

Solution Concepts for Linear DAEs: A Survey

Stephan Trenn

Abstract This survey aims at giving a comprehensive overview of the solution theory of linear differential-algebraic equations (DAEs). For classical solutions a complete solution characterization is presented including explicit solution formulas similar to the ones known for linear ordinary differential equations (ODEs). The problem of inconsistent initial values is treated and different approaches are discussed. In particular, the common Laplace-transform approach is discussed in the light of more recent distributional solution frameworks.

Keywords Differential algebraic equations · Descriptor systems · Distributional solution theory · Laplace transform

Mathematics Subject Classification (2010) 34A09 · 34A12 · 34A05 · 34A25

1 Introduction

Modeling physical phenomena relates physical variables via differential equations as well as algebraic equations leading in general to a system description of the form

$$F(t, \dot{x}, x) = 0,$$

a *differential-algebraic equation* (DAE). However, this survey will not treat this most general system description but it will consider its linear counterpart

$$E\dot{x} = Ax + f, \tag{1.1}$$

where $E, A \in \mathbb{R}^{m \times n}$, $m, n \in \mathbb{N}$, are constant matrices and $f : \mathbb{R} \rightarrow \mathbb{R}^m$ is some inhomogeneity. If the matrix E is square and invertible, the DAE is equivalent to an ordinary differential equation (ODE) of the form

$$\dot{x} = Ax + f. \tag{1.2}$$

S. Trenn (✉)
University of Kaiserslautern, 67663 Kaiserslautern, Germany
e-mail: trenn@mathematik.uni-kl.de

For this ODE the solution theory is well understood and there have been no disputes or different viewpoints on it in the last five or more decades. In fact, the solution formula can concisely be expressed with the matrix exponential:

$$x(t) = e^{At} x_0 + \int_0^t e^{A(t-\tau)} f(\tau) d\tau, \quad x_0 \in \mathbb{R}^n; \quad (1.3)$$

although the Jordan canonical form of A is essential to grasp the whole of the possibilities of solution behaviors. Some features of the solutions of an ODE are highlighted:

Existence. For every initial condition $x(0) = x_0$, $x_0 \in \mathbb{R}^n$, and each (locally integrable) inhomogeneity f there exists a solution.

Uniqueness. For any fixed inhomogeneity f the initial value $x(0)$ uniquely determines the whole solution; in fact each single value $x(t)$, $t \in \mathbb{R}$, determines the solution on the whole time axis.

Inhomogeneity. The solution is always one degree “smoother” than the inhomogeneity, i.e. if f is differentiable then x is at least twice differentiable, in particular, non-smoothness of f does not prevent the ODE of having a solution (at least in the sense of Carathéodory).

In Sects. 2.4 and 2.5 solution formulas similar to (1.3) will be presented for *regular* DAEs; however, for general DAEs none of these three properties have to hold anymore as the following example shows.

Example 1.1 Consider the DAE

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \dot{x} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} x + f$$

which implies $x_2 = -f_2$, $x_1 = \dot{x}_2 - f_1 = -\dot{f}_2 - f_1$ and $f_3 = 0$. In particular, not for all initial values or all inhomogeneities there exists a solution. Furthermore, x_3 is not restricted at all, hence uniqueness of solutions is not present. Finally, x_1 contains the derivative of the inhomogeneity so that the solution is “less smooth” than the inhomogeneity which could lead to non-existence of solutions if the inhomogeneities is not sufficiently smooth.

The aim of this survey is twofold: (1) to present a fairly complete classical solution theory for the DAE (1.1) also for the singular case; (2) to discuss the approaches to treat inconsistent initial values and the corresponding distributional solution concepts. In particular, a rigorous discussion of the so-called Laplace-transform approach to treat inconsistent initial values and its connection to distributional solution concepts is carried out. This is a major difference with the already available survey by Lewis [32], which is not so much concerned with distributional solutions. The focus of Lewis’ survey is more on system theoretic topics like controllability, observability, stability and feedback, which are not treated here.

This survey is structured as follows. In Sect. 2 classical (i.e. differentiable) solutions of (1.1) are studied. It is shown how the Weierstraß and Kronecker canonical form of the matrix pencil $sE - A \in \mathbb{R}^{m \times n}[s]$ can be used to fully characterize the solutions. Solution formulas which do not need the complete knowledge of the canonical forms will be presented, too. A short overview over the situation for time-varying DAEs is given as well. Inconsistent initial values are the most discussed topics concerning DAEs and different arguments how to treat them have been proposed. One common approach to treat inconsistent values is the application of the Laplace transform to (1.1); the details are explained in Sect. 4. However, the latter approach led to much confusion and therefore a time-domain approach based on distributional solutions was developed and studied by a number of authors, see Sect. 5.

2 Classical Solutions

In this section classical solutions of the DAE (1.1) are considered:

Definition 2.1 (Classical solution) A classical solution of the DAE (1.1) is any differential function $x \in \mathcal{C}^1(\mathbb{R} \rightarrow \mathbb{R}^n)$ such that $E\dot{x}(t) = Ax(t) + f(t)$ holds for all $t \in \mathbb{R}$.

It will turn out that existence of a classical solution in general also depends on the smoothness properties of the inhomogeneity; if not mentioned otherwise it will be assumed therefore in the following that the inhomogeneity f is sufficiently smooth, e.g. by assuming that f is in fact smooth (i.e. arbitrarily often differentiable).

2.1 The Kronecker and Weierstraß Canonical Forms

The first appearance of DAEs (1.1) with a complete solution discussion seems to be the one by Gantmacher [21] (Russian original 1953), where he considered classical solutions. His analysis is based on the following notion of equivalence of matrix pairs (called strict equivalence by him):

$$(E_1, A_1) \cong (E_2, A_2) \\ \Leftrightarrow \exists S \in \mathbb{R}^{m \times m}, T \in \mathbb{R}^{n \times n} \text{ both invertible: } (E_1, A_1) = (SE_2T, SA_2T).$$

It is clear that for equivalent matrix pairs (E_1, A_1) and (E_2, A_2) (via the transformation matrices S and T) the following equivalence holds:

$$x \text{ solves } E_1\dot{x} = A_1x + f \quad \Leftrightarrow \quad z = T^{-1}x \text{ solves } E_2\dot{z} = Az + S^{-1}f.$$

Gantmacher's focus is actually on matrix pencils $sE - A \in \mathbb{R}^{m \times n}[s]$ and the derivation of a canonical form corresponding to the above equivalence—the *Kronecker canonical form* (KCF). The solution theory of the DAE (1.1) is a mere application of the KCF. In particular, he does not consider inconsistent initial values or non-smooth inhomogeneities. The existence and representation of the KCF is formulated with the following result.

Theorem 2.1 (Kronecker canonical form [21, 28]) *For every matrix pencil $sE - A \in \mathbb{R}^{m \times n}[s]$ there exist invertible matrices $S \in \mathbb{C}^{m \times m}$ and $T \in \mathbb{C}^{n \times n}$ such that, for $a, b, c, d \in \mathbb{N}$ and $\varepsilon_1, \dots, \varepsilon_a, \rho_1, \dots, \rho_b, \sigma_1, \dots, \sigma_c, \eta_1, \dots, \eta_d \in \mathbb{N}$,*

$$S(sE - A)T = \text{diag}(\mathcal{P}_{\varepsilon_1}(s), \dots, \mathcal{P}_{\varepsilon_a}(s), \mathcal{J}_{\rho_1}(s), \dots, \mathcal{J}_{\rho_b}(s), \mathcal{N}_{\sigma_1}(s), \dots, \mathcal{N}_{\sigma_c}(s), \mathcal{Q}_{\eta_1}(s), \dots, \mathcal{Q}_{\eta_d}(s)), \quad (2.1)$$

where

$$\mathcal{P}_\varepsilon(s) = s \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix} \in \mathbb{R}^{\varepsilon \times (\varepsilon+1)}[s], \quad \varepsilon \in \mathbb{N},$$

$$\mathcal{J}_\rho(s) = sI - \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix} \in \mathbb{C}^{\rho \times \rho}[s], \quad \rho \in \mathbb{N}, \lambda \in \mathbb{C},$$

$$\mathcal{N}_\sigma(s) = s \begin{bmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix} - I \in \mathbb{R}^{\sigma \times \sigma}[s], \quad \sigma \in \mathbb{N},$$

$$\mathcal{Q}_\eta(s) = s \begin{bmatrix} 0 & & \\ 1 & \ddots & \\ & \ddots & 0 \\ & & 1 \end{bmatrix} - \begin{bmatrix} 1 & & \\ 0 & \ddots & \\ & \ddots & 1 \\ & & 0 \end{bmatrix} \in \mathbb{R}^{(\eta+1) \times \eta}[s], \quad \eta \in \mathbb{N}.$$

The block-diagonal form (2.1) is unique up to reordering of the blocks and is called *Kronecker canonical form* (KCF) of the matrix pencil $(sE - A)$.

Note that in the KCF $\mathcal{P}_\varepsilon(s)$ -blocks with $\varepsilon = 0$ and $\mathcal{Q}_\eta(s)$ -blocks with $\eta = 0$ are possible, which results in zero columns (for $\varepsilon = 0$) and/or zero rows (for $\eta = 0$) in the KCF, see the following example.

Example 2.1 (KCF of Example 1.1) By a simple interchanging of rows and columns the KCF is obtained from Example 1.1 and has the following form

$$\left[\begin{array}{c|cc} \hline 0 & -1 & 0 \\ \hline 0 & s & -1 \\ 0 & 0 & 0 \\ \hline \end{array} \right]$$

i.e. the KCF consists of one $\mathcal{P}_0(s)$ -block, one $\mathcal{N}_2(s)$ -block and one $\mathcal{Q}_0(s)$ -block.

In the canonical coordinates the solution analysis is now rather straightforward because each block on the diagonal and the associated variables can be considered separately. The four different block types lead to the following solution characterizations:

$\mathcal{P}_\varepsilon(s)$ -block If $\varepsilon = 0$ then this simply means that the corresponding variable does not appear in the equations and is therefore free and can be chosen arbitrarily. For $\varepsilon > 0$ consider the differential equation $\mathcal{P}_\varepsilon(\frac{d}{dt})(x) = f$ which equivalently can be written as the ODE

$$\begin{pmatrix} \dot{x}_2 \\ \dot{x}_3 \\ \vdots \\ \dot{x}_{\varepsilon+1} \end{pmatrix} = \begin{bmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_{\varepsilon+1} \end{pmatrix} + \begin{pmatrix} f_1 \\ f_2 \\ \dots \\ f_\varepsilon \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix} x_1.$$

Hence for any x_1 and any inhomogeneity f there exist solutions for $x_2, x_3, \dots, x_{\varepsilon+1}$ uniquely determined by the initial values $x_2(0), \dots, x_3(0)$. In particular, for all initial values and all inhomogeneities there exist solutions which are not unique because x_1 can freely be chosen.

