Electronic Transactions on Numerical Analysis.
Volume 4, pp. 138-157, December 1996.

ISSN 1068-9613.

ETNA
Kent State University
etna@mcs.kent.edu

LOCAL AND GLOBAL INVARIANTS OF LINEAR DIFFERENTIAL-ALGEBRAIC EQUATIONS AND THEIR RELATION*

PETER KUNKEL† AND VOLKER MEHRMANN‡

Abstract. We study local and global invariants of linear differential-algebraic equations with variable coefficients and their relation. In particular, we discuss the connection between different approaches to the analysis of such equations and the associated indices, which are the differentiation index and the strangeness index. This leads to a new proof of an existence and uniqueness theorem as well as to an adequate numerical algorithm for the solution of linear differential-algebraic equations.

Key words. differential-algebraic equations, invariants, differentiation index, strangeness index, normal form, existence and uniqueness.

AMS subject classification. 34A09.

1. Introduction.

In this paper we study the behaviour of linear differential-algebraic equations (DAE's)

$$E(t)\dot{x} = A(t)x + f(t), \tag{1.1}$$

possibly together with an initial condition

$$x(t_0) = x_0, \tag{1.2}$$

where

$$E, A \in C(\mathbb{I}, \mathbb{C}^{n,n}),\ f \in C(\mathbb{I}, \mathbb{C}^n),\ \mathbb{I} \subseteq \mathbb{R} \text{ a (closed) interval},\ t_0 \in \mathbb{I},\ x_0 \in \mathbb{C}^n, \tag{1.3}$$

with respect to existence and uniqueness of solutions (denoting the set of i-times continuously differentiable functions from the interval $\mathbb{I}$ into the complex $m \times n$ matrices by $C^i(\mathbb{I}, \mathbb{C}^{m,n})$).

Most approaches to this question (see, e.g., [5, 6, 9, 13, 15]) require a number of matrix functions to have constant rank. For example, it is common to require that the rank of E does not change on $\mathbb{I}$. On the other hand, it is well known that there are simple problems (see, e.g., the example at the end of Section 4) which have a unique solution for consistent initial values but which do not satisfy this constant rank condition. If numerical algorithms are based on such analytical theories (as, e.g., [10] on [9]) then these methods may fail or at least require some additional considerations.

Another approach to the analysis of (1.1) is based on the so-called differentiation index (see [3] or [1] and references therein). This approach avoids most constant rank assumptions, but numerical algorithms using this concept often do not exhibit the correct solution behaviour. Especially, they tend to violate equality constraints that are contained in (1.1), even in the simplest case $E(t) = 0$. This behaviour is often called drift-off.

* Received May 3, 1996. Accepted for publication November 6, 1996. Communicated by G. Ammar.
† Fachbereich Mathematik, Carl von Ossietzky Universität, Postfach 2503, D–26111 Oldenburg, Fed. Rep. Germany
‡ Fakultät für Mathematik, Technische Universität Chemnitz-Zwickau, D–09107 Chemnitz, Fed. Rep. Germany.
This work has been supported by *Deutsche Forschungsgemeinschaft, Research grant Me 790/5-1 Differentiell-algebraische Gleichungen*.

In principle, these two different approaches can be interpreted as follows. In the first case, one looks for local invariants which are numerically accessible and then requires them to be global to obtain analytical results. In the second case, one looks for global invariants which are suitable for the analytical treatment, with the disadvantage that a proper numerical treatment of (1.1) is not directly possible because the global invariants are not computable.

Of course, for a proper numerical treatment of (1.1), the algorithms should be based on an existence and uniqueness theorem which is as general as possible. The above observations, however, suggest that this will be difficult to achieve, since the more local approaches, which can calculate local information to obtain the correct solution behaviour, have the drawback that they are not applicable to some well-behaved problems, and the global approach, which would cover all well-behaved problems, has the drawback of drift-off and, what seems to be more important, of not being able to compute or check the global invariants.

The present paper therefore studies in detail the relation between local and global invariants in order to obtain a deeper insight in the difference between these two approaches. In particular, we show how the differentiation index of [1, 3] and the strangeness index of [9] are related and give a new proof of an existence and uniquenesss theorem, first stated in [3], under weaker smoothness assumptions. We then show that the numerical procedure of [10], with a slightly different termination criterion, is suitable for computing any unique solution of (1.1) and (1.2), provided the coefficients are as smooth as the new existence and uniqueness theorem requires.

The paper is organized as follows. In Section 2, we first give a brief outline of previous results in [9, 10] on the so-called strangeness index and of the results in [1, 3] on the differentiation index. We then discuss the relation between these two approaches in Section 3. In Section 4, we finally give a new existence and uniqueness result, generalizing the results in [9] and [3], and we discuss the numerical relevance of the obtained results.

2. Basic results. In the following, we briefly describe the results of [9, 10] on one side and of [1, 3] on the other side, to give the background and to introduce the necessary notation.

We start the presentation of the ideas in [9, 10] with the observation that (1.1) can be transformed into a DAE, of equal solution behaviour, by scaling the equation and the unknown by pointwise nonsingular matrix functions, leading to the following equivalence relation.

Definition 2.1. Two pairs of matrix functions (E_i, A_i), $E_i, A_i \in C(\mathbb{I}, \mathbb{C}^{n,n})$, $i = 1, 2$ are called *(globally) equivalent* if there exist pointwise nonsingular matrix functions $P \in C(\mathbb{I}, \mathbb{C}^{n,n})$ and $Q \in C^1(\mathbb{I}, \mathbb{C}^{n,n})$ such that

$$E_2 = PE_1Q, \quad A_2 = PA_1Q - PE_1\dot{Q}. \tag{2.1}$$

Since for a fixed $t \in \mathbb{I}$ we can choose P and Q in such a way that they assume prescribed values $P(t)$, $Q(t)$, and $\dot{Q}(t)$, this equivalence relation possesses a local version.

Definition 2.2. Two pairs of matrices (E_i, A_i), $E_i, A_i \in \mathbb{C}^{n,n}$, $i = 1, 2$ are called *(locally) equivalent* if there are matrices $P, Q, B \in \mathbb{C}^{n,n}$ with P, Q nonsingular such that

$$E_2 = PE_1Q, \quad A_2 = PA_1Q - PE_1B. \tag{2.2}$$

ETNA
Kent State University
etna@mcs.kent.edu

Given $E, A \in \mathbb{C}^{n,n}$, it was shown in [9] that the quantities (with the convention $\operatorname{rank} \emptyset = 0$)

$$r = \operatorname{rank} E, \quad a = \operatorname{rank}(Z^* A T), \quad s = \operatorname{rank}(V^* Z^* A T') \tag{2.3}$$

are invariants with respect to local equivalence. Here the columns of T, T', Z, and V span kernel E, cokernel E, range E, and corange(Z^*AT) respectively. Applying this equivalence transformation pointwise to matrix functions $E, A \in C(\mathbb{I}, \mathbb{C}^{n,n})$ yields functions $r, a, s \colon \mathbb{I} \to \mathbb{N}_0$. If one requires the constant rank condition

$$r(t) \equiv r, \quad a(t) \equiv a, \quad s(t) \equiv s \tag{2.4}$$

on $\mathbb{I}$, i.e., if one requires that the local invariants bear global information, then the pair (E, A) can be transformed to the globally equivalent pair

$$\left(\begin{bmatrix} I_s & 0 & 0 & 0 & 0 \\ 0 & I_d & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & A_{12} & 0 & A_{14} & A_{15} \\ 0 & 0 & 0 & A_{24} & A_{25} \\ 0 & 0 & I_a & 0 & 0 \\ I_s & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right) \begin{array}{c} s \\ d \\ a \\ s \\ u \end{array}, \tag{2.5}$$

where we have used the abbreviations $d = r - s$ and $u = n - r - a - s$. In terms of the DAE (1.1) this reads

$$\begin{array}{ll} \text{(a)} & \dot{x}_1 = A_{12}(t)x_2 + A_{14}(t)x_4 + A_{15}(t)x_5 + f_1(t), \\ \text{(b)} & \dot{x}_2 = A_{24}(t)x_4 + A_{25}(t)x_5 + f_2(t), \\ \text{(c)} & 0 = x_3 + f_3(t), \\ \text{(d)} & 0 = x_1 + f_4(t), \\ \text{(e)} & 0 = f_5(t). \end{array} \tag{2.6}$$

Differentiating the fourth equation and eliminating $\dot{x}_1$ in the first equation does not alter the solution set. This corresponds to the replacement of the identity in the upper left corner of (2.5) by zero. Starting with $(E_0, A_0) = (E, A)$, and repeating the procedure of transformation to the form (2.6) and inserting the differentiated equation (2.6d) into equation (2.6a) leads to an iterative definition of sequences (E_i, A_i) of pairs of matrix functions and (r_i, a_i, s_i) of characteristic values when one requires constant rank assumptions as (2.4) in each step. Since $r_{i+1} = r_i - s_i$, the process terminates when s_i becomes zero and there is nothing left to be differentiated. Note that the characteristic values are invariant under global equivalence transformations and hence the following value is also invariant under global equivalence transformations.

