokotović, Hassan K. Khalil, and John O'Reilly. *Singular perturbation methods in control: Analysis and applications*. Academic Press Inc., 1986.

Peter Kunkel and Volker Mehrmann. *Differential-algebraic equations, analysis and numerical solution*. European Mathematical Society, 2006.

R. M. M Mattheij and P. M. E. J. Wijckmans. Sensitivity of solutions of linear DAE to perturbations of the system matrices. *Numerical Algorithms*, 19(1–4), 1998.

Andreas Steinbrecher. *Numerical solution of quasi-linear differential-algebraic equations and industrial simulation of multibody systems*. PhD thesis, Technischen Universität Berlin, 2006.

Charles F. Van Loan. The sensitivity of the matrix exponential. *SIAM J. Numer. Anal.*, 14(6):971–981, December 1977.

Appendix A. PRELIMINARIES

In this section, we state two bounds on the norm of the matrix exponential. They are much more simple than tight.

Lemma 12. Let A be a linear map from an n-dimensional space to itself. Let $\alpha(A)$ denote the largest real part of the eigenvalues of A. Then

$$\left\|\mathrm{e}^{A\,t}\right\| \leq \mathrm{e}^{\alpha(A)\,t} \sum_{i=0}^{n-1} \frac{(2\,\|A\|)^i\,t^i}{i!} \tag{A.1}$$

Proof. Let $Q^{\mathrm{H}} A Q = D + N$ be a Schur decomposition of A, meaning that Q is unitary, D diagonal, and N nilpotent. The following bound, derived in Van Loan (1977),

$$\left\|\mathrm{e}^{A\,t}\right\| \leq \mathrm{e}^{\alpha(A)\,t} \sum_{i=0}^{n-1} \frac{\|N\|^i\,t^i}{i!}$$

readily gives the result since $\|N\| = \|Q^{\mathrm{H}} A Q - D\| \leq \|A\| + \|A\|$. $\square$

Lemma 13. If the map A is Hurwitz, that is, $\alpha(A) < 0$, then for $t \geq 0$,

$$\left\|\mathrm{e}^{A\,t}\right\| \leq \mathrm{e}^{2\,\mathrm{e}^{-1}\,n\,\frac{\|A\|}{-\alpha(A)}}$$

Proof. Let $f(t) \overset{\triangle}{=} \left\|\mathrm{e}^{A\,t}\right\|$. From lemma 12 we have that

$$f(t) \leq \sum_{i=0}^{n-1} \frac{(2\,\|A\|)^i\,t^i}{i!} \mathrm{e}^{\alpha(A)\,t} =: \sum_i f_i(t)$$

Each $f_i(t)$ can easily be bounded globally since they are smooth, tend to 0 from above as $t \to \infty$, and the only stationary point is given by $f_i'(t)$. From

$$f_i'(t) = \mathrm{e}^{\alpha(A)\,t} \frac{(2\,\|A\|)^i\,t^{i-1}}{i!} \left(t\,\alpha(A) + i\right)$$

it follows that the stationary point is $t = -\frac{i}{\alpha(A)}$. Hence,

$$f_i(t) \leq f_i\left(-\frac{i}{\alpha(A)}\right) = \frac{\left(\frac{2\,\|A\|}{-\alpha(A)}\right)^i i^i}{i!} \mathrm{e}^{-i} \leq \frac{\left(2\,\mathrm{e}^{-1}\,n\,\frac{\|A\|}{-\alpha(A)}\right)^i}{i!}$$

and it follows that

$$f(t) \leq \sum_{i=0}^{n-1} \frac{\left(2\,\mathrm{e}^{-1}\,n\,\frac{\|A\|}{-\alpha(A)}\right)^i}{i!} \leq \sum_{i=0}^{\infty} \frac{\left(2\,\mathrm{e}^{-1}\,n\,\frac{\|A\|}{-\alpha(A)}\right)^i}{i!} = \mathrm{e}^{2\,\mathrm{e}^{-1}\,n\,\frac{\|A\|}{-\alpha(A)}}$$

$\square$