$\mathcal{I}_\rho(s)$ -block The differential equation $\mathcal{I}_\rho(\frac{d}{dt})(x) = f$ is a standard linear ODE, i.e. it holds that for all initial values and all inhomogeneities a unique solution.

$\mathcal{N}_\rho(s)$ -block Write $\mathcal{N}_\rho(s) = sN - I$, then it is easily seen that the differential operator $\mathcal{N}_\rho(\frac{d}{dt}) : \mathcal{C}^\infty \rightarrow \mathcal{C}^\infty$ is invertible with inverse

$$\left(N \frac{d}{dt} - I \right)^{-1} = - \sum_{i=0}^{\rho-1} N^i \frac{d^i}{dt^i}. \tag{2.2}$$

In particular for any smooth inhomogeneity the solution of the differential equation $\mathcal{N}(\frac{d}{dt})(x) = f$ is uniquely given by

$$x = - \sum_{i=0}^{\rho-1} N^i f^{(i)} = \begin{pmatrix} -f_1 \\ -f_2 - \dot{f}_1 \\ \vdots \\ -f_\rho - \dot{f}_{\rho-1} - \dots - f_1^{(\rho-1)} \end{pmatrix}. \tag{2.3}$$

In particular it is not possible to specify the initial values arbitrarily—they are completely determined by the inhomogeneity.

$\mathcal{Q}_\eta(s)$ -**block** If $\eta = 0$ then no variable is present and the equation reads $0 = f$, hence not for all inhomogeneities the overall DAE is solvable. If $\eta > 0$ then the solution of the differential equation $\mathcal{Q}_\eta(\frac{d}{dt})(x) = f$ is given by (2.3) with ρ replaced by η but only if the inhomogeneity fulfills

$$f_{\eta+1} = \dot{x}_\eta = -\dot{f}_\eta - \ddot{f}_\eta - \dots - f_1^{(\eta)}.$$

In particular not for all inhomogeneities and not for all initial values solutions exist. However, when solutions exist they are uniquely given by (2.3).

A consequence of the above blockwise analysis is the following result.

Corollary 2.2 (Existence and uniqueness of solutions) *The DAE (1.1) has a smooth solution x for all smooth inhomogeneities f if, and only if, in the KCF the $\mathcal{Q}_\eta(s)$ -blocks are not present. Any solution x of (1.1) with fixed inhomogeneity f is uniquely determined by the initial value $x(0)$ if, and only if, in the KCF the $\mathcal{P}_\varepsilon(s)$ -blocks are not present.*

The KCF without the $\mathcal{P}_\varepsilon(s)$ and $\mathcal{Q}_\eta(s)$ blocks is also called the *Weierstraß canonical form* (WCF) and can be characterized directly in terms of the original matrices. For this the notion of regularity is needed.

Definition 2.2 (Regularity) The matrix pencil $sE - A \in \mathbb{R}^{m \times n}[s]$ is called regular if, and only if, $n = m$ and $\det(sE - A)$ is not the zero polynomial. The matrix pair (E, A) and the corresponding DAE (1.1) is called regular whenever $sE - A$ is regular.

Theorem 2.3 (Weierstraß canonical form [49]) *The matrix pencil $sE - A \in \mathbb{R}^{n \times n}[s]$ is regular if, and only if, there exist invertible matrices $S, T \in \mathbb{C}^{n \times n}$ such that $sE - A$ is transformed into the Weierstraß canonical form (WCF)*

$$S(sE - A)T = s \begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} - \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix},$$

where $J \in \mathbb{C}^{n_1 \times n_1}$, $N \in \mathbb{C}^{n_2 \times n_2}$, $n_1 + n_2 = n$, are matrices in Jordan canonical form and N is nilpotent.

In conclusion, if one aims at similar solution properties as for classical linear ODEs the class of regular DAEs is exactly the one to consider, see also Sects. 2.4 and 2.5. In the classical solution framework there is still a gap between ODEs and regular DAEs because (1.1) does not have solutions for all initial values and not for insufficiently smooth inhomogeneities. However, in a distributional solution framework these two missing properties can also be recaptured, see Sect. 5.

2.2 Solution Formulas Based on the Wong Sequences: General Case

For practical problems the above solution characterization is not so useful as the determination of the KCF is numerically ill posed. Therefore, solution formulas which do not need the complete KCF are of interest. One of the first work in this direction is the one by Wilkinson [50], who presents an iterative algorithm to obtain the solutions. More geometrical approaches can be traced back to Dieudonné [15] and Wong [51]; the latter introduced the two important subspace sequences for a matrix pair $(E, A) \in (\mathbb{R}^{m \times n})^2$:

$$\begin{aligned} \mathcal{V}_0 &= \mathbb{R}^n, & \mathcal{V}_{i+1} &= A^{-1}(E\mathcal{V}_i), \quad i = 0, 1, 2, \dots, \\ \mathcal{W}_0 &= \{0\}, & \mathcal{W}_{j+1} &= E^{-1}(A\mathcal{W}_j), \quad j = 0, 1, 2, \dots, \end{aligned} \tag{2.4}$$

which therefore will be called *Wong sequences* in the following. It is easily seen that the Wong sequences are nested and terminate after finitely many steps, i.e.

$$\begin{aligned} \exists i^* \in \{0, 1, \dots, n\}: & \quad \mathcal{V}^* := \bigcap_{i \in \mathbb{N}} \mathcal{V}_i = \mathcal{V}_{i^*}, \\ \exists j^* \in \{0, 1, \dots, n\}: & \quad \mathcal{W}^* := \bigcup_{j \in \mathbb{N}} \mathcal{W}_j = \mathcal{W}_{j^*}. \end{aligned}$$

Bernhard [6] used the first Wong sequence in his geometrical analysis of (1.1) where the inhomogeneity has the special form $f = Bu$ for some suitable matrix B . Utilizing both Wong sequences Armentano [2] was able to obtain a Kronecker like form. However, his arguments are purely geometrical and it is not apparent how to characterize the solutions of (1.1) because the necessary transformation matrices are not given explicitly. This problem was resolved recently in [4], where the following connection between the Wong sequences and a quasi-Kronecker form was established.

Theorem 2.4 (Quasi Kronecker form (QKF) [4]) *Consider the DAE (1.1) and the corresponding limits \mathcal{V}^* and \mathcal{W}^* of the Wong sequences (2.4). Choose any invertible matrices $[P_1, R_1, Q_1] \in \mathbb{R}^{n \times n}$ and $[P_2, R_2, Q_2] \in \mathbb{R}^{m \times m}$ such that*

$$\begin{aligned} \text{im } P_1 &= \mathcal{V}^* \cap \mathcal{W}^*, & \text{im}[P_1, R_1] &= \mathcal{V}^* + \mathcal{W}^*, \\ \text{im } P_2 &= E\mathcal{V}^* \cap A\mathcal{W}^*, & \text{im}[P_2, R_2] &= E\mathcal{V}^* + A\mathcal{W}^*, \end{aligned}$$

then $T = [P_1, R_1, Q_1]$, $S = [P_2, R_2, Q_2]^{-1}$ put the matrix pencil $sE - A$ into quasi-Kronecker triangular form (QKTF):

$$S(sE - A)T = \begin{bmatrix} sE_P - A_P & sE_{PR} - A_{PR} & sE_{PQ} - A_{PQ} \\ 0 & sE_R - A_R & sE_{RQ} - A_{RQ} \\ 0 & 0 & sE_Q - A_Q \end{bmatrix}, \tag{2.5}$$

where $\lambda E_P - A_P$ has full row rank for all $\lambda \in \mathbb{C} \cup \{\infty\}$, $sE_R - A_R$ is regular, and $\lambda E_Q - A_Q$ has full column rank for all $\lambda \in \mathbb{C} \cup \{\infty\}$. Furthermore, the following generalized Sylvester equations are solvable:

$$\begin{aligned} 0 &= E_{RQ} + E_R F_1 + F_2 E_Q, & 0 &= A_{RQ} + A_R F_1 + F_2 A_Q, \\ 0 &= E_{PR} + E_P G_1 + G_2 E_R, & 0 &= A_{PR} + A_P G_1 + G_2 A_R, \\ 0 &= (E_{PQ} + E_{PR} F_1) + E_P H_1 + H_2 E_Q, \\ 0 &= (A_{PQ} + A_{PR} F_1) + A_P H_1 + H_2 A_Q, \end{aligned}$$

and any solutions $F_1, F_2, G_1, G_2, H_1, H_2$ yield a quasi-Kronecker form (QKF) via

$$\begin{aligned} & \begin{bmatrix} I & -G_2 & -H_2 \\ 0 & I & -F_2 \\ 0 & 0 & I \end{bmatrix}^{-1} S(sE - A)T \begin{bmatrix} I & G_1 & H_1 \\ 0 & I & F_1 \\ 0 & 0 & I \end{bmatrix} \\ &= \begin{bmatrix} sE_P - A_P & 0 & 0 \\ 0 & sE_R - A_R & 0 \\ 0 & 0 & sE_Q - A_Q \end{bmatrix}, \end{aligned} \quad (2.6)$$

where the diagonal block entries are the same as in (2.5).

The solution analysis can now be carried out via analyzing the blocks in the QKF (2.6) individually:

- $sE_P - A_P$: Due to the full rank assumption there exists a unimodular¹ matrix $[M_P(s), K_P(s)]$ such that

$$(sE_P - A_P)[M_P(s), K_P(s)] = [I, 0], \quad (2.7)$$

see e.g. [4, Lem. 3.1]. The solutions x_P of the DAE $E_P \dot{x}_P = A_P x_P + f_P$ are given by

$$x_P = M_P \left(\frac{d}{dt} \right) (f_P) + K_P \left(\frac{d}{dt} \right) (u)$$

where $u : \mathbb{R} \rightarrow \mathbb{R}^{n_P - m_P}$ is an arbitrary (sufficiently smooth) function and where $m_P \times n_P$ with $m_P < n_P$ is the size of the matrix pencil $sE_P - A_P$. Furthermore, each initial condition $x_P(0) = x_P^0$ can be achieved by an appropriate choice of u .

- $sE_R - A_R$: The solution behavior for a regular DAE was already discussed at the end of Sect. 2.1, a further discussion is carried out in Sects. 2.4 and 2.5.