Definition 2.3. The value

$$\mu = \min\{ i \in \mathbb{N}_0 \mid s_i = 0 \} \tag{2.7}$$

is called the *strangeness index* of (E, A) or of (1.1).

If μ is well-defined, i.e., if the above process can be executed until $s_i = 0$ is reached, we have transformed (1.1) into a DAE of the form

$$\begin{array}{ll} \text{(a)} & \dot{x}_1 = A_{13}(t)x_3 + f_1(t), \\ \text{(b)} & 0 = x_2 + f_2(t), \\ \text{(c)} & 0 = f_3(t), \end{array} \tag{2.8}$$

whose solution set is in one-to-one correspondence with that of the original problem via a transformation with a pointwise nonsingular matrix function. The sizes of

ETNA
Kent State University
etna@mcs.kent.edu

the unknowns x_1, x_2, x_3 and of the inhomogeneities f_1, f_2, f_3 are given by d_μ, a_μ, u_μ, respectively. The basic properties of (1.1) can now be read off directly from (2.8); see [9].

THEOREM 2.4. *Let μ be well-defined for the sufficiently smooth pair (E, A) in (1.1) and let $f \in C^{\mu+1}(\mathbb{I}, \mathbb{C}^n)$. Then the following holds.*

1. *The equation (1.1) is solvable, i.e., it has at least one solution $x \in C^1(\mathbb{I}, \mathbb{C}^n)$, if and only if the u_μ functional consistency conditions*
$$f_3(t) \equiv 0 \tag{2.9}$$
are satisfied.
2. *An initial condition (1.2) is consistent, i.e., the corresponding initial value problem has at least one solution, if and only if in addition (1.2) implies the a_μ conditions*
$$x_2(t_0) = -f_2(t_0). \tag{2.10}$$
3. *The initial value problem (1.1) with (1.2) is uniquely solvable, if and only if in addition we have*
$$u_\mu = 0. \tag{2.11}$$

Moreover, one has a normal form of (E, A) with respect to global equivalence (cf. [9]).

THEOREM 2.5. *Let μ be well-defined for the sufficiently smooth pair (E, A) in (1.1) and let (r_i, a_i, s_i), $i = 0, \ldots, \mu$, be the corresponding sequence of characteristic values. Furthermore, let*
$$w_0 = u_0, \qquad w_i = u_i - u_{i-1}, \quad i = 1, \ldots, \mu, \tag{2.12}$$
and
$$c_0 = a_0 + s_0, \qquad c_i = s_{i-1} - w_i, \quad i = 1, \ldots, \mu, \tag{2.13}$$
with $d_i = r_i - s_i$ and $u_i = n - r_i - a_i - s_i$. Then (E, A) is globally equivalent to a matrix pair of the form
$$\left(\left[\begin{array}{c|ccc|cccc} I & 0 & \cdots & 0 & 0 & * & \cdots & * \\ \hline 0 & 0 & \cdots & 0 & 0 & F_\mu & & * \\ \vdots & \vdots & & \vdots & & \ddots & \ddots & \\ \vdots & \vdots & & \vdots & & & \ddots & F_1 \\ 0 & 0 & \cdots & 0 & & & & 0 \\ \hline 0 & 0 & \cdots & 0 & 0 & G_\mu & & * \\ \vdots & \vdots & & \vdots & & \ddots & \ddots & \\ \vdots & \vdots & & \vdots & & & \ddots & G_1 \\ 0 & 0 & \cdots & 0 & & & & 0 \end{array}\right], \left[\begin{array}{c|ccc|cccc} * & * & \cdots & * & 0 & \cdots & \cdots & 0 \\ \hline 0 & 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & & \vdots \\ \vdots & \vdots & & \vdots & \vdots & & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ \hline 0 & 0 & \cdots & 0 & I & & & \\ \vdots & \vdots & & \vdots & & \ddots & & \\ \vdots & \vdots & & \vdots & & & \ddots & \\ 0 & 0 & \cdots & 0 & & & & I \end{array}\right] \right) \begin{array}{c} d_\mu \\ w_\mu \\ \vdots \\ \vdots \\ w_0 \\ c_\mu \\ \vdots \\ \vdots \\ c_0 \end{array} \tag{2.14}$$
where
$$\operatorname{rank} \left[\begin{array}{c} F_i \\ G_i \end{array}\right] = c_i + w_i = s_{i-1} \le c_{i-1}, \tag{2.15}$$

ETNA
Kent State University
etna@mcs.kent.edu

which means that the matrix functions F_i and G_i together have pointwise full row rank.

The above theoretical approach does not allow for a numerical treatment of (1.1), since the global equivalence transformations cannot be determined numerically. In [10], a method was developed that allows for the numerical determination of the invariants as well as for the numerical treatment of (1.1) in the case of a well-defined strangeness index. This method is based on an idea of Campbell (see, e.g., [3]) of building *inflated DAE's*

$$M_\ell(t)\dot{z}_\ell = N_\ell(t)z_\ell + g_\ell(t), \tag{2.16}$$

where

$$\begin{aligned} (M_\ell)_{i,j} &= \tbinom{i}{j}E^{(i-j)} - \tbinom{i}{j+1}A^{(i-j-1)},\ i,j = 0,\dots,\ell, \\ (N_\ell)_{i,j} &= \begin{cases} A^{(i)} & \text{for } i = 0,\dots,\ell,\ j = 0, \\ 0 & \text{else,} \end{cases} \\ (z_\ell)_i &= x^{(i)},\ i = 0,\dots,\ell, \\ (g_\ell)_i &= f^{(i)},\ i = 0,\dots,\ell, \end{aligned} \tag{2.17}$$

is obtained by successive differentiation of (1.1) (with the convention that $\binom{i}{j} = 0$ for $i < 0,\ j < 0$ or $j > i$). The key observation in [10] is that if μ is well-defined, then the inflated pairs of two globally equivalent pairs of matrix functions are locally equivalent for each $\ell \in \mathbb{N}_0$ and each $t \in \mathbb{I}$.

LEMMA 2.6. *Let the pairs (E, A) and $(\tilde{E}, \tilde{A})$ of matrix functions be sufficiently smooth and globally equivalent via*

$$\tilde{E} = PEQ, \quad \tilde{A} = PAQ - PE\dot{Q}. \tag{2.18}$$

Suppose, furthermore, that the strangeness index μ is well-defined and let (M_ℓ, N_ℓ) and $(\tilde{M}_\ell, \tilde{N}_\ell)$ be the corresponding inflated pairs. Then the matrix pairs $(M_\ell(t), N_\ell(t))$ and $(\tilde{M}_\ell(t), \tilde{N}_\ell(t))$ are locally equivalent for each $\ell \in \mathbb{N}_0$ and each $t \in \mathbb{I}$ via

$$(\tilde{M}_\ell(t), \tilde{N}_\ell(t)) = \Pi_\ell(t)(M_\ell(t), N_\ell(t)) \begin{bmatrix} \Theta_\ell(t) & -\Psi_\ell(t) \\ 0 & \Theta_\ell(t) \end{bmatrix}, \tag{2.19}$$

where

$$\begin{aligned} (\Pi_\ell)_{i,j} &= \tbinom{i}{j}P^{(i-j)}, \quad (\Theta_\ell)_{i,j} = \tbinom{i+1}{j+1}Q^{(i-j)}, \\ (\Psi_\ell)_{i,j} &= \begin{cases} Q^{(i+1)} & \text{for } i = 0,\dots,\ell,\ j = 0, \\ 0 & \text{else.} \end{cases} \end{aligned} \tag{2.20}$$

Thus, the local characteristic values $(\tilde{r}_\ell, \tilde{a}_\ell, \tilde{s}_\ell)$ of $(M_\ell(t), N_\ell(t))$ for $\ell = 0,\dots,\mu$ are also characteristic values for (E, A). Moreover, the relations

$$\begin{aligned} \tilde{r}_\ell &= (\ell+1)n - (c_0 + \cdots + c_\ell) - (u_0 + \cdots + u_\ell), \\ \tilde{a}_\ell &= s_{\ell-1} - w_\ell - s_\ell = c_\ell - s_\ell, \\ \tilde{s}_\ell &= s_\ell + (c_0 + \cdots + c_{\ell-1}), \\ \tilde{d}_\ell &= \tilde{r}_\ell - s_\ell = (\ell+1)n - c_\ell - (u_0 + \cdots + u_\ell), \\ \tilde{u}_\ell &= (\ell+1)n - \tilde{r}_\ell - \tilde{a}_\ell - \tilde{s}_\ell = u_0 + \cdots + u_\ell, \end{aligned} \tag{2.21}$$

$\ell = 0,\dots,\mu$, hold between the local characteristic values $(\tilde{r}_\ell, \tilde{a}_\ell, \tilde{s}_\ell)$ of $(M_\ell(t), N_\ell(t))$ and the global characteristic values (r_i, a_i, s_i) of (E, A); see [10]. Since local characteristic values (2.3) are numerically computable by three successive rank determinations, the sequence $(\tilde{r}_\ell, \tilde{a}_\ell, \tilde{s}_\ell)$, and therefore also the sequence (r_i, a_i, s_i), are numerically computable as well. Moreover, this even leads to a different (local) definition of the characteristic values of (E, A) in the form of functions $\mu, r_i, a_i, s_i \colon \mathbb{I} \to \mathbb{N}_0$.