¹A polynomial matrix is called unimodular if it is invertible and its inverse is again a polynomial matrix.

- $sE_Q - A_Q$: Analogous to the $sE_P - A_P$ block there exists a unimodular matrix $\begin{bmatrix} M_Q(s) \\ K_Q(s) \end{bmatrix}$ such that

$$\begin{bmatrix} M_Q(s) \\ K_Q(s) \end{bmatrix} (sE_Q - A_Q) = \begin{bmatrix} I \\ 0 \end{bmatrix}. \quad (2.8)$$

Then $E_Q \dot{x}_Q = A_Q x_Q + f_Q$ is solvable if, and only if,

$$K_Q \left(\frac{d}{dt} \right) (f_Q) = 0$$

and the solution is uniquely determined by

$$x_Q = M_Q \left(\frac{d}{dt} \right) (f_Q).$$

In particular, the initial values cannot be specified as they are already fixed by $x_Q(0) = M_Q \left(\frac{d}{dt} \right) (f_Q)(0)$.

In summary, the QKF decouples the corresponding DAE into the underdetermined part (existence but non-uniqueness of solutions), the regular part (existence and uniqueness of solutions) and the overdetermined part (uniqueness of solution but possible non-existence). Furthermore, the above solution characterization can also be carried out directly with the QKTF (2.5), where the analysis for the $sE_Q - A_Q$ block remains unchanged, for the regular block the inhomogeneity f_R is replaced by $f_R + (E_{RQ} \frac{d}{dt} - A_{RQ})(x_Q)$ and for the $sE_P - A_P$ block the inhomogeneity f_P is replaced by $f_P + (E_{PR} \frac{d}{dt} - A_{PR})(x_R) + (E_{PQ} \frac{d}{dt} - A_{PQ})(x_Q)$.

Remark 2.1 (Refinement of QKF [3]) If R_1 and R_2 in Theorem 2.4 are chosen in the special way $R_1 = [R_1^J, R_1^N]$ and $R_2 = [R_2^J, R_2^N]$ where

$$\text{im}[P_1, R_1^J] = \mathcal{V}^*, \quad \text{im}[P_2, R_2^J] = E\mathcal{V}^*,$$

then a decoupling of the regular part in (2.5) corresponding the WCF is obtained as well. In particular, applying the Wong sequences again to the regular part (see next section) is not necessary for a further analysis.

2.3 Existence and Uniqueness of Solutions with Respect to In- and Outputs

In practical application the inhomogeneity f in the DAE (1.1) is often generated by a lower dimensional input u , i.e. $f = Bu$ for some suitable matrix B ; furthermore, an output $y = Cx + Du$ is introduced to represent the signals of the systems which are available for measurement and/or are of interest. The resulting DAE is then

often called *descriptor system* [52] (other common names are singular systems [8] or generalized state-space system [48])

$$\begin{aligned} E\dot{x} &= Ax + Bu, \\ y &= Cx + Du. \end{aligned} \tag{2.9}$$

Clearly, a solution theory for general DAEs (1.1) is also applicable to descriptor systems (2.9). In particular, regularity of the matrix pair (E, A) guarantees existence and uniqueness of solutions for any sufficiently smooth input. However, existence and uniqueness of solutions with respect to the input and output might hold for descriptor systems even when the matrix pair (E, A) is not regular as the following example shows.

Example 2.2 Consider the following descriptor system:

$$\begin{aligned} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \dot{x} &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u, \\ y &= \begin{bmatrix} 1 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \end{bmatrix} u, \end{aligned}$$

which has for any input u the unique output $y = -u$. However, the corresponding matrix pair $(E, A) = \left(\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right)$ is not regular.

It is therefore useful to define the notion of *external regularity*.

Definition 2.3 (External regularity) The descriptor system (2.9) and the corresponding matrix tuple (E, A, B, C, D) are called *externally regular* if, only if, for all sufficiently smooth inputs u there exist (classical) solutions x of (2.9) and the output y is uniquely determined by u and $x(0)$.

With the help of the quasi-Kronecker form it is now possible to prove the following characterization of external regularity.

Theorem 2.5 (Characterization of external regularity) *The descriptor system (2.9) is externally regular if, and only if,*

$$\text{rk}[sE - A, B] = \text{rk}[sE - A] = \text{rk} \begin{bmatrix} sE - A \\ C \end{bmatrix} \tag{2.10}$$

for infinitely many $s \in \mathbb{C}$.

Proof The rank of a matrix does not change when multiplied with invertible matrices (from the left and the right), hence it can be assumed that the matrix pair (E, A) is already in QKF (2.6) with corresponding transformation matrices S and T . According to the block size in (2.6) let $SB = [B_P^\top, B_R^\top, B_Q^\top]^\top$ and $CT =$

$[C_P, C_R, C_Q]$. Then (2.10) is equivalent to

$$\text{rk}[sE_Q - A_Q, B_Q] = \text{rk}[sE_Q - A_Q] \quad \text{and} \quad \text{rk} \begin{bmatrix} sE_P - A_P \\ C_P \end{bmatrix} = \text{rk}[sE_P - A_P]$$

for infinitely many $s \in \mathbb{C}$. The rank is also invariant under multiplication with unimodular polynomial matrices, hence (2.10) is also equivalent to, invoking (2.8) and (2.7),

$$\text{rk} \begin{bmatrix} I & M_Q(s)B_Q \\ 0 & K_Q(s)B_Q \end{bmatrix} = \text{rk} \begin{bmatrix} I \\ 0 \end{bmatrix} \quad \text{and} \quad \text{rk} \begin{bmatrix} I & 0 \\ C_P M_P(s) & C_P K_P(s) \end{bmatrix} = \text{rk} \begin{bmatrix} I & 0 \end{bmatrix}.$$

Because a polynomial matrix is zero if and only if it is zero at infinitely values it follows that (2.10) is equivalent to the condition $K_Q(s)B_Q \equiv 0$ and $C_P K_P(s) \equiv 0$. Taking into account the solution characterization given in conclusion to Theorem 2.4 the characterization of external regularity is shown. \square

Note that condition (2.10) already appears in the survey paper by Lewis [32] based on arguments in the frequency domain.

2.4 Solution Formulas Based on the Wong Sequences: Regular Case

If the Wong sequences (2.4) are applied to a regular matrix pencil $sE - A \in \mathbb{R}^{n \times n}[s]$ then the limits \mathcal{V}^* and \mathcal{W}^* fulfill (see [2, 5, 51])

$$\begin{aligned} \mathcal{V}^* \cap \mathcal{W}^* &= \{0\}, & \mathcal{V}^* + \mathcal{W}^* &= \mathbb{R}^n, \\ E\mathcal{V}^* \cap A\mathcal{W}^* &= \{0\}, & E\mathcal{V}^* + A\mathcal{W}^* &= \mathbb{R}^n. \end{aligned}$$

In particular $[V, W]$ and $[EV, AW]$ are invertible matrices for all basis matrices V and W of \mathcal{V}^* and \mathcal{W}^* . In fact, any of these invertible matrices yield a transformation which put the matrix pencil $sE - A$ into a quasi-Weierstraß form (QWF):

Theorem 2.6 (Quasi Weierstraß form (QWF) [2, 5]) *Consider a regular matrix pencil $sE - A \in \mathbb{R}^{n \times n}[s]$ and the corresponding Wong sequences with limits \mathcal{V}^* and \mathcal{W}^* . For any full rank matrices V, W with $\text{im } V = \mathcal{V}^*$ and $\text{im } W = \mathcal{W}^*$ let $T = [V, W]$ and $S = [EV, AW]^{-1}$. Then*

$$S(sE - A)T = s \begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} - \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix}, \tag{2.11}$$

where $J \in \mathbb{R}^{n_1 \times n_1}$, $n_1 \in \mathbb{N}$, is some matrix and $N \in \mathbb{R}^{n_2 \times n_2}$, $n_2 = n - n_1$, is nilpotent. In particular, \mathcal{V}^* is exactly the space of consistent initial values, i.e. for all $x_0 \in \mathcal{V}^*$ there exists a unique (classical) solution x of $E\dot{x} = Ax$ with $x(0) = x_0$.

The difference to the WCF from Theorem 2.3 is that J and N are not assumed to be in Jordan canonical form. Furthermore, the transformation matrices for the QWF can be chosen easily; it is only necessary to calculate the Wong sequences.

The knowledge of the two limiting spaces \mathcal{V}^* and \mathcal{W}^* is enough to obtain an explicit solution formula similar to the solution formula (1.3) for ODEs as the next result shows. To formulate the explicit solution formula it is necessary to define certain projectors as follows.

Definition 2.4 (Consistency, differential and impulse projector[43]) Consider a regular matrix pair (E, A) and use the same notation as in Theorem 2.6. The *consistency projector* is given by

$$\Pi_{(E,A)} := T \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} T^{-1},$$

the *differential projector* is given by

$$\Pi_{(E,A)}^{\text{diff}} := T \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} S,$$

and the *impulse projector* is given by

$$\Pi_{(E,A)}^{\text{imp}} := T \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} S,$$

where the block structure is as in the QWF (2.11). Furthermore, let

$$A^{\text{diff}} := \Pi_{(E,A)}^{\text{diff}} A \quad \text{and} \quad E^{\text{imp}} = \Pi_{(E,A)}^{\text{imp}} E.$$

Note that the above defined matrices do not depend on the specific choice of the matrices V and W , because when choosing different basis matrices \tilde{V} and \tilde{W} it must hold that $V = \tilde{V}Q$ and $W = \tilde{W}P$ for some invertible P and Q . Hence

$$\tilde{T} = [\tilde{V}, \tilde{W}] = T \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix} \quad \text{and} \quad \tilde{S} = [E\tilde{V}, A\tilde{W}]^{-1} = \begin{bmatrix} P^{-1} & 0 \\ 0 & Q^{-1} \end{bmatrix} S$$

and the invariance of the above definitions with respect to the choice of V and W is obvious. Furthermore, the differential and impulse projectors are not projectors in the usual sense because they are in general not idempotent.

Theorem 2.7 (Explicit solution formula based on Wong sequences [47]) *Let (E, A) be a regular matrix pair and use the notation from Definition 2.4. Then all solutions of (1.1) are given by, for $c \in \mathbb{R}^n$,*

$$x(t) = e^{A^{\text{diff}}t} \Pi_{(E,A)} c + \int_0^t e^{A^{\text{diff}}(t-\tau)} \Pi_{(E,A)}^{\text{diff}} f(\tau) d\tau - \sum_{i=0}^{n-1} (E^{\text{imp}})^i \Pi_{(E,A)}^{\text{imp}} f^{(i)}(t). \quad (2.12)$$

In particular,

$$x(0) = \Pi_{(E,A)}c - \sum_{i=0}^{n-1} (E^{\text{imp}})^i \Pi_{(E,A)}^{\text{imp}} f^{(i)}(0)$$

i.e. $c \in \mathbb{R}^n$ implicitly specifies the initial value (but in general $x(0) \neq c$ even when $c \in \mathcal{V}^*$).

In the homogeneous case the following equivalence holds [43]:

$$E\dot{x} = Ax \quad \Leftrightarrow \quad \dot{x} = A^{\text{diff}}x \wedge x(0) \in \mathcal{V}^*,$$

which motivates the name differential projector. There is also a motivation for the name of the impulse projector, see the end of Sect. 4 as well as Sect. 5.

The Wong sequences appeared sporadically in the DAE literature: For example, Yip and Sincovec [52] used them to characterize regularity of the matrix pencil, Owens and Debeljkovic [36] characterized the space of consistent initial values via the Wong sequences; they are also included in the text books [1, 29] but not in the text books [7–9, 14, 30]. In general it seems that the connection between the Wong sequences and the (quasi-)Weierstraß/Kronecker form and their role in the solution characterization is not well known or appreciated in the DAE community (especially in the case of singular matrix pencils).

2.5 The Drazin Inverse Solution Formula

Another explicit solution formula was proposed by Campbell et al. [11] already in 1976 and is based on the Drazin inverse.

Definition 2.5 (Drazin inverse [17]) For $M \in \mathbb{R}^{n \times n}$ a matrix $D \in \mathbb{R}^{n \times n}$ is called Drazin inverse if, and only if,

1. $DM = MD$,
2. $D = DMD$,
3. $\exists \nu \in \mathbb{N} : M^\nu = M^{\nu+1}D$.

In [17] it is shown that the Drazin inverse is unique and it is easy to see that the Drazin inverse of M is given by

$$M^D = T \begin{bmatrix} J^{-1} & 0 \\ 0 & 0 \end{bmatrix} T^{-1},$$

where the invertible matrix T is such that

$$M = T \begin{bmatrix} J & 0 \\ 0 & N \end{bmatrix} T^{-1},$$

J is invertible and N is nilpotent. In particular, for invertible M the Drazin inverse is just the classical inverse, i.e. $M^{-1} = M^D$.

The following solution formula for the DAE (1.1) based on the Drazin inverse needs commutativity of the matrices E and A , however, as also regularity is assumed the following result shows that this is not a restriction of generality.

Lemma 2.8 (Commutativation of (E, A) [11]) *Assume (E, A) is regular and chose $\lambda \in \mathbb{R}$ such that $\lambda E - A$ is invertible. Then*

$$(\lambda E - A)^{-1} E \quad \text{and} \quad (\lambda E - A)^{-1} A$$

commute, i.e. the whole equation (1.1) can simply be multiplied from the left with $(\lambda E - A)^{-1}$ which will not change the solution properties but will guarantee commutativity of the coefficient matrices.

Theorem 2.9 (Explicit solution formula based on the Drazin inverse [11]) *Consider the regular DAE (1.1) with $EA = AE$. Then all solutions x are given by*

$$\begin{aligned} x(t) = & e^{E^D A t} E^D E c + \int_0^t e^{E^D A(t-\tau)} E^D f(\tau) d\tau \\ & - (I - E^D E) \sum_{i=0}^{n-1} (EA^D)^i A^D f^{(i)}(t). \end{aligned} \quad (2.13)$$

A direct comparison of the solution formula (2.12) based on the Wong sequences and (2.13) indicates that $E^D A$ plays the role of A^{diff} , $E^D E$ plays the role of the consistency projector and E^D plays the role of the differential projector. However, the connection between the impulse projector and E^{imp} to the expressions involving the Drazin inverse of A is not immediately clear. The following result justifies the previous observations.

Lemma 2.10 (Wong sequences and Drazin inverse [5]) *Consider the regular matrix pair (E, A) with $EA = AE$ and use the notation from Theorem 2.6. Then*

$$E^D = T \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} S \quad \text{and} \quad A^D = T \begin{bmatrix} J^D & 0 \\ 0 & I \end{bmatrix} S.$$

In particular, also taking into account $E = S^{-1} \begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} T^{-1}$ and $A = S^{-1} \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix} T^{-1}$,

$$E^D = \Pi_{(E,A)}^{\text{diff}},$$

$$E^D A = \Pi_{(E,A)}^{\text{diff}} A = A^{\text{diff}},$$

$$E^D E = T \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} S S^{-1} \begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} T^{-1} = \Pi_{(E,A)},$$

$$(EA^D)^i = \left(S^{-1} \begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} T^{-1} T \begin{bmatrix} J^D & 0 \\ 0 & I \end{bmatrix} S \right)^i = S^{-1} \begin{bmatrix} (J^D)^i & 0 \\ 0 & N^i \end{bmatrix} S, \quad i \in \mathbb{N},$$

and with some more effort, using $E\mathcal{V}^* = \mathcal{V}^*$ and $A\mathcal{W}^* = \mathcal{W}^*$ in the commuting case (see [5]), it follows that

$$(I - E^D E)(EA^D)^i A^D = T \begin{bmatrix} 0 & 0 \\ 0 & N^i \end{bmatrix} S = (E^{\text{imp}})^i \Pi_{(E,A)}^{\text{imp}}.$$

This shows that indeed the two solution formulas (2.12) and (2.13) are identical in the commuting case. Note that in the solution formula (2.13) the Drazin inverse A^D appears and one might therefore think that the occurrence of zero eigenvalues in A plays some special role for the solution. However, this is just an artifact and it turns out that in the expression

$$A^D = T \begin{bmatrix} J^D & 0 \\ 0 & I \end{bmatrix} S$$

the matrix J^D can be replaced by an arbitrary matrix without changing the result of the solution formula (2.13). One canonical choice is to replace J^D by the zero matrix which yields the impulse projector and which makes the ‘‘correction term’’ $(I - E^D E)$ superfluous.

2.6 Time-Varying DAEs

In this section the time-varying version of (1.1), i.e.

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

is briefly discussed.

Campbell and Petzold [10] proved that if $E(\cdot)$ and $A(\cdot)$ have real analytical entries then a solution characterization similar to Corollary 2.2 holds. In particular, they showed that unique solvability is equivalent to finding time-varying (analytical) transformation matrices $S(\cdot)$, $T(\cdot)$, such that

$$(S(t)E(t)T(t), S(t)A(t)T(t) - S(t)E(t)T'(t)) = \left(\begin{bmatrix} I & 0 \\ 0 & N(t) \end{bmatrix}, \begin{bmatrix} J(t) & 0 \\ 0 & I \end{bmatrix} \right),$$

where $N(t)$ is a strictly lower triangular (and hence nilpotent) matrix. In particular, as in the time-invariant case, the DAE decouples into an ODE part and a pure DAE part. It is easily seen that for a strictly lower triangular matrix $N(t)$ also the differential operator $N(\cdot)\frac{d}{dt}$ is nilpotent, hence the inverse operator of $(N(\cdot)\frac{d}{dt} - I)$ can be calculated nearly identically as in (2.2):

$$\left(N(\cdot)\frac{d}{dt} - I \right)^{-1} = - \sum_{i=0}^{\nu-1} \left(N(\cdot)\frac{d}{dt} \right)^i,$$

where $\nu \in \mathbb{N}$ is the nilpotency index of the operator $N(\cdot)\frac{d}{dt}$.

If the coefficient matrices are not analytical the situation is not so clear anymore and different approaches have been proposed. Most methods have their motivation in numerical simulations and a detailed description and discussion is outside the scope of this survey. The interested reader is referred to the nice survey by Rabier and Rheinboldt [38], and to the text book by Kunkel and Mehrmann [30] as well as the recent monograph by Lamour, März and Tischendorf [31]. However, all these approaches do not allow for discontinuous coefficient matrices. These are studied in [46] and because of the connection to inconsistent initial value problems the problem of discontinuous coefficient matrices is further discussed in Sect. 5.

3 Inconsistent Initial Values and Distributional Solutions

After having presented a rather extensive discussion of classical solutions, this section presents an introductory discussion of the problem of inconsistent initial values. From the above derived solution formulas for (1.1) it becomes apparent that $x(0)$ cannot be chosen arbitrarily, certain parts of $x(0)$ are already fixed by the DAE and the inhomogeneity, cf. Theorem 2.7. In the extreme case that the QWF of (E, A) only consists of the nilpotent part, the initial value $x(0)$ is completely determined by the inhomogeneity and no freedom to choose the initial value is left. However, there are situations where one wants to study the response of a system described by a DAE when an *inconsistent initial value* is given. Examples are electrical circuits which are switched on at a certain time [48]. There have been different approaches to deal with inconsistent initial values, e.g. [12, 18, 35, 37, 39, 42], some of them will be presented in detail in the later sections. All have in common that jumps as well as Dirac impulses may occur in the solutions. The Dirac impulse is a distribution (a generalized function), hence one must enlarge the considered solution space to also include distributions. In fact, also the presence of non-smooth inhomogeneities (or inputs) can lead to distributional solutions. However, the latter do not produce conceptual difficulties as the solution characterization of the previous section basically remains unchanged.

In order to be able to make mathematical precise statements the classical distribution theory [41] is revised first. The space of *test functions* is given by

$$\mathcal{C}_0^\infty := \{ \varphi : \mathbb{R} \rightarrow \mathbb{R} \mid \varphi \in \mathcal{C}^\infty \text{ has compact support} \},$$

which is equipped with a certain topology.² The space of distributions, denoted by \mathbb{D} , is then the dual of the space of test functions, i.e.

$$\mathbb{D} := \{ D : \mathcal{C}_0^\infty \rightarrow \mathbb{R} \mid D \text{ is linear and continuous} \}.$$

²The topology is such that a sequence $(\varphi_k)_{k \in \mathbb{N}}$ of test functions converges to zero if, and only if, (1) the supports of all φ_k are contained within one common compact set $K \subseteq \mathbb{R}$ and (2) for all $i \in \mathbb{N}$, $\varphi_k^{(i)}$ converges uniformly to zero as $k \rightarrow \infty$.