In the case of a well-defined strangeness index μ and $u_\mu = 0$ (i.e., the initial value problem for consistent initial conditions has a unique solution), it was then shown in [10] that the pair (E, A) satisfies the following hypothesis by setting $\hat{\mu} = \mu$, $\hat{a} = a_\mu$, and $\hat{d} = d_\mu$.

Hypothesis 2.7. *There exist integers $\hat{\mu}$, $\hat{a}$, and $\hat{d}$ such that the inflated pair $(M_{\hat{\mu}}, N_{\hat{\mu}})$ associated with (E, A) has the following properties:*

1. *For all $t \in \mathbb{I}$ it holds that $\operatorname{rank} M_{\hat{\mu}}(t) = (\hat{\mu}+1)n - \hat{a}$, such that there exists a smooth matrix function $\hat{Z}_2$ with orthonormal columns and size $((\hat{\mu}+1)n, \hat{a})$ satisfying $\hat{Z}_2^* M_{\hat{\mu}} = 0$.*
2. *For all $t \in \mathbb{I}$ it holds that $\operatorname{rank} \hat{A}_2(t) = \hat{a}$, where $\hat{A}_2 = \hat{Z}_2^* N_{\hat{\mu}} [I_n \, 0 \, \cdots \, 0]^*$, such that there exists a smooth matrix function $\hat{T}_2$ with orthonormal columns and size $(n, \hat{d})$, $\hat{d} = n - \hat{a}$, satisfying $\hat{A}_2 \hat{T}_2 = 0$.*
3. *For all $t \in \mathbb{I}$ it holds that $\operatorname{rank} E(t)\hat{T}_2(t) = \hat{d}$, such that there is a smooth matrix function $\hat{Z}_1$ with orthonormal columns and size $(n, \hat{d})$ yielding that $\hat{E}_1 = \hat{Z}_1^* E$ has constant rank $\hat{d}$.*

Note that it has been proved in [10] that the above hypothesis is invariant under global equivalence transformations.

Additionally setting $\hat{A}_1 = \hat{Z}_1^* A$ and $\hat{f}_1 = \hat{Z}_1^* f$, $\hat{f}_2 = \hat{Z}_2^* g_{\hat{\mu}}$, for sufficiently smooth f, yields a new DAE

$$\begin{bmatrix} \hat{E}_1(t) \\ 0 \end{bmatrix} \dot{x} = \begin{bmatrix} \hat{A}_1(t) \\ \hat{A}_2(t) \end{bmatrix} x + \begin{bmatrix} \hat{f}_1(t) \\ \hat{f}_2(t) \end{bmatrix} \tag{2.22}$$

of the same size as (1.1). We also use (2.22) in the notation $\hat{E}(t)\dot{x} = \hat{A}(t)x + \hat{f}(t)$ with

$$\hat{E} = Z^* M_{\hat{\mu}} [I_n \, 0 \, \cdots \, 0]^*, \quad \hat{A} = Z^* N_{\hat{\mu}} [I_n \, 0 \, \cdots \, 0]^*, \quad \hat{f} = Z^* g_{\hat{\mu}}. \tag{2.23}$$

It has also been shown in [10], that for well-defined μ, the system (2.22) has the same solutions as (1.1). Since (2.22) is numerically computable (up to a scaling from the left) and has local characteristic values $(\hat{r}, \hat{a}, \hat{s}) = (d_\mu, a_\mu, 0)$, it can be solved numerically by all integration schemes, which are suited for general so-called index-1 problems, like BDF (as, e.g., implemented in `DASSL` of [14]) or special Runge-Kutta schemes (as, e.g., implemented in `RADAU5` of [8]); see also `GELDA` of [12].

The second part of our review is concerned with the relevant results of [1, 3] on the so-called differentiation index.

Definition 2.8. The DAE (1.1) has *differentiation index* $\nu \in \mathbb{N}_0$ if ν is the smallest number ℓ such that (2.16) determines $\dot{x}$ as a function of t and x.

In other words, for well-defined differentiation index, every solution of the DAE is also a solution of an ODE

$$\dot{x} = h(t, x), \tag{2.24}$$

the so-called *underlying ODE*. It is clear from Theorem 2.4 that the differentiation index cannot be well-defined for problems with well-defined strangeness index μ and $u_\mu \neq 0$, because of the infinite-dimensional solution space of the corresponding homogeneous problem.

ETNA
Kent State University
etna@mcs.kent.edu

Closely connected with the differentiation index is the notion of 1-fullness for block matrices.

DEFINITION 2.9. Given $n \in \mathbb{N}$, a matrix $M \in \mathbb{C}^{kn,ln}$ with $k, l \in \mathbb{N}$ is called *1-full* if there is a nonsingular matrix $R \in \mathbb{C}^{kn,kn}$ such that

$$RM = \begin{bmatrix} I_n & 0 \\ 0 & H \end{bmatrix}. \tag{2.25}$$

In particular one has the following result.

THEOREM 2.10. *The differentiation index ν is the smallest number ℓ such that $M_\ell(t)$ is 1-full for all $t \in \mathbb{I}$ and $\operatorname{rank} M_\ell(t)$ is constant on $\mathbb{I}$. Moreover, $\operatorname{rank}(M_\ell(t), N_\ell(t)) = (\ell+1)n$ for all $t \in \mathbb{I}$ and for all $\ell = 0, \ldots, \nu$.*

Note that all these properties of (E, A) are invariant under global equivalence. Here one has the following existence and uniqueness result; see [1, 3].

THEOREM 2.11. *Let E, A, and f in (1.1) be sufficiently smooth and so that (1.1) has the differentiation index ν. Then the following holds.*

1. *An initial condition (1.2) is consistent if and only if*

$$g_\nu(t_0) - N_\nu(t_0)[I_n \; 0 \; \cdots \; 0]^* x_0 \in \operatorname{range} M_\nu(t_0), \tag{2.26}$$

   where M_ν, N_ν, g_ν are defined as in (2.16).
2. *The solutions of (1.1) coincide with those of the underlying ODE (2.24) for consistent initial conditions.*
3. *Let (1.2) be consistent. Then the solutions of (1.1) and (1.2) coincide with those of (2.24) and (1.2). In particular, the solution is unique.*

This finishes our brief summary of two approaches to the analysis of linear DAE's with variable coefficients which lead to two different existence and uniqueness results. In the next section we will discuss the relationship between these approaches.

3. Relation between local and global invariants. Both approaches sketched in the previous section seem to cover different aspects of linear DAE's but neither of them contains the other. The approach based on the strangeness index includes undetermined solution components but requires a number of constant rank conditions, whereas the approach based on the differentiation index does not need such constant rank conditions but excludes undetermined solution components by construction. Another difference, already discussed in the introduction, is that the concept behind the strangeness index is to start with local invariants and require them to be global, whereas the differentiation index is, per construction, a global invariant. It is the aim of this section to study in detail the connection between these two approaches. In particular, we discuss the relation between local and global invariants.

We start with the following observation for the rank of continuous matrix functions, see, e.g., [4, Ch. 10].

THEOREM 3.1. *Let $\mathbb{I} \subseteq \mathbb{R}$ be a closed interval and $M \in C(\mathbb{I}, \mathbb{C}^{m,n})$. Then, there exist open intervals $\mathbb{I}_j \subseteq \mathbb{I}$, $j \in \mathbb{N}$, with*

$$\overline{\bigcup_{j\in\mathbb{N}} \mathbb{I}_j} = \mathbb{I}, \quad \mathbb{I}_i \cap \mathbb{I}_j = \emptyset \text{ for } i \neq j, \tag{3.1}$$

and integers $r_j \in \mathbb{N}_0$, $j \in \mathbb{N}$, such that

$$\operatorname{rank} M(t) = r_j \text{ for all } t \in \mathbb{I}_j. \tag{3.2}$$

Applying this property of a continuous matrix function to the construction leading to the strangeness index given in Section 2, one immediately obtains the following result.

COROLLARY 3.2. *Let $\mathbb{I} \subseteq \mathbb{R}$ be a closed interval and $E, A: \mathbb{I} \to \mathbb{C}^{n,n}$ be sufficiently smooth. Then there exist open intervals $\mathbb{I}_j$, $j \in \mathbb{N}$, as in Theorem 3.1 such that the strangeness index of (E, A), restricted to $\mathbb{I}_j$, is well-defined for every $j \in \mathbb{N}$.*

Note that a similar result cannot hold for the differentiation index. A necessary condition for ν to be defined on $\mathbb{I}_j$ is that $u_\mu = 0$ on $\mathbb{I}_j$. The results of [9] show that this is also sufficient.

THEOREM 3.3. *Let (E, A) be sufficiently smooth. Furthermore, let the strangeness index μ be well-defined for (E, A) and let $u_\mu = 0$. Then the differentiation index ν is well-defined for (E, A) as well and we have*

$$\nu = \begin{cases} 0 & \text{for } a_\mu = 0, \\ \mu + 1 & \text{for } a_\mu \neq 0. \end{cases} \tag{3.3}$$

Proof. For well-defined μ with $u_\mu = 0$, Hypothesis 2.7 holds and (1.1) can be transformed to (2.22). If $a_\mu = 0$, then the matrix on the left hand side of (2.22) is $\hat{E}_1$, and Hypothesis 2.7 guarantees that it is nonsingular. If $a_\mu \neq 0$, then differentiation of the algebraic equation in (2.22) yields an equation of the form

$$\begin{bmatrix} \hat{E}_1 \\ -\hat{A}_2 \end{bmatrix} \dot{x} = \begin{bmatrix} \hat{A}_1 \\ \dot{\hat{A}}_2 \end{bmatrix} x + \begin{bmatrix} \hat{f}_1 \\ \dot{\hat{f}}_2 \end{bmatrix},$$

and Hypothesis 2.7 again guarantees that the matrix on the left hand side is nonsingular (see [10]). In both cases, multiplying with the inverse yields an ODE, i.e., in both cases the differentiation index is well-defined. In [9] it has been shown that then (3.3) holds. ▯

COROLLARY 3.4. *Let (E, A) be sufficiently smooth. If the differentiation index ν is well-defined for (E, A), it is well-defined for every restriction of (E, A) on $\mathbb{I}_j$. Let ν_j denote the differentiation index and μ_j the strangeness index on $\mathbb{I}_j$. Then we have*

$$\mu_j = \max\{0, \nu_j - 1\}, \tag{3.4}$$

and

$$\nu_j \leq \nu. \tag{3.5}$$

Proof. The first relation follows from Theorem 3.3. The second relation holds, since on a smaller interval a smaller number of differentiations may be sufficient to obtain an underlying ODE. ▯

Our next aim is to show that a pair (E, A), for which the differentiation index is well-defined, also satisfies Hypothesis 2.7 for some choice of $\hat{\mu}$, $\hat{a}$, and $\hat{d}$. To do so, we must first determine the corange (left nullspace) of M_ν. According to [3, 9, 10] we are allowed to restrict ourselves to the normal form (2.14) of (E, A) which can be written as

$$(E, A) = \left(\begin{bmatrix} I & C \\ 0 & G \end{bmatrix}, \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix} \right) \tag{3.6}$$

when we work on a specific interval $\mathbb{I}_j$. Let ν be the corresponding differentiation index. The quantity G in (3.6) then is a matrix function that is strictly upper triangular such that every arbitrary ν-fold product of G and its derivatives vanishes.

Since all diagonal blocks of the lower block triangular matrix function M_ν are E itself, the corange vectors must have zero entries where they encounter the identity I of E. Thus we may further restrict ourselves to the case

$$(E, A) = (G, I). \tag{3.7}$$

We now consider the infinite matrix function

$$M = \begin{bmatrix} G & & & \\ \dot{G} - I & G & & \\ \ddot{G} & 2\dot{G} - I & G & \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix}, \tag{3.8}$$

built according to (2.17). Looking for a matrix function Z of maximal rank satisfying $Z^*M = 0$, we must solve

$$[Z_0^* \; Z_1^* \; Z_2^* \; \cdots] \left\{ \begin{bmatrix} G & & & \\ \dot{G} & G & & \\ \ddot{G} & 2\dot{G} & G & \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix} - \begin{bmatrix} 0 & & & \\ I & 0 & & \\ & I & 0 & \\ & & \ddots & \ddots \end{bmatrix} \right\} = 0,$$

where $Z^* = [Z_0^* \; Z_1^* \; Z_2^* \; \cdots]$. Setting $Z_0 = I$, a simple manipulation yields

$$[Z_1^* \; Z_2^* \; \cdots] = [G \, 0 \cdots] \left\{ \begin{bmatrix} I & & & \\ & I & & \\ & & \ddots & \\ & & & \ddots \end{bmatrix} - \begin{bmatrix} \dot{G} & G & & \\ \ddot{G} & 2\dot{G} & G & \\ \vdots & \ddots & \ddots & \ddots \\ \vdots & & \ddots & \ddots \end{bmatrix} \right\}^{-1} \tag{3.9}$$

showing that M has a corange whose dimension equals the size of the blocks. Observe that the infinite matrix on the right hand side is indeed invertible, since it is of the form of an identity matrix minus a nilpotent matrix and that, although we formally treat infinite matrices, all expression become finite when we apply the requirement that all ν-fold products of G and its derivatives vanish.

Using the Neumann series and an induction argument on the number of factors of G and its derivatives in the first block row of the inverse in (3.9) yields that Z_j is a sum of at least j-fold products; hence,

$$Z_j = 0 \text{ for } j \geq \nu. \tag{3.10}$$

Now let $\tilde{M}$ be the matrix that is obtained from M by discarding its first block row and block column, i.e., let $\tilde{M} = S^*MS$ with the block up-shift matrix

$$S = \begin{bmatrix} 0 & & & \\ I & 0 & & \\ & I & 0 & \\ & & \ddots & \ddots \end{bmatrix}. \tag{3.11}$$

The same arguments as for M then show that the dimension of the corange of $\tilde{M}$ equals the dimension of the corange of M. The relation between these coranges can be described as follows.

LEMMA 3.5. *Let $Z^*M = 0$ hold for M as in (3.8) with smooth Z. Then,*

$$(Z^* + \dot{Z}^*S)\tilde{M} = 0. \tag{3.12}$$

Proof. By definition we have

$$\begin{aligned}(SM + S\dot{M}S)_{i,j} &= M_{i-1,j} + \dot{M}_{i-1,j+1} \\ &= \tbinom{i-1}{j}E^{(i-j-1)} - \tbinom{i-1}{j+1}A^{(i-j-2)} + \tbinom{i-1}{j+1}E^{(i-j-1)} - \tbinom{i-1}{j+2}A^{(i-j-2)} \\ &= \tbinom{i}{j+1}E^{(i-j-1)} - \tbinom{i}{j+2}A^{(i-j-2)} = M_{i,j+1} = (MS)_{i,j};\end{aligned}$$

hence, $MS = SM + S\dot{M}S$ and therefore $\tilde{M} = S^*MS = S^*(SM + S\dot{M}S) = M + \dot{M}S$ such that

$$\begin{aligned}(Z^* + \dot{Z}^*S)(M + \dot{M}S) &= Z^*M + Z^*\dot{M}S + \dot{Z}^*SM + \dot{Z}^*S\dot{M}S \\ &= Z^*\dot{M}S + \dot{Z}^*MS = \tfrac{d}{dt}(Z^*M)S = 0.\end{aligned}$$

□

Note that Z^* can be retrieved from $Z^* + \dot{Z}^*S$ by observing that

$$W^* = Z^* + \dot{Z}^*S \quad \Longleftrightarrow \quad Z^* = \sum_{k\geq 0}(-1)^k \left(\frac{d}{dt}\right)^k W^*S^k. \tag{3.13}$$

With these preparations we find the following properties of the inflated matrices when the differentiation index is well-defined.

LEMMA 3.6. *Let (E, A) be sufficiently smooth and let the differentiation index ν be well-defined for (E, A) with $\nu \geq 1$. Then,*

$$\operatorname{corank} M_\nu(t) = \operatorname{corank} M_{\nu-1}(t) \text{ for all } t \in \mathbb{I}. \tag{3.14}$$

Proof. From Theorem 2.10 we have that $\operatorname{corank} M_\nu(t)$ is constant on $\mathbb{I}$. Because row rank and column rank are equal, property (2.25) for $M_\nu(t)$ implies that, for every $t \in \mathbb{I}$, the matrix H has constant corank equal to that of $M_\nu(t)$. Because H is obtained by row operations on $M_\nu(t)$ with first block row and first block columm discarded, the corank of H equals the size of G in the normal form (3.6) of (E, A) on $\mathbb{I}_j$. But since $Z_\nu = 0$ from (3.10), the corank of $M_{\nu-1}(t)$ already equals the size of G on $\mathbb{I}_j$. Since $\operatorname{corank} M_\nu(t) \geq \operatorname{corank} M_{\nu-1}(t)$ by construction, it follows $\operatorname{rank} M_{\nu-1}(t) \geq \nu n - \operatorname{corank} M_\nu(t)$ and equality holds on a dense subset of $\mathbb{I}$. Because $\operatorname{corank} M_\nu(t)$ is constant on $\mathbb{I}$ and the rank is continuous from below, equality holds on the whole interval $\mathbb{I}$. □

THEOREM 3.7. *Let (E, A) be sufficiently smooth and let the differentiation index ν be well-defined for (E, A). Then (E, A) satisfies Hypothesis 2.7 with the setting*

$$\hat{\mu} = \max\{0, \nu - 1\}, \quad \hat{a} = \left\{\begin{array}{ll} 0 & \text{for } \nu = 0, \\ \operatorname{corank} M_{\nu-1}(t) & \text{otherwise,}\end{array}\right. , \quad \hat{d} = n - \hat{a}. \tag{3.15}$$