A large class of ordinary functions, namely locally integrable functions, can be embedded into \mathbb{D} via the following injective³ homomorphism:

$$f \mapsto f_{\mathbb{D}}, \quad \text{with } f_{\mathbb{D}}(\varphi) := \int_{\mathbb{R}} f \varphi.$$

The main feature of distributions is the ability to take derivatives for any distribution $D \in \mathbb{D}$ via

$$D'(\varphi) := -D(\varphi').$$

Simple calculations show that this is consistent with the classical derivative, i.e. if f is differentiable, then

$$(f_{\mathbb{D}})' = (f')_{\mathbb{D}}.$$

In particular, the Heaviside unit step $\mathbb{1}_{[0, \infty)}$ has a distributional derivative which can easily be calculated to be

$$(\mathbb{1}_{[0, \infty)}_{\mathbb{D}})'(\varphi) = \varphi(0) =: \delta(\varphi),$$

hence it results in the well known Dirac impulse δ (at $t = 0$). In general, the Dirac impulse δ_t at time $t \in \mathbb{R}$ is given by $\delta_t(\varphi) := \varphi(t)$. Furthermore, if g is a piecewise differentiable function with one jump at $t = t_j$, i.e. g is given as

$$g(t) = \begin{cases} g_1(t), & t < t_j, \\ g_2(t), & t \geq t_j, \end{cases}$$

where g_1 and g_2 are differentiable functions and

$$g^1(t) := \begin{cases} g'_1(t), & t < t_j, \\ g'_2(t), & t \geq t_j, \end{cases}$$

then

$$(g_{\mathbb{D}})' = (g^1)_{\mathbb{D}} + (g(t_{j+}) - g(t_{j-}))\delta_{t_j}. \tag{3.1}$$

In other words, taking derivatives of a general jump results in a Dirac impulse at the jump position whose amplitude is the height of the jump.

Finally, distributions can be multiplied with smooth functions α :

$$(\alpha D)(\varphi) = D(\alpha\varphi)$$

and it is easily seen that this multiplication is consistent with the pointwise multiplication of functions and that the Leibniz product rule holds:

$$(\alpha D)' = \alpha' D + \alpha D'.$$

³Two locally integrable functions which only differ on a set of measure zero are identified with each other.

Now it is no problem to consider the DAE (1.1) in a distributional solution space, instead of x and f being vectors of functions they are now vectors of distributions, i.e. $x \in \mathbb{D}^n$ and $f \in \mathbb{D}^m$ where $m \times n$ is the size of the matrices E and A . The definition of the matrix vector product remains unchanged⁴ so that (1.1) reads as m equations in \mathbb{D} .

Considering distributional solutions, however, does not help to treat inconsistent initial value; au contraire, distributions cannot be evaluated at a certain time because they are not functions of time, so writing $x(0) = x_0$ makes no sense. Even when assuming that a pointwise evaluation is well defined for certain distributions, the DAE (1.1) will still not exhibit (distributional) solution with arbitrary initial values. This is easily seen when considering the DAE $N\dot{x} = x + f$ with nilpotent N . Then also in the distributional solution framework the operator $N\frac{d}{dt} - I : \mathbb{D} \rightarrow \mathbb{D}$ is invertible with inverse as in (2.2) and there exists a unique (distributional) solution given by

$$x = - \sum_{i=0}^{n-1} N^i f^{(i)},$$

hence the initial value of x cannot be assigned arbitrarily (i.e. independently of the inhomogeneity).

So what does it then mean to speak of a solution of (1.1) with inconsistent initial value? The motivation for inconsistent initial values is the situation that the system descriptions gets active at the initial time $t = 0$ and before that the system was governed by different (maybe unknown) rules. This viewpoint was also expressed by Doetsch [16, p. 108] in the context of distributional solutions for ODEs:

The concept of “initial value” in the physical science can be understood only when the past, that is, the interval $t < 0$, has been included in our considerations. This occurs naturally for distributions which, without exception, are defined on the entire t -axis.

So mathematically, there is some given past trajectory x^0 for x up to the initial time and the DAE (1.1) only holds on the interval $[0, \infty)$. This means that a solution of the following *initial trajectory problem* (ITP) is sought:

$$\begin{aligned} x_{(-\infty, 0)} &= x^0_{(-\infty, 0)}, \\ (E\dot{x})_{[0, \infty)} &= (Ax + f)_{[0, \infty)}, \end{aligned} \tag{3.2}$$

where $x^0 \in \mathbb{D}^n$ is an arbitrary past trajectory and D_I for some interval $I \subseteq \mathbb{R}$ and $D \in \mathbb{D}$ denotes a distributional restriction generalizing the restrictions of functions given by

$$f_I(t) = \begin{cases} f(t), & t \in I, \\ 0, & t \notin I. \end{cases}$$

⁴Some authors [30, 38] use a different definition for the matrix vector product which is due to the different viewpoint of a distributional vector x as a map from $(\mathcal{C}_0^\infty)^n$ to \mathbb{R} instead of a map from \mathcal{C}_0^∞ to \mathbb{R}^n . The latter seems the more natural approach in view of applying it to (1.1), but it seems that both approaches are equivalent at least with respect to the solution theory of DAEs.

A fundamental problem is the fact (see Lemma 5.1) that such a distributional restriction does not exist!

This problem was resolved especially in older publication [8, 9, 48] by ignoring it and/or by arguing with the Laplace transform (see the next section). Cobb [13] seems to be the first to be aware of this problem and he resolved it by introducing the space of piecewise-continuous distributions; Geerts [22, 23] was the first to use the space of impulsive-smooth distributions (introduced in [27]) as a solution space for DAEs. Seemingly unaware of these two approaches, Tolsa and Salichs [44] developed a distributional solution framework which can be seen as a mixture between the approaches of Cobb and Geerts. The more comprehensive space of piecewise-smooth distributions was later introduced [45] to combine the advantages of the piecewise-continuous and impulsive-smooth distributional solution spaces. The details are discussed in Sect. 5.

Cobb [12] also presented another approach by justifying the impulsive response due to inconsistent initial values via his notion of *limiting solutions*. The idea is to replace the singular matrix E in (1.1) by a “disturbed” version E_ε which is invertible for all $\varepsilon > 0$ and $E_\varepsilon \rightarrow E$ as $\varepsilon \rightarrow 0$. If the solutions of the corresponding initial value ODE problem $\dot{x} = E_\varepsilon^{-1}Ax$, $x(0) = x_0$ converges to a distribution, then Cobb calls this the limiting solution. He is then able to show that the limiting solution is unique and equal to the one obtained via the Laplace-transform approach. Campbell [9] extends this result also to the inhomogeneous case.

4 Laplace Transform Approaches

Especially in the signal theory community it is common to study systems like (1.1) or (2.9) in the so called *frequency domain* (in contrast to the *time domain*). In particular, when the input-output mapping is of interest the frequency domain approach significantly simplifies the analysis. The transformation between time and frequency domain is given by the *Laplace transform* defined via the Laplace integral:

$$\hat{g}(s) := \int_0^\infty e^{-st} g(t) dt \quad (4.1)$$

for some function g and $s \in \mathbb{C}$. Note that in general the Laplace integral is not well defined for all $s \in \mathbb{C}$ and a suitable domain for \hat{g} must be chosen [16]. If a suitable domain exists, then $\hat{g} = \mathcal{L}\{g\}$ is called the *Laplace transform* of g and, in general, $\mathcal{L}\{\cdot\}$ denotes the Laplace transform operator. Again note that it is not specified at this point which class of functions have a Laplace transform and which class of functions are obtained as the image of $\mathcal{L}\{\cdot\}$. The main feature of the Laplace transform is the following property, where g is a differentiable function for which g and g' have Laplace transforms:

$$\mathcal{L}\{g'\}(s) = s\mathcal{L}\{g\}(s) - g(0), \quad (4.2)$$

which is a direct consequence of the definition of the Laplace integral invoking partial differentiation. If g is not continuous at $t = 0$ but $g(0+)$ exists and g' denotes the derivative of g on $\mathbb{R} \setminus \{0\}$, then (4.2) still holds in a slightly altered form:

$$\mathcal{L}\{g'\}(s) = s\mathcal{L}\{g\}(s) - g(0+). \quad (4.3)$$

In particular, the Laplace transform does not take into account at all how g behaved for $t < 0$ which is a trivial consequence of the definition of the Laplace integral. This observation will play an important role when studying inconsistent initial values.

Taking into account the linearity of the Laplace transform the descriptor system (2.9) is transformed into

$$\begin{aligned} sE\hat{x}(s) &= A\hat{x}(s) + B\hat{u}(s) + Ex(0+), \\ \hat{y}(s) &= C\hat{x}(s) + D\hat{u}(s). \end{aligned} \quad (4.4)$$

If the matrix pair (E, A) is regular and $x(0+) = 0$, the latter can be solved easily algebraically:

$$\hat{y}(s) = (C(sE - A)^{-1}B + D)\hat{u}(s) =: G(s)\hat{u}(s), \quad (4.5)$$

where $G(s)$ is a matrix over the field of rational functions and is usually called transfer function. As there are tables of functions and its Laplace transforms it is often possible to find the solutions of descriptor system with given input simply by plugging the Laplace transform of the input in the above formula and lookup the resulting output $\hat{y}(s)$ to obtain the solution $y(t)$ in the time domain. Furthermore, many important system properties can be deduced from properties (like the zeros and poles) of the transfer function directly.

A first systematic treatment of descriptor systems in the frequency domain was carried out by Rosenbrock [40]. He, however, only considered zero initial values and the input-output behavior. In particular, he was not concerned with a solution theory for general DAEs (1.1) with possible inconsistent values. Furthermore, he restricted attention to inputs which are exponentially bounded (guaranteeing existence of the Laplace transform), hence formally his framework could not deal with arbitrary (sufficiently smooth) inputs.

The definition of the Laplace transform can be extended to be well defined for certain distributions as well [16], therefore consider the following class of distributions:

$$\mathbb{D}_{\geq 0, k} := \{D = (g_{\mathbb{D}})^{(k)} \mid \text{where } g : \mathbb{R} \rightarrow \mathbb{R} \text{ is continuous and } g(t) = 0 \text{ on } (-\infty, 0)\}.$$

For $D \in \mathbb{D}_{\geq 0, k}$ with $D = (g_{\mathbb{D}})^{(k)}$ the (distributional) Laplace transform is now given by

$$\mathcal{L}_{\mathbb{D}}\{D\}(s) := s^k \mathcal{L}\{g\}(s)$$

on a suitable domain in \mathbb{C} . Note that $\delta \in \mathbb{D}_{\geq 0, 2}$ and it is easily seen that

$$\mathcal{L}_{\mathbb{D}}\{\delta\} = 1. \quad (4.6)$$

Furthermore, for every locally integrable function g for which $\mathcal{L}\{g\}$ is defined on a suitable domain it holds that

$$\mathcal{L}_{\mathbb{D}}\{g_{\mathbb{D}}\} = s\mathcal{L}\left\{\int_0^{\cdot} g\right\} \stackrel{[16]}{=} s\frac{1}{s}\mathcal{L}\{g\} = \mathcal{L}\{g\}, \tag{4.7}$$

i.e. the distributional Laplace transform coincides with the classical Laplace transform defined by (4.1).