Proof. The claim is trivial for $\nu = 0$. We therefore assume $\nu \geq 1$. Lemma 3.6 implies that $\hat{a} = \operatorname{corank} M_{\nu-1}(t)$ is constant on $\mathbb{I}$ such that $\hat{Z}_2$ and $\hat{T}_2$ can be chosen according to the requirements of Hypothesis 2.7. If (E, A) is in the normal form (3.6), we obtain $\hat{T}_2(t) = [I \;\; 0]^*$ and $\operatorname{rank} E(t)\hat{T}_2(t) = n - \hat{a}$ on a dense subset of $\mathbb{I}$ and therefore on the whole interval $\mathbb{I}$. The claim follows, since all relevant quantities are invariant under global equivalence. □

COROLLARY 3.8. *Let (E, A) be sufficiently smooth and let it satisfy Hypothesis 2.7 with $\hat{\mu}$ and $\hat{a}$. Then the differentiation index ν is well-defined with $\nu = 0$ for $\hat{\mu} = 0$, $\hat{a} = 0$ and $\nu \leq \hat{\mu} + 1$ otherwise. If $\hat{\mu}$ is minimally chosen then equality holds in the latter relation.*

Proof. The proof is trivial for the first part of the claim. As in the proof of Theorem 3.3 it follows for the second part that the differentiation index ν is well-defined with $\nu \leq \hat{\mu} + 1$. The previous theorem then shows that equality holds when $\hat{\mu}$ is minimally chosen. □

Having discussed the connection between the different global characteristic values, the question remains which information is available locally, especially information that can be used in a numerical algorithm. We therefore examine now the local invariants $\mu, r_i, a_i, s_i \colon \mathbb{I} \to \mathbb{N}_0$ defined by (2.21). Note again that, per construction, the global invariants leading to the strangeness index are local when we restrict (E, A) to an interval $\mathbb{I}_j$. Hence, we must pay attention only to the boundary of the union of the intervals $\mathbb{I}_j$.

THEOREM 3.9. *Let (E, A) be sufficiently smooth and let the differentiation index ν be well-defined for (E, A). Then (2.21) defines local invariants $\mu, r_i, a_i, s_i, d_i, u_i \colon \mathbb{I} \to \mathbb{N}_0$, $i \in \mathbb{N}_0$, satisfying*

$$\begin{aligned} \mu(t) \leq \max\{0, \nu - 1\},\;\; r_{\mu(t)}(t) = d_{\mu(t)}(t) = \hat{d}, \\ a_{\mu(t)}(t) = \hat{a},\;\; s_{\mu(t)}(t) = u_{\mu(t)}(t) = 0 \end{aligned} \tag{3.16}$$

for all $t \in \mathbb{I}$ where $\hat{a}$ is taken from Theorem 3.7 and $\hat{d} = n - \hat{a}$.

Proof. Again the claim is trivial for $\nu = 0$, so we may assume that $\nu \geq 1$. Since the size of G in (3.6) equals a_{μ_j} on $\mathbb{I}_j$, we have that (3.16) holds on a dense subset of $\mathbb{I}$. To show (3.16) for the whole interval $\mathbb{I}$, let $t \in \mathbb{I}$ be fixed and $\mu = \nu - 1$. Since $\tilde{a}_\ell + \tilde{s}_\ell$, as in (2.21), is the rank of the part of $N_\ell(t)$ that belongs to the corange of $M_\ell(t)$ and this part also occurs in the corange of $M_{\ell+1}(t)$, we have

$$\tilde{a}_{\ell+1} + \tilde{s}_{\ell+1} \geq \tilde{a}_\ell + \tilde{s}_\ell.$$

By (2.21), we obtain (omitting arguments)

$$c_\ell \geq 0,$$

and since $(M_\ell(t), N_\ell(t))$ has full row rank (see Theorem 2.10), we have $\tilde{u}_\ell = 0$ for all ℓ, which implies that

$$u_\ell = 0, \quad w_\ell = 0, \quad s_{\ell-1} = c_\ell, \quad s_{\ell-1} \geq 0.$$

Since $\tilde{a}_\ell \geq 0$, we find $c_\ell \geq s_\ell$ and therefore $s_{\ell-1} \geq s_\ell$. By assumption, we have

$$\tilde{a}_\mu + \tilde{s}_\mu = \hat{a},$$

and since $\tilde{a}_{\mu+1} + \tilde{s}_{\mu+1} = \hat{a}$ by Lemma 3.6, we get

$$c_{\mu+1} = s_\mu = 0.$$

Hence, (2.21) defines a local strangeness index $\mu(t) \leq \mu$ with

$$\tilde{a}_{\mu(t)} + \tilde{s}_{\mu(t)} = a_{\mu(t)}(t) = \hat{a}.$$

The claim follows because of $s_{\mu(t)}(t) = u_{\mu(t)}(t) = 0$ together with $d_i = r_i - s_i$ and $u_i = n - r_i - a_i - s_i$. ☐

We finish this section with a remark concerning the so-called *perturbation index*, first introduced by [7], see also [8].

REMARK 3.10. According to [11], problem (2.22) with initial condition $x(t_0) = 0$ can be written in operator form as

$$Dx = \hat{f} \tag{3.17}$$

with

$$D: \mathcal{X} \to \mathcal{Y}, \quad Dx(t) = \hat{E}(t)\dot{x}(t) - \hat{A}(t)x(t).$$

In the notation of (2.23), the Banach spaces $\mathcal{X}$ and $\mathcal{Y}$ are given by

$$\begin{aligned} &\mathcal{X} = \{x \in C(\mathbb{I}, \mathbb{C}^n) \mid \hat{E}^+\hat{E}x \in C^1(\mathbb{I}, \mathbb{C}^n),\ \hat{E}^+\hat{E}x(t_0) = 0\}, \\ &\mathcal{Y} = C(\mathbb{I}, \mathbb{C}^n) \end{aligned}$$

equipped with the norms

$$\|x\|_{\mathcal{X}} = \|x\|_{\mathcal{Y}} + \|\frac{d}{dt}(\hat{E}^+\hat{E}x)\|_{\mathcal{Y}}, \quad \|f\|_{\mathcal{Y}} = \max_{t\in\mathbb{I}} \|f(t)\|_\infty.$$

Note that homogeneous initial conditions can be obtained, without loss of generality, by replacing $x(t)$ with $x(t) - x_0$. The operator $\hat{E}^+\hat{E}$ is defined pointwise by $\hat{E}^+\hat{E}x(t) = \hat{E}(t)^+\hat{E}(t)x(t)$ where $\hat{E}(t)^+$ denotes the Moore-Penrose pseudoinverse of $\hat{E}(t)$.

The results of [11] in particular show that D has a continuous inverse in the context of the present paper. Let now $x \in \mathcal{X}$ be a solution of (1.1) with $x(t_0) = 0$ and let $\hat{x} \in C(\mathbb{I}, \mathbb{C}^n)$ be a function such that

$$E(t)\dot{\hat{x}}(t) - A(t)\hat{x}(t) - f(t) = \delta(t), \quad \hat{x}(t_0) = \hat{x}_0$$

with some defect $\delta \in \mathcal{Y}$. Shifting to a homogeneous initial condition, as above, yields

$$E(t)(\dot{\hat{x}}(t) - \dot{\hat{x}}_0) - A(t)(\hat{x}(t) - \hat{x}_0) - (f(t) + A(t)\hat{x}_0) = \delta(t), \quad \hat{x}(t_0) - \hat{x}_0 = 0.$$

Using Hypothesis 2.7 now gives

$$\hat{E}(t)(\dot{\hat{x}}(t) - \dot{\hat{x}}_0) - \hat{A}(t)(\hat{x}(t) - \hat{x}_0) - (\hat{f}(t) + \hat{A}(t)\hat{x}_0) = \hat{\delta}(t), \quad \hat{x}(t_0) - \hat{x}_0 = 0.$$

with $\hat{\delta} = Z^*(\delta^*, \dot{\delta}^*, \ldots, \delta^{(\hat{\mu})*})^*$ according to (2.23), or (with all composed functions defined pointwise)

$$D(\hat{x} - \hat{x}_0) = \hat{f} + \hat{\delta} + \hat{A}\hat{x}_0 \tag{3.18}$$

such that it is reasonable to require $\hat{x} - \hat{x}_0 \in \mathcal{X}$ and $\hat{f} + \hat{\delta} + \hat{A}\hat{x}_0 \in \mathcal{Y}$. Recalling that $x_0 = 0$, we then obtain

$$\|(\hat{x} - x) - (\hat{x}_0 - x_0)\|_{\mathcal{X}} = \|D^{-1}(\hat{\delta} + \hat{A}(\hat{x}_0 - x_0))\|_{\mathcal{X}} \leq \hat{C}(\|\hat{x}_0 - x_0\|_\infty + \|\hat{\delta}\|_{\mathcal{Y}}).$$

This implies

$$\big|\,\|\hat{x}-x\|_{\mathcal{X}}-\|\hat{x}_0-x_0\|_{\mathcal{X}}\,\big|\le \hat{C}(\|\hat{x}_0-x_0\|_\infty+\|\hat{\delta}\|_{\mathcal{Y}})$$

or

$$\|\hat{x}-x\|_{\mathcal{X}}\le \tilde{C}(\|\hat{x}_0-x_0\|_\infty+\|\hat{\delta}\|_{\mathcal{Y}})$$

with positive constants $\hat{C}$ and $\tilde{C}$. Using the definition of $\hat{\delta}$ we finally get the estimate

$$\|\hat{x}-x\|_{\mathcal{X}}\le C(\|\hat{x}_0-x_0\|_\infty+\|\delta\|_{\mathcal{Y}}+\|\dot{\delta}\|_{\mathcal{Y}}+\cdots+\|\delta^{(\hat{\mu})}\|_{\mathcal{Y}}). \tag{3.19}$$

Omitting the trivial case $\nu=0$, the perturbation index is defined to be the smallest number $\hat{\mu}+1$ such that this estimate holds for all $\hat{x}-\hat{x}_0$ in a neighborhood of $x-x_0$. Since the minimal choice yields $\hat{\mu}=\nu-1$ (see Corollary 3.8), the perturbation index equals the differentiation index ν. To include the trivial case $\nu=0$, the definition of the perturbation index needs an extension of some integral form. For details, we refer to [8]. Working with Hypothesis 2.7 and the quantity $\hat{\mu}$, such an extension is not necessary. In particular, we can formulate the above result as follows. Provided $\hat{\mu}$ is well-defined and chosen minimally, it is the smallest number such that (3.19) holds with some (positive) constant C for all $\hat{x}-\hat{x}_0$ in a neighborhood of $x-x_0$ (with respect to the topology of $\mathcal{X}$).

4. Existence and uniqueness. In this section, we develop an existence and uniqueness theorem for linear DAE's satisfying Hypothesis 2.7. Provided that the problem is sufficiently smooth, Theorem 3.7 and Corollary 3.8 state that Hypothesis 2.7 is equivalent to requiring that the differentiation index ν is well-defined. Thus, we would be in the situation of Theorem 2.11 which is due to [1, 3]. But note that Hypothesis 2.7 only uses $M_{\nu-1}$ instead of M_ν. So there is a difference in the smoothness requirements which will turn out to be even larger when dealing with an existence and uniqueness result. We therefore give an alternative approach to the results of [3].

To begin with, we observe that, under Hypothesis 2.7, every solution of (1.1) is also a solution of (2.22), since (1.1) implies (2.22). The problem is to prove that the reverse implication is valid.

In the notation of (2.23), the key result that we will show is that there exists a smooth pointwise nonsingular matrix function R such that on every subinterval $\mathbb{I}_j$

$$R\begin{bmatrix} Z^* \\ \dot{Z}^*+Z^*S^* \end{bmatrix}=\begin{bmatrix} I_n & 0 \\ 0 & H \end{bmatrix}, \tag{4.1}$$

i.e., that the above matrix function is *smoothly 1-full* on $\mathbb{I}$. Here S is again the block up-shift matrix. We first show that this property is invariant under global equivalence transformations.

LEMMA 4.1. *Let (E,A) and $(\tilde{E},\tilde{A})$ be globally equivalent and let Hypothesis 2.7 hold with $\hat{\mu}$, $\hat{a}$, and $\hat{d}$. Let $(M_{\hat{\mu}},N_{\hat{\mu}})$ and $(\tilde{M}_{\hat{\mu}},\tilde{N}_{\hat{\mu}})$ be the associated inflated matrices and let $Z=(Z_1,Z_2)$ with $Z_1^*=[Z_{10}^*\,0\cdots 0]^*$ and $T=(T_1,T_2)$ be given such that*

$$\begin{array}{ll} Z_2^*M_{\hat{\mu}}=0, & \operatorname{rank} Z_2=\hat{a},\\ Z_2^*N_{\hat{\mu}}[I_n\;0\;\cdots\;0]^*T_2=0, & \operatorname{rank} T_2=\hat{d},\\ & \operatorname{rank} Z_{10}^*ET_2=\hat{d}. \end{array} \tag{4.2}$$

Let $\tilde Z = (\tilde Z_1, \tilde Z_2)$, $\tilde T = (\tilde T_1, \tilde T_2)$ be the correponding subspaces associated to $(\tilde M_{\hat\mu}, \tilde N_{\hat\mu})$. If

$$\begin{bmatrix} \tilde Z^* \\ \dot{\tilde Z}^* + \tilde Z^* S^* \end{bmatrix} \tag{4.3}$$

is smoothly 1-full, then also

$$\begin{bmatrix} Z^* \\ \dot Z^* + Z^* S^* \end{bmatrix} \tag{4.4}$$

is smoothly 1-full.

Proof. According to (2.19), we have (omitting subscripts)

$$\tilde M = \Pi M \Theta, \quad \tilde N = \Pi N \Theta - \Pi M \Psi.$$

From

$$\tilde Z_2^* \tilde M = \tilde Z_2^* \Pi M \Theta = 0,$$

it follows that

$$Z_2^* = V_2^* \tilde Z_2^* \Pi$$

with some pointwise nonsingular V_2. Since N has only nonvanishing entries in the first block column, we have

$$\begin{aligned} \tilde Z_2^* \tilde N [I_n\, 0\, \cdots\, 0]^* \tilde T_2 &= \tilde Z_2^* \Pi N \Theta [I_n\, 0\, \cdots\, 0]^* \tilde T_2 \\ &= \tilde Z_2^* \Pi N [Q\, 2\dot Q\, \cdots\, (\hat\mu+1) Q^{(\hat\mu)}]^* \tilde T_2 \\ &= \tilde Z_2^* \Pi N [I_n\, 0\, \cdots\, 0]^* Q \tilde T_2 = 0. \end{aligned}$$

This implies that

$$T_2 = Q \tilde T_2 W_2$$

for some pointwise nonsingular W_2. Now from

$$\operatorname{rank} \tilde Z_{10}^* \tilde E \tilde T_2 = \operatorname{rank} \tilde Z_{10}^* P E Q \tilde T_2 = \operatorname{rank} \tilde Z_{10}^* P E T_2 W_2^{-1} = \hat d,$$

we obtain

$$Z_{10}^* = V_1^* \tilde Z_{10}^* P$$

or

$$Z_1^* = V_1^* \tilde Z_1^* \Pi$$

for some pointwise nonsingular V_1. Hence,

$$Z^* = V^* \tilde Z^* \Pi$$

for some pointwise nonsingular V. Applying row operations we get

$$\begin{aligned} \begin{bmatrix} Z^* \\ \dot Z^* + Z^* S^* \end{bmatrix} &= \begin{bmatrix} V^* \tilde Z^* \Pi \\ \dot V^* \tilde Z^* \Pi + V^* \dot{\tilde Z}^* \Pi + V^* \tilde Z^* \dot\Pi + V^* \tilde Z^* \Pi S^* \end{bmatrix} \rightarrow \\ &\rightarrow \begin{bmatrix} \tilde Z^* \Pi \\ \dot{\tilde Z}^* \Pi + \tilde Z^* \dot \Pi + \tilde Z^* \Pi S^* \end{bmatrix}. \end{aligned}$$

ETNA
Kent State University
etna@mcs.kent.edu

Since

$$\begin{aligned}(\dot\Pi + \Pi S^*)_{i,j} &= \dot\Pi_{i,j} + \Pi_{i,j-1} = \tbinom{i}{j}P^{(i-j+1)} + \tbinom{i}{j-1}P^{(i-j+1)} \\ &= \tbinom{i+1}{j}P^{(i-j+1)} = \Pi_{i+1,j} = (S^*\Pi)_{i,j},\end{aligned}$$

the relation $\dot\Pi + \Pi S^* = S^*\Pi$ holds and with (4.1) we conclude that

$$\begin{aligned}\begin{bmatrix} Z^* \\ \dot Z^* + Z^*S^* \end{bmatrix} &\to \begin{bmatrix} \tilde Z^*\Pi \\ \dot{\tilde Z}^*\Pi + \tilde Z^* S^*\Pi \end{bmatrix} = \begin{bmatrix} \tilde Z^* \\ \dot{\tilde Z}^* + \tilde Z^* S^* \end{bmatrix}\Pi \\ &= \tilde R^{-1}\begin{bmatrix} I_n & 0 \\ 0 & \tilde H \end{bmatrix}\begin{bmatrix} P & 0 \\ * & * \end{bmatrix} \to \begin{bmatrix} P & 0 \\ * & * \end{bmatrix} \to \begin{bmatrix} I_n & 0 \\ 0 & H \end{bmatrix}.\end{aligned}$$

□

Thus, we may assume that (E, A) is in the normal form (3.6) when working on $\mathbb{I}_j$.