A direct consequence of the definition of $\mathcal{L}_{\mathbb{D}}$ is the following derivative rule for all $D \in \bigcup_k \mathbb{D}_{\geq 0, k}$:

$$\mathcal{L}_{\mathbb{D}}\{D'\}(s) = s\mathcal{L}_{\mathbb{D}}\{D\} \tag{4.8}$$

which seems to be in contrast to the derivative rule (4.3), because no initial value occurs. The latter can actually not be expected because general distributions do not have a well defined function evaluation at a certain time t . However, the derivative rule (4.8) is consistent with (4.3); to see this let g be a function being zero on $(-\infty, 0)$, differentiable on $(0, \infty)$ with well defined value $g(0+)$. Denote with g' the (classical) derivative of g on $\mathbb{R} \setminus \{0\}$, then (invoking linearity of $\mathcal{L}_{\mathbb{D}}$)

$$\begin{aligned} \mathcal{L}_{\mathbb{D}}\{(g_{\mathbb{D}})'\}(s) &\stackrel{(3.1)}{=} \mathcal{L}_{\mathbb{D}}\{(g')_{\mathbb{D}} + g(0+)\delta\}(s) \\ &= \mathcal{L}_{\mathbb{D}}\{(g')_{\mathbb{D}}\}(s) + g(0+)\mathcal{L}_{\mathbb{D}}\{\delta\}(s) \stackrel{(4.6),(4.7)}{=} \mathcal{L}\{g'\} + g(0+), \end{aligned}$$

which shows equivalence of (4.8) and (4.3). The key observation is that the distributional derivative takes into account the jump at $t = 0$ whereas the classical derivative ignores it, i.e. in the above context

$$(g_{\mathbb{D}})' \neq (g')_{\mathbb{D}}.$$

As it is common to identify g with $g_{\mathbb{D}}$ (even in [16]), the above distinction is difficult to grasp, in particular for inexperienced readers. As this problem plays an important role when dealing with inconsistent initial values, it is not surprising that researchers from the DAE community who are simply using the Laplace transform as a tool, struggle with the treatment of inconsistent initial values, cf. [34].

Revisiting the treatment of the descriptor system (2.9) in the frequency domain one has now to decide whether to use the usual Laplace transform resulting in (4.4) or the distributional Laplace transform resulting in

$$\begin{aligned} sE\hat{x}(s) &= A\hat{x}(s) + B\hat{u}(s), \\ \hat{y}(s) &= C\hat{x}(s) + D\hat{u}(s), \end{aligned} \tag{4.9}$$

where the initial value $x(0+)$ does not occur anymore. In particular, if the matrix pair (E, A) is regular, the only solution of (4.9) is given by (4.5) independently of $x(0+)$. In particular, if $u = 0$ the only solution of (4.9) is $\hat{x}(s) = 0$ and $\hat{y}(s) = 0$. Assuming a well defined inverse Laplace transform this implies that the only solution of (2.9) with $u = 0$ is the trivial solution, which is of course not true in general. Altogether the following dilemma occurs.

Dilemma (Discrepancy between time domain and frequency domain) Consider the regular DAE (1.1) or more specifically (2.9) with zero inhomogeneity (input) but non-zero initial value.

- An ad hoc analysis calls for *distributional solutions* in response to inconsistent initial values. For consistent initial value there exist classical (non-zero) solutions.
- Using the *distributional* Laplace transform to analyze the (distributional) solutions of (1.1) or (2.9) reveals that the *only* solution is the trivial one. In particular, no initial values (neither inconsistent nor consistent ones) are taken into account at all.

This problem was already observed in [16, p. 108] and is based on the definition of the distributional Laplace transform which is only defined for distributions vanishing on $(-\infty, 0)$. The following “solution” to this dilemma was suggested [16, p. 129]: Define for $D \in \bigcup_k \mathbb{D}_{\geq 0, k}$ the “past-aware” derivative operator $\frac{d_-}{dt}$:

$$\frac{d_-}{dt} D := D' - d_0^- \delta, \quad (4.10)$$

where $d_0^- \in \mathbb{R}$ is interpreted as a “virtual” initial value for $D(0-)$. Note, however, that, by definition, $D(0-) = 0$ for every $D \in \bigcup_k \mathbb{D}_{\geq 0, k}$; hence at this stage it is not clear why this definition makes sense. This problem was also pointed out by Cobb [12]. Nevertheless, a motivation for this choice will be given in Sect. 5.

Using now the past-aware derivative in the distributional formulation of (1.1) one obtains

$$\begin{aligned} Ex' &= Ax + Bu + Ex_0^- \delta, \\ y &= Cx + Du, \end{aligned} \quad (4.11)$$

where $x_0^- \in \mathbb{R}^n$ is the virtual (possible inconsistent) initial value for $x(0-)$ and solutions are sought in the space $(\bigcup_k \mathbb{D}_{\geq 0, k})^n$, i.e. x is assumed to be zero on $(-\infty, 0)$. Applying the distributional Laplace transform to (4.11) yields

$$\begin{aligned} sE\hat{x}(s) &= A\hat{x}(s) + B\hat{u}(s) + Ex_0^-, \\ \hat{y}(s) &= C\hat{x}(s) + D\hat{u}(s). \end{aligned} \quad (4.12)$$

In contrast to (4.4), x_0^- is not the initial value for $x(0+)$ but is the virtual initial value for $x(0-)$. If the matrix pair (E, A) is regular, the solution of (4.12) can now be obtained via

$$\hat{x}(s) = (sE - A)^{-1}(B\hat{u}(s) + Ex_0^-)$$

and using the inverse Laplace transform. Because E is not invertible in general, the rational matrix $(sE - A)^{-1}$ may contain polynomial entries resulting in polynomial parts in \hat{x} corresponding to Dirac impulses in the time domain, for details see the end of this section.

The solution formula for $\hat{x}(s)$ is possible to calculate analytically when the matrices E , A , and B are known and for suitable inputs u the inverse Laplace transform of $\hat{x}(s)$ can also be obtained analytically. This is the main advantage of the Laplace transform approach. There are, however, the following major drawbacks:

1. Within the frequency domain it is not possible to motivate the incorporation of the (inconsistent) initial values as in (4.11); in fact, Doetsch [16] who seems to have introduced this notion, needs to argue with the help of the distributional derivative and (4.10) within the time domain!
2. The Laplace transform ignores everything that was in the past, i.e. on the interval $(-\infty, 0)$; this is true for the classical Laplace transform (by definition of the Laplace integral) as well as for the distributional Laplace transform (by only considering distributions which vanish for $t < 0$). Hence the natural viewpoint of an initial trajectory problem (3.2) as also informally advocated by Doetsch cannot possibly be treated with the Laplace transform approach.
3. A frequency domain analysis gets useless when the original system is time-varying or nonlinear, whereas (linear) time-domain methods may in principle be extended to also treat time-variance and certain non-linearities. In fact, the piecewise-smoothly distributional solution framework as presented in Sect. 5 can be used without modification for linear time-varying DAEs [46] and also for certain non-linear DAEs [33].
4. Making statements about existence and uniqueness of solution with the help of the frequency domain heavily depends on an isomorphism between the time-domain and the frequency domain; there are, however, only a few special isomorphisms between certain special subspaces of the frequency and time domain, no general isomorphism is available, see also the discussion concerning (4.9).

This section on the Laplace domain concludes with the calculation of the re-initialization of the inconsistent initial value as well as the resulting Dirac impulses occurring in the solution. Therefore, consider the “distributional version” (following Doetsch) of (1.1):

$$E\dot{x} = Ax + f_{\mathbb{D}} + Ex_0^- \delta, \tag{4.13}$$

where $x_0^- \in \mathbb{R}^n$, and its corresponding Laplace transformed version in frequency domain

$$sE\hat{x}(s) = A\hat{x}(s) + \hat{f}(s) + Ex_0^-. \tag{4.14}$$

The unique solution of (4.14) in frequency domain is given by

$$\hat{x}(s) = (sE - A)^{-1}(\hat{f}(s) + Ex_0^-),$$

which needs regularity of the matrix pair (E, A) to be well defined, which will therefore be assumed in the following. Applying a coordinate transformation $x = T \begin{pmatrix} v \\ w \end{pmatrix}$

according to the QWF (2.11), the solution in the new coordinates is given by

$$\begin{aligned} \begin{pmatrix} \hat{v}(s) \\ \hat{w}(s) \end{pmatrix} &= T^{-1}(sE - A)^{-1} \left(\hat{f}(s) + ET \begin{pmatrix} v_0^- \\ w_0^- \end{pmatrix} \right) \\ &= (sSET - SAT)^{-1} \left(S\hat{f}(s) + SET \begin{pmatrix} v_0^- \\ w_0^- \end{pmatrix} \right), \end{aligned}$$

where $x_0^- =: T \begin{pmatrix} v_0^- \\ w_0^- \end{pmatrix}$. Hence, invoking the QWF (2.11), the solution formula decouples into

$$\begin{aligned} \hat{v}(s) &= (sI - J)^{-1}(\hat{f}_1(s) + v_0^-), \\ \hat{w}(s) &= (sN - I)^{-1}(\hat{f}_2(s) + Nw_0^-) = - \sum_{i=0}^{v-1} N^i s^i (\hat{f}_2(s) + Nw_0^-), \end{aligned}$$

where $Sf =: \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}$ and $v \in \mathbb{N}$ is the nilpotency index of N . Since $(sI - J)^{-1}$ is a strictly proper rational matrix, the solution for v (resulting from taking the inverse Laplace transform) is the corresponding standard ODE solution (1.3). In particular, $v(0+) = v_0^-$ and no Dirac impulses occur in v . Applying the inverse Laplace transformation on the solution formula for $\hat{w}(s)$, one obtains the solution $w = w_f + w_i$, where w_f is the response with respect to the inhomogeneity given by

$$w_f := - \sum_{i=0}^{v-1} N^i (f_{2\mathbb{D}})^{(i)}$$

and w_i consists of Dirac impulses at $t = 0$ produced by the inconsistent initial value:

$$w_i := - \sum_{i=0}^{v-1} N^{i+1} w_0^- \delta^{(i)}.$$

Note that in order to obtain w_f by using the correspondence (4.8), the distributional derivatives of f_2 have to be considered. As the (distributional) Laplace transform can only be applied to distributions vanishing on $(-\infty, 0)$, the inhomogeneity f_2 will in general have a jump at $t = 0$, hence w_f will also contain Dirac impulses depending on $f_2^{(i)}(0+)$, $i = 0, 1, \dots, v - 1$. In summary:

Theorem 4.1 (Solution formula obtained via the Laplace transform approach) *Consider the regular DAE (1.1) with its “distributional version” (4.13). Let $v \in \mathbb{N}$ be the nilpotency index of N in the QWF (2.11) of the matrix pair (E, A) . Assume $f : \mathbb{R} \rightarrow \mathbb{R}^n$ is zero on $(-\infty, 0)$ and $v - 1$ times differentiable on $(0, \infty)$ with well defined values $f^{(i)}(0+)$, $i = 0, 1, \dots, v - 1$. Use the notation from Definition 2.4. Then $x \in (\bigcup_k \mathbb{D}_{\geq 0, k})^n$ given by (2.12) on $(0, \infty)$ with $c = x_0^-$ and by the impulsive*

part at $t = 0$, denoted by $x[0]$,

$$\begin{aligned}
 x[0] = & - \sum_{i=0}^{v-2} (E^{\text{imp}})^{i+1} \sum_{j=0}^i \Pi_{(E,A)}^{\text{imp}} f^{(i-j)}(0+) \delta^{(j)} \\
 & - \sum_{i=0}^{v-2} (E^{\text{imp}})^{i+1} (I - \Pi_{(E,A)}) x_0^- \delta^{(i)}
 \end{aligned} \tag{4.15}$$

is the unique solution of (4.13) obtained via solving (4.14). In particular,

$$x(0+) = \Pi_{(E,A)} x_0^- + \sum_{i=0}^{n-1} (E^{\text{imp}})^i \Pi_{(E,A)}^{\text{imp}} f^{(i)}(0+), \tag{4.16}$$

hence if $f \equiv 0$ then the consistent reinitialization is given by the consistency projector $\Pi_{(E,A)}$ via

$$x(0+) = \Pi_{(E,A)} x_0^-.$$

Proof Invoking (3.1), one obtains

$$(f_{2\mathbb{D}})^{(i)}[0] = \sum_{j=0}^{i-1} f_2^{(i-1-j)}(0+) \delta^{(j)},$$

hence

$$w_f[0] = - \sum_{i=0}^{v-2} N^{i+1} \sum_{j=0}^i f_2^{(i-j)}(0+) \delta^{(j)}.$$

Now using the identities, cf. [47],

$$\begin{aligned}
 A^{\text{diff}} &= T \begin{bmatrix} J & 0 \\ 0 & 0 \end{bmatrix} T^{-1}, & E^{\text{imp}} &= T \begin{bmatrix} 0 & 0 \\ 0 & N \end{bmatrix} T^{-1}, \\
 T \begin{pmatrix} f_1 \\ 0 \end{pmatrix} &= \Pi_{(E,A)}^{\text{diff}} f, & T \begin{pmatrix} 0 \\ f_2 \end{pmatrix} &= \Pi_{(E,A)}^{\text{imp}} f, & T \begin{pmatrix} v_0^- \\ 0 \end{pmatrix} &= \Pi_{(E,A)} x_0^-, \\
 T \begin{pmatrix} 0 \\ w_0^- \end{pmatrix} &= (I - \Pi_{(E,A)}) x_0^-
 \end{aligned}$$

yields the claimed solution formula. □

5 Distributional Solutions

The previous section introduced distributional solutions in order to treat inconsistent initial values with the help of the Laplace transform. This leads to the consideration

of the distributional space $\bigcup_k \mathbb{D}_{\geq 0, k}$ which contains all distributions which can be written as a (distributional) k th derivative, $k \in \mathbb{N}$, of a continuous function being zero on $(-\infty, 0)$ and of which a Laplace transform exists. This choice is motivated by the applicability of the Laplace transform and is actually not motivated by dealing with inconsistent initial values. In fact, as was pointed out in the previous section, the Laplace transform ignores by definition/design all what has happened before $t < 0$ and is therefore in principle not suitable to treat inconsistent initial values coming from the past. Most researchers in the field agree with the notion that an inconsistent initial is due to a past which was *not* governed by the system description (1.1). One way of formalizing this viewpoint is the ITP (3.2). In general, having a past which obeys different rules than the present means that the overall system description is *time-variant* which gives another reason why the Laplace-transform approach runs into difficulties.

5.1 The Problem of Distributional Restrictions

Treating the ITP (3.2) in a distributional solution framework is, however, also not straightforward, because (as already mentioned above) the distributional restriction used in (3.2) is not well defined.

Lemma 5.1 (Bad distribution [45]) *Let D be the (distributional, i.e. weak*) limit of the distributions:*

$$D_k := \sum_{i=0}^k d_i \delta_{d_i}, \quad \text{where } d_i := \frac{(-1)^i}{i+1}, \quad i, k \in \mathbb{N}.$$

Then the restriction (in the sense of [45]) of D to the interval $[0, \infty)$ is not a well-defined distribution.

Proof Clearly,

$$D_{[0, \infty)} = \sum_{j=0}^{\infty} d_{2j} \delta_{d_{2j}},$$

however, applying $D_{[0, \infty)}$ to a test function φ which is identically one on $[0, 1]$ yields

$$D_{[0, \infty)}(\varphi) = \sum_{j=0}^{\infty} d_{2j} \delta_{d_{2j}}(\varphi) = \sum_{j=0}^{\infty} \frac{1}{2^j} = \infty,$$

which shows that $D_{[0, \infty)}$ is not a well defined distribution. \square

Remark 5.1 (Restriction to open intervals) The above results remain true when considering restriction to *open* intervals. However, it should be mentioned here that

nevertheless the equation $F_I = G_I$ makes sense for arbitrary distributions $F, G \in \mathbb{D}$ and any open interval $I \subseteq \mathbb{R}$ by defining:

$$F_I = G_I \quad :\Leftrightarrow \quad \forall \varphi \in \mathcal{C}_0^\infty \text{ with } \text{supp } \varphi \subseteq I : F(\varphi) = G(\varphi).$$

In fact, this definition is consistent with the restriction-definition to be established in the following for a special class of distributions [45, Prop. 2.2.10]. Nevertheless, restricting the second equation in the ITP (3.2) to the *closed* interval $[0, \infty)$ is essential. Taking an open restriction in both equations of (3.2) would imply that the past and the present are decoupled so that the initial trajectory would not influence the future trajectory. To be more precise: Any (distributional) solution x of (3.2) will exhibit a jump at $t = 0$ in response to an inconsistent value $x_0(0-)$, but the derivative of this jump appears as a Dirac impulse in the expression $E\dot{x}$. While the restriction to the open interval $(0, \infty)$ would neglect this Dirac impulse, the restriction to the closed interval $[0, \infty)$ keeps the Dirac impulse in the second equation of the ITP (3.2) and hence the past can influence the present.

5.2 Cobb's Space of Piecewise-Continuous Distributions

The need to define a restriction for distributions was already advocated by Cobb [13]; although his motivation was not the ITP (3.2) but a rigorous definition of the impulsive term $D[t]$ of a distribution D at time $t \in \mathbb{R}$ which can be viewed as a restriction to the interval $[t, t]$. To this end, Cobb first defined the space of piecewise-continuous distributions given by

$$\mathbb{D}_{\text{pw}\mathcal{C}^0} := \left\{ D \in \mathbb{D} \left| \begin{array}{l} \exists T = \{t_i \in \mathbb{R} \mid i \in \mathbb{Z}\} \text{ ordered and locally finite} \\ \exists g \in \mathcal{C}_{\text{pw}}^0 \forall i \in \mathbb{Z} : D_{(t_i, t_{i+1})} = (g_{\mathbb{D}})_{(t_i, t_{i+1})} \\ \text{in the sense of Remark 5.1} \end{array} \right. \right\},$$

where $\mathcal{C}_{\text{pw}}^0$ denotes the space of piecewise-continuous functions, in particular, for any $g \in \mathcal{C}_{\text{pw}}^0$ the values $g(t+)$ and $g(t-)$ are well defined for all $t \in \mathbb{R}$.

Definition 5.1 (Cobb's distributional restriction [13]) Let $D \in \mathbb{D}_{\text{pw}\mathcal{C}^0}$ with $g \in \mathcal{C}_{\text{pw}}^0$ and $T = \{t_i \in \mathbb{R} \mid i \in \mathbb{Z}\}$ such that D coincides with $g_{\mathbb{D}}$ on each interval (t_i, t_{i+1}) , $i \in \mathbb{Z}$. For any $\tau \in \mathbb{R}$ choose $\varepsilon > 0$ such that $(\tau - \varepsilon, \tau) \subseteq (t_i, t_{i+1})$ for some $i \in \mathbb{Z}$. Then the restriction of D to the interval $[\tau, \infty)$ is defined via

$$D_{[\tau, \infty)}(\varphi) = \begin{cases} 0, & \text{if } \text{supp } \varphi \subseteq (-\infty, \tau], \\ D(\varphi) - \int_{\tau-\varepsilon}^{\tau} g(t)\varphi(t) dt, & \text{if } \text{supp } \varphi \subseteq [\tau - \varepsilon, \infty), \\ D_{[\tau, \infty)}(\varphi^\varepsilon), & \text{otherwise,} \end{cases}$$

where $\varphi^\varepsilon \in \mathcal{C}_0^\infty$ is such that $\varphi = \varphi_\tau + \varphi^\varepsilon$ with $\text{supp } \varphi_\tau \subseteq (-\infty, \tau]$ and $\text{supp } \varphi^\varepsilon \subseteq [\tau - \varepsilon, \infty)$.

It is easily seen that this definition does not depend on the specific choice of φ^ε , hence $D_{[\tau, \infty)}$ is a well defined (continuous) operator on \mathcal{C}_0^∞ and therefore a distribution. In fact, $D_{[\tau, \infty)} \in \mathbb{D}_{\text{pw}\mathcal{C}^0}$ with $g_{[\tau, \infty)}$ as the corresponding piecewise-continuous function. The restriction to the closed interval $(-\infty, \tau]$ is defined analogously, and the restriction to arbitrary intervals can be defined as follows, $s, t \in \mathbb{R} \cup \{\infty\}$:

$$\begin{aligned} D_{(s, t)} &= D - D_{[t, \infty)} - D_{(-\infty, s]}, \\ D_{[s, t]} &= D_{[s, \infty)} - D_{(t, \infty)}, \\ D_{(s, t]} &= D_{[s, \infty)} - D_{[t, \infty)}, \\ D_{[s, t)} &= D_{(s, \infty)} - D_{(t, \infty)}. \end{aligned}$$

It is worth noting that it is not difficult to show that

$$\mathbb{D}_{\text{pw}\mathcal{C}^0} = \left\{ D = g_{\mathbb{D}} + \sum_{t \in T} D_t \left| \begin{array}{l} g \in \mathcal{C}_{\text{pw}}^0, T \subseteq \mathbb{R} \text{ is locally finite, } \forall t \in T \\ \exists n_t \in \mathbb{N}, \alpha_1^t, \dots, \alpha_{n_t}^t \in \mathbb{R} : D_t = \sum_{k=0}^{n_t} \alpha_k^t \delta_t^{(k)} \end{array} \right. \right\}$$

and the restriction of $D \in \mathbb{D}_{\text{pw}\mathcal{C}^0}$ with the above representation $D = g_{\mathbb{D}} + \sum_{t \in T} D_t$ to an interval $I \in \mathbb{R}$ is given by

$$D_I = g_{I\mathbb{D}} + \sum_{t \in T \cap I} D_t.$$

The space of piecewise-continuous distributions also allows a pointwise evaluation in the following three senses, for $t \in \mathbb{R}$ and $D \in \mathbb{D}_{\text{pw}\mathcal{C}^0}$ with corresponding $g \in \mathcal{C}_{\text{pw}}^0$:

- the right sided evaluation: $D(t+) := g(t+)$,
- the left sided evaluation: $D(t-) = g(t-)$,
- the impulsive part: $D[t] := D_{[t, t]}$.

The following relates the restriction with the derivative.

Lemma 5.2 (Derivative of a restriction [13, Prop. 1]) *Let $D \in \mathbb{D}_{\text{pw}\mathcal{C}^0}$ and assume $D' \in \mathbb{D}_{\text{pw}\mathcal{C}^0}$ as well. Then, for any $\tau \in \mathbb{R}$,*

$$(D_{[\tau, \infty)})' = (D')_{[\tau, \infty)} + D(\tau-)\delta_\tau.$$

Note that Cobb did not include the assumption $D' \in \mathbb{D}_{\text{pw}\mathcal{C}^0}$ in his result; however, without this assumption the restriction of D' to some interval is not defined, because in general D' is not a piecewise-continuous distributions anymore (actually Cobb claims that the result is “obvious”; this is quite often a hint that there might be something wrong).

Remark 5.2 (A distributional motivation of Doetsch’s past-aware derivative) Lemma 5.2 now gives a justification of the past-aware derivative (4.10) as propagated by Doetsch, because $D_{[0,\infty)}$ as well as $(D')_{[0,\infty)}$ are elements of the space $\bigcup_k \mathbb{D}_{\geq 0,k}$, however, D can still be non-zero on $(-\infty, 0)$ and $D(0-) \neq 0$ in general.

A connection between (consistent) distributional solution of (1.1) and the solutions of “distributional” DAEs (4.13) was established in [13, Prop. 2], a clearer connection, also allowing for inconsistent initial values, will be formulated in the context of piecewise-smooth distributions (see Sect. 5.4).

5.3 Impulsive-Smooth Distributions as Solution Space

The space of impulsive-smooth distributions was introduced by Hautus [26] (without denoting them as such) and was first used by this name in the context of optimal control problems [27]. Geerts [22–24] was then the first to use them as a solution space for DAEs. The space of impulsive-smooth distributions \mathcal{C}_{imp} is defined in this earlier work as follows:

$$\mathcal{C}_{\text{imp}} := \left\{ D = g_{[0,\infty)\mathbb{D}} + D_{\text{imp}} \mid g \in \mathcal{C}^\infty, D_{\text{imp}} = \sum_{i=0}^k \alpha_i \delta^{(i)}, k \in \mathbb{N}, \alpha_0, \dots, \alpha_k \in \mathbb{R} \right\}.$$

Similar as in the Laplace transform approach, Geerts considers the distributional version (4.13) instead of (2.9) and he rewrites the (distributional) derivative as the convolution with δ' :

$$\delta' * Ex = Ax + f + Ex_0\delta. \tag{5.1}$$

By viewing \mathcal{C}_{imp} as a commutative algebra with convolution as multiplication, the distributional DAE can now be written as

$$pEx = Ax + f + Ex_0,$$

where $p = \delta'$ and δ is the unit with respect to convolution and hence denoted by one. The (time-domain) equation is now algebraically identically to the one obtained by the Laplace transformation approach without the need to think about problems like the existence of the Laplace transform and domain of convergence. In particular, existence and uniqueness results directly apply because no isomorphism between different solution spaces is needed. Nevertheless, the definition of \mathcal{C}_{imp} still assumes that all involved variables are identically zero on $(-\infty, 0)$, hence speaking of inconsistent initial values is conceptually as difficult as for the Laplace transform approach. In summary, viewing x_0 in (5.1) as the initial value for $x(0-)$ cannot be motivated within the impulsive-smooth distributional framework, because, by definition, $x(0-) = 0$.

In fact, there is no reason to consider variables which have to vanish on $(-\infty, 0)$: Rabier and Rheinboldt [37] were the first to use the space of impulsive-smooth distributions which can also be non-zero in the past. The formal definition is

$$\mathcal{C}_{\text{imp}}(\mathbb{R}^*) := \left\{ D = f_{(-\infty, 0)\mathbb{D}}^- + D_{\text{imp}} + f_{(0, \infty)\mathbb{D}}^+ \left| \begin{array}{l} f^-, f^+ \in \mathcal{C}^\infty, D_{\text{imp}} = \sum_{i=0}^k \alpha_i \delta^{(i)}, \\ k \in \mathbb{N}, \alpha_0, \dots, \alpha_k \in \mathbb{R} \end{array} \right. \right\}.$$

Clearly,

$$\mathcal{C}_{\text{imp}} \subset \mathcal{C}_{\text{imp}}(\mathbb{R}^*) \subset \mathbb{D}_{\text{pw}}\mathcal{C}^0 \subset \mathbb{D},$$

in particular, the three types of evaluation defined for piecewise-continuous distributions are also well defined for impulsive-smooth distribution as well as the distributional restriction. The main difference to the space of piecewise-continuous distribution is the fact that the space of impulsive-smooth distribution is closed under differentiation. In particular, impulsive-smooth distributions are arbitrarily often differentiable within the space of impulsive-smooth distributions.

Within the impulsive-smooth distributional framework the ITP (3.2)

$$\begin{aligned} x_{(-\infty, 0)} &= x_{(-\infty, 0)}^0, \\ (E\dot{x})_{[0, \infty)} &= (Ax + f)_{[0, \infty)} \end{aligned}$$

is well defined for all initial trajectories $x^0 \in \mathcal{C}_{\text{imp}}(\mathbb{R}^*)^n$, all inhomogeneities $f \in \mathcal{C}_{\text{imp}}(\mathbb{R}^*)^m$ and solutions x are sought in $\mathcal{C}_{\text{imp}}(\mathbb{R}^*)^n$. In fact, the following result holds, which finally gives a satisfying and rigorous motivation for the incorporation of the (inconsistent) initial value as in (4.13).

Theorem 5.3 (Equivalent description of the ITP (3.2)) *Consider the ITP (3.2) within the impulsive-smooth distributional solution framework with fixed initial trajectory $x^0 \in \mathcal{C}_{\text{imp}}(\mathbb{R}^*)^n$ and inhomogeneity $f \in \mathcal{C}_{\text{imp}}(\mathbb{R}^*)^m$. Then $x \in \mathcal{C}_{\text{imp}}(\mathbb{R}^*)^n$ solves the ITP (3.2) if, and only if, $z := x - x_{(-\infty, 0)}^0 = x_{[0, \infty)}$ solves*

$$\begin{aligned} z_{(-\infty, 0)} &= 0, \\ (E\dot{z})_{[0, \infty)} &= (Az + f)_{[0, \infty)} + Ex^0(0-)\delta. \end{aligned} \tag{5.2}$$

Proof Let x be a solution of the ITP (3.2) and let $z = x_{[0, \infty)}$. Then, clearly, $z_{(-\infty, 0)} = 0$. Furthermore,

$$\begin{aligned} (E\dot{z})_{[0, \infty)} &= (E\dot{x})_{[0, \infty)} - (E(x_{(-\infty, 0)}^0))'_{[0, \infty)} = (Ax + f)_{[0, \infty)} + Ex^0(0-)\delta \\ &= (Az + f)_{[0, \infty)} + Ex^0(0-)\delta, \end{aligned}$$

which shows that $z = x_{[0,\infty)}$ is indeed a solution of (5.2). On the other hand, let z be a solution of (5.2) and define $x := z + x_{(-\infty,0)}^0$. Then, clearly, $x_{(-\infty,0)} = x_{(-\infty,0)}^0$. Furthermore,

$$\begin{aligned} (E\dot{x})_{[0,\infty)} &= (E\dot{z})_{[0,\infty)} + (E(x_{(-\infty,0)}^0))'_{[0,\infty)} \\ &= (Az + f)_{[0,\infty)} + Ex^0(0-)\delta - Ex^0(0-)\delta \\ &= (Ax + f)_{[0,\infty)}. \end{aligned} \quad \square$$

Remark 5.3

1. If (5.2) is considered within the one-sided impulsive-smooth distributional framework, i.e. $f \in (\mathcal{C}_{\text{imp}})^m$ and $z \in (\mathcal{C}_{\text{imp}})^n$ then (5.2) simplifies to

$$E\dot{z} = Az + f + Ex^0(0-)\delta. \tag{5.3}$$

2. Comparing the result of Theorem 5.3 with the result of Cobb [13, Prop. 2] reveals three main differences: (1) Cobb only states one direction and not the equivalence, (2) instead of the ITP (3.2) Cobb just considers the original DAE (1.1), hence his result concerns only consistent solutions, (3) Cobb assumes that (5.3) has a unique solution.
3. Regularity of the matrix pair (E, A) is not assumed; in particular, neither is it assumed that for all inhomogeneities f there exist solutions to (3.2) and (5.2), nor is it assumed that solutions of (3.2) and (5.2) are uniquely given for fixed initial trajectory and fixed inhomogeneity. However, due to the established equivalence all existence and uniqueness