LEMMA 4.2. *Let Hypothesis 2.7 hold for (E, A). Then Z, as given in Lemma 4.1, is smoothly 1-full.*

Proof. Using Lemma 4.1 we may assume without loss of generality that our problem is in the normal form (3.6). Due to the previous computations, we can choose Z, in the notation of (3.9), as

$$Z^* = \left(\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & Z_1^* \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & Z_2^* \end{bmatrix}, \ldots \right).$$

Due to its special structure, it is sufficient to show the claim for the subproblem (3.7), i.e., to look at

$$Z^* = [I\ Z_1^*\ Z_2^*\ \cdots].$$

Using row transformations we obtain

$$\begin{aligned}\begin{bmatrix} Z^* \\ \dot Z^* + Z^*S^* \end{bmatrix} &= \begin{bmatrix} I & Z_1^* & Z_2^* & \cdots \\ 0 & I + \dot Z_1^* & Z_1^* + \dot Z_2^* & \cdots \end{bmatrix} \to \\ &\to \begin{bmatrix} I & Z_1^* & Z_2^* & \cdots \\ 0 & I & (I+\dot Z_1^*)^{-1}(Z_1^* + \dot Z_2^*) & \cdots \end{bmatrix} \to \\ &\to \begin{bmatrix} I & 0 & Z_2^* - Z_1^*(I+\dot Z_1^*)^{-1}(Z_1^* + \dot Z_2^*) & \cdots \\ 0 & I & * & \cdots \end{bmatrix},\end{aligned}$$

where the invertibility of $I + \dot Z_1^*$ follows, since $\dot Z_1$ is nilpotent. Thus, it suffices to show that

$$Z_j^* = Z_1^*(I + \dot Z_1^*)^{-1}(Z_{j-1}^* + \dot Z_j^*) \text{ for } j \ge 2.$$

Working again with infinite matrices, we first use (3.9) in the form

$$[Z_1^*\ Z_2^*\ \cdots] = [G\,0\ \cdots](I - X)^{-1}$$

and

$$[\dot Z_1^*\ \dot Z_2^*\ \cdots] = [\dot G\,0\ \cdots](I - X)^{-1} + [G\,0\ \cdots](I - X)^{-1}\dot X(I - X)^{-1},$$

where

$$X = \begin{bmatrix} \dot{G} & G & & \\ \ddot{G} & 2\dot{G} & G & \\ \vdots & \ddots & \ddots & \ddots \\ \vdots & & \ddots & \ddots \end{bmatrix}.$$

We then find that

$$S^*X = \dot{X} + XS^*.$$

Subtracting S^* on both sides we obtain

$$(I-X)^{-1}S^* = S^*(I-X)^{-1} - (I-X)^{-1}\dot{X}(I-X)^{-1}.$$

With this, we have

$$\begin{aligned}
[Z_1^*\, Z_2^*\, \cdots] &= [G\,0\,\cdots](I-X)^{-1} = [G\,0\,\cdots]\sum_{k\geq 0} X^k \\
&= [G\,0\,\cdots] + [G\,0\,\cdots]\sum_{k\geq 1} X^k = [G\,0\,\cdots] + [G\dot{G}\,G^2\,0\,\cdots]\sum_{k\geq 0} X^k \\
&= [G\,0\,\cdots] + G[\dot{G}\,0\,\cdots](I-X)^{-1} + G[G\,0\,\cdots]S^*(I-X)^{-1} \\
&= [G\,0\,\cdots] + G[\dot{Z}_1^*\,\dot{Z}_2^*\,\cdots] + G[0\,Z_1^*\,Z_2^*\,\cdots].
\end{aligned}$$

From the first entry, we obtain

$$Z_1^* = G + G\dot{Z}_1^*$$

or

$$Z_1^*(I+\dot{Z}_1^*)^{-1} = G,$$

and from the other entries, we get

$$Z_j^* = G(Z_{j-1}^* + \dot{Z}_j^*) \text{ for } j \geq 2.$$

This proves our claim. □

Now we are in the situation to give our new existence and uniqueness theorem.

THEOREM 4.3. *Let (E,A) satisfy Hypothesis 2.7 with $\hat{\mu}$ and $\hat{a}$. In particular, let $E, A \in C^{\hat{\mu}+1}(\mathbb{I}, \mathbb{C}^{n,n})$ and $f \in C^{\hat{\mu}+1}(\mathbb{I}, \mathbb{C}^n)$. Then, the following holds.*

1. *An initial condition (1.2) is consistent if and only if (1.2) implies the $\hat{a}$ conditions*

$$\hat{A}_2(t_0)x_0 + \hat{f}_2(t_0) = 0. \tag{4.5}$$

2. *Every initial value problem with consistent initial condition has a unique solution.*

Proof. Each solution of (1.1) must satisfy (2.22). For the reverse direction, let x be a solution of (2.22). Differentiating once yields the inflated system

$$\begin{bmatrix} \hat{E} & 0 \\ \dot{\hat{E}} - \hat{A} & \hat{E} \end{bmatrix} \begin{pmatrix} \dot{x} \\ \ddot{x} \end{pmatrix} = \begin{bmatrix} \hat{A} & 0 \\ \dot{\hat{A}} & 0 \end{bmatrix} \begin{pmatrix} x \\ \dot{x} \end{pmatrix} + \begin{pmatrix} \hat{f} \\ \dot{\hat{f}} \end{pmatrix},$$

ETNA
Kent State University
etna@mcs.kent.edu

in the notation of (2.23). Inserting (2.23) gives (using the abbreviation $V = [I\, 0 \cdots]^*$)

$$\begin{bmatrix} Z^* M_{\hat\mu} V & 0 \\ \dot Z^* M_{\hat\mu} V + Z^* \dot M_{\hat\mu} V - Z^* N_{\hat\mu} V & Z^* M_{\hat\mu} V \end{bmatrix} \begin{pmatrix} \dot x \\ \ddot x \end{pmatrix} = \begin{bmatrix} Z^* N_{\hat\mu} V & 0 \\ \dot Z^* N_{\hat\mu} V + Z^* \dot N_{\hat\mu} V & 0 \end{bmatrix} \begin{pmatrix} x \\ \dot x \end{pmatrix} + \begin{pmatrix} Z^* g_{\hat\mu} \\ \dot Z^* g_{\hat\mu} + Z^* \dot g_{\hat\mu} \end{pmatrix}.$$

From (2.16) we have

$$\dot M_{\hat\mu} V = S^* M_{\hat\mu} V + N_{\hat\mu} V, \quad \dot N_{\hat\mu} V = S^* N_{\hat\mu} V, \quad \dot g_{\hat\mu} = S^* g_{\hat\mu},$$

and thus we obtain

$$\begin{bmatrix} Z^* M_{\hat\mu} V & 0 \\ \dot Z^* M_{\hat\mu} V + Z^* S^* M_{\hat\mu} V & Z^* M_{\hat\mu} V \end{bmatrix} \begin{pmatrix} \dot x \\ \ddot x \end{pmatrix} = \begin{bmatrix} Z^* N_{\hat\mu} V & 0 \\ \dot Z^* N_{\hat\mu} V + Z^* S^* N_{\hat\mu} V & 0 \end{bmatrix} \begin{pmatrix} x \\ \dot x \end{pmatrix} + \begin{pmatrix} Z^* g_{\hat\mu} \\ \dot Z^* g_{\hat\mu} + Z^* S^* g_{\hat\mu} \end{pmatrix},$$

or

$$\begin{bmatrix} Z^* & 0 \\ \dot Z^* + Z^* S^* & Z^* \end{bmatrix} \begin{bmatrix} M_{\hat\mu} V & 0 \\ 0 & M_{\hat\mu} V \end{bmatrix} \begin{pmatrix} \dot x \\ \ddot x \end{pmatrix} = \begin{bmatrix} Z^* & 0 \\ \dot Z^* + Z^* S^* & 0 \end{bmatrix} \begin{bmatrix} N_{\hat\mu} V & 0 \\ 0 & N_{\hat\mu} V \end{bmatrix} \begin{pmatrix} x \\ \dot x \end{pmatrix} + \begin{bmatrix} Z^* \\ \dot Z^* + Z^* S^* \end{bmatrix} g_{\hat\mu}.$$

Utilizing the 1-fullness of the matrix

$$\begin{bmatrix} Z^* \\ \dot Z^* + Z^* S^* \end{bmatrix}$$

on $\mathbb{I}_j$, we see from Lemma 4.2 that (in the nilpotent part of the normal form)

$$\begin{bmatrix} Z^* & 0 \\ \dot Z^* + Z^* S^* & Z^* \end{bmatrix} \to \begin{bmatrix} Z^* - G(\dot Z^* + Z^* S^*) & -GZ^* \\ \dot Z^* + Z^* S^* & Z^* \end{bmatrix} \to \begin{bmatrix} I & 0 & -GZ^* \\ 0 & H & * \end{bmatrix},$$

and hence

$$R \begin{bmatrix} Z^* & 0 \\ \dot Z^* + Z^* S^* & Z^* \end{bmatrix} \begin{bmatrix} M_{\hat\mu} V & 0 \\ 0 & M_{\hat\mu} V \end{bmatrix} = \begin{bmatrix} I & 0 & -GZ^* \\ 0 & H & * \end{bmatrix} \begin{bmatrix} E & 0 \\ * & 0 \\ 0 & M_{\hat\mu} V \end{bmatrix} = \begin{bmatrix} E & 0 \\ * & * \end{bmatrix}.$$

By Lemma 4.1, this also holds in the general case. Applying now (4.1) to all terms of the above inflated system, we deduce from the first block row

$$E\dot x = Ax + f.$$

Thus x satisfies (1.1) on a dense subset of $\mathbb{I}$ and therefore, by continuity, on the whole interval $\mathbb{I}$. □

To compare this result with that of [1, 3], note that in [1, 3], it is required that the coefficients E, A, f are at least $3n$-times continuously differentiable. Summarizing the results presented in this paper, we need them to be $\hat\mu$-times continuously differentiable with $\hat\mu = \nu - 1$ if $\nu \geq 1$ to obtain $\hat A, \hat E, \hat f$ as defined in (2.23), and these coefficients must be once continuously differentiable to conclude that the solutions of (2.22) also solve the original problem (1.1). Together this sums up to a number of only ν instead of $3n$. Of course, in the trivial case $\nu = 0$ the coefficients need not to be differentiated at all.

To illustrate the various aspects of the obtained results, we give a simple worked-out example.

EXAMPLE 4.4. Let (E, A) belong to the DAE

$$\begin{bmatrix} 0 & t \\ 0 & 0 \end{bmatrix} \dot{x} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x + \begin{bmatrix} f_1(t) \\ f_2(t) \end{bmatrix}, \quad \mathbb{I} = [-1, 1],$$

see, e.g., [15]. Obviously E has a rank drop at $t = 0$ so that μ cannot be well-defined. Nevertheless, we can smoothly transform to the form (2.5) by interchanging the columns when we do not require the identity in the upper left corner. This weaker form still allows for the differentiation and elimination step described for (2.5) and gives the (unique) solution

$$x_1(t) = -(f_1(t) + t\dot{f}_2(t)), \quad x_2(t) = -f_2(t).$$

In this way $t = 0$ seems to be no exceptional point. Rewriting the above DAE by means of the product rule into

$$\frac{d}{dt}(tx_2) - x_2 = x_1 + f_1(t), \quad 0 = x_2 + f_2(t)$$

yields the (unique) solution

$$x_1(t) = f_2(t) - (f_1(t) + \frac{d}{dt}(tf_2(t))), \quad x_2(t) = -f_2(t),$$

which makes sense if tf_2 is continuously differentiable, e.g., for $f_2(t) = |t|$. In this way $t = 0$ remains exceptional and is reflected by changes in the characteristic values, according to

$$\begin{aligned} &r_0 = 1,\, a_0 = 0,\, s_0 = 1, \\ &r_1 = 0,\, a_1 = 2,\, s_1 = 0 \text{ for } t \neq 0, \\ &r_0 = 0,\, a_0 = 2,\, s_0 = 0 \text{ for } t = 0. \end{aligned}$$

This behaviour corresponds to the splitting

$$[-1, 1] = \{-1\} \cup (-1, 0) \cup \{0\} \cup (0, 1) \cup \{1\}$$

of the interval $\mathbb{I}$ along the lines of Corollary 3.2. Examining

$$(M_1(t), N_1(t)) = \left(\left[\begin{array}{cc|cc} 0 & t & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline -1 & 1 & 0 & t \\ 0 & -1 & 0 & 0 \end{array} \right], \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right] \right),$$

we first find a constant corank $\hat{a} = 2$. Choosing

$$Z_2^*(t) = \left[\begin{array}{cc|cc} 1 & 0 & 0 & t \\ 0 & 1 & 0 & 0 \end{array} \right],$$

we then get

$$\hat{A}_2(t) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \hat{f}_2(t) = \begin{bmatrix} f_1(t) + t\dot{f}_2(t) \\ f_2(t) \end{bmatrix}.$$

Consequently (E, A) satisfies Hypothesis 2.7 with $\hat{\mu} = 1$, $\hat{a} = 2$, and $\hat{d} = 0$, the latter meaning that $\hat{E}_1(t) = \emptyset$, $\hat{A}_1(t) = \emptyset$. This also tells us that the differentiation index is well-defined with $\nu = 2$. In addition,

$$\left[\begin{array}{c} Z_2^*(t) \\ \dot{Z}_2^*(t) + Z_2^*(t)S^* \end{array}\right] = \left[\begin{array}{cc|cc|cc} 1 & 0 & 0 & t & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 & t \\ 0 & 0 & 0 & 1 & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{cc|cc|cc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 & t \\ 0 & 0 & 0 & 1 & 0 & 0 \end{array}\right]$$

is indeed smoothly 1-full. Finally we have

$$\mu(t) = \begin{cases} 0 \text{ for } t = 0, \\ 1 \text{ for } t \neq 0, \end{cases} \qquad a_{\mu(t)}(t) = 2, \quad d_{\mu(t)}(t) = 0,$$

where $\mu(0) = 0$ indicates that $Z_2^*(0)$ has vanishing last block column.

We conclude this section with a remark on the implications of the presented results for a numerical method.

REMARK 4.5. Hypothesis 2.7 was suggested by the numerical procedure given in [10], where the starting point was to require that the local invariants are also global, i.e., that the strangeness index μ is well-defined on the whole interval $\mathbb{I}$. In this case, the algorithm can work without information from the user on the properties of the problem. It itself can compute the relevant structural information. In the presence of structural rank changes, however, the numerical method must or should terminate. The disadvantage of such a procedure of course is that it may terminate, although the problem to be solved behaves well, as in Example 4.4. To avoid this, we can take the other extremum and supply the global information ν of the differentiation index, or equivalently $\hat{\mu}$ of Hypothesis 2.7. Of course, this requires a priori knowledge from the user. Given $\hat{\mu}$, the quantities $\hat{a}$ and $\hat{d}$ of Hypothesis 2.7 are still local such that this part can be done by the algorithm. Again the program must or should terminate if they change. But due to Theorem 3.9 there is a compromise. It does not seem necessary to terminate the first kind of algorithm when the strangeness index $\mu(t)$ changes provided $a_{\mu(t)}(t)$ and $d_{\mu(t)}(t)$ do not change. An algorithm with this strategy implemented would not terminate in Example 4.4 even if it would hit $t = 0$ exactly. With this altered termination criterion, the procedure of [10] and the production code of [12] based on that approach will integrate all systems that satisfy Hypothesis 2.7. Recall that these are, up to smoothness requirements, those systems for which the differentiation index is well-defined. In contrast to the algorithms developed in [2], this approach does not exhibit drift-off, since it uses the information given by $\hat{a}$ and $\hat{d}$ such that all constraints are included into the system (2.22) that is integrated. In this sense, it represents a method which works for all problems that are well-posed (i.e., for all problems that have a unique solution, depending smoothly on consistent initial values) and it furthermore obeys all invariants that are present.

5. Conclusion. In this paper we have examined the relationship between the differentiation index as defined in [1, 3] and the strangeness index as defined in [9]. We have given a new existence and uniqueness result for linear DAE's with variable coefficients that has weaker smoothness assumptions than the result in [1, 3] and weaker constant rank assumptions than the result in [9]. Based on these new results one can easily modify the numerical procedure developed in [10] and the production code of [12] to be applicable to all well-posed problems that have a unique solution depending smoothly on consistent initial values.

REFERENCES

[1] K. E. Brenan, S. L. Campbell, and L. R. Petzold, *Numerical Solution of Initial-Value Problems in Differential Algebraic Equations*, Elsevier, North Holland, New York, 1989.

[2] S. L. Campbell, *The numerical solution of higher index linear time varying singular systems of differential equations*, SIAM J. Sci. Statist. Comput., 6 (1985), pp. 334–348.

[3] ———, *A general form for solvable linear time varying singular systems of differential equations*, SIAM J. Math. Anal., 18 (1987), pp. 1101–1115.

[4] S. L. Campbell and C. D. Meyer, *Generalized Inverses of Linear Transformations*, Pitman, London, 1979.

[5] E. Griepentrog and R. März, *Differential-Algebraic Equations and Their Numerical Treatment*, Teubner Verlag, Leipzig, 1986.

[6] ———, *Basic properties of some differential algebraic equations*, Z. Anal. Anwendungen., 8 (1989), pp. 25–40.

[7] E. Hairer, C. Lubich, and M. Roche, *The Numerical Solution of Differential-Algebraic Systems by Runge-Kutta Methods*, Springer Verlag, Berlin, 1989.

[8] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations II*, Springer-Verlag, Berlin, 1991.

[9] P. Kunkel and V. Mehrmann, *Canonical forms for linear differential-algebraic equations with variable coefficients*, J. Comput. Appl. Math., 56 (1994), pp. 225–251.

[10] ———, *A new class of discretization methods for the solution of linear differential-algebraic equations*, SIAM J. Numer. Anal., 33 (1996), pp. 1941–1961.

[11] ———, *Generalized inverses of differential-algebraic operators*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 426–442.

[12] P. Kunkel, V. Mehrmann, W. Rath, and J. Weickert, *A new software package for linear differential-algebraic equations*, SIAM J. Sci. Comput., 18 (1997), pp. 115–138.

[13] R. März, *Some new results concerning index-3 differential-algebraic equations*, J. Math. Anal. Appl., 140 (1989), pp. 177–199.

[14] L. R. Petzold, *A description of DASSL: A differential/algebraic systems solver*, in Scientific Computing, R. S. Stepleman et al., eds., North-Holland, Amsterdam, 1983, pp. 65–68.

[15] P. J. Rabier and W. C. Rheinboldt, *Classical and generalized solutions of time-dependent linear differential algebraic equations*, Linear Algebra Appl., 245 (1996), pp. 259–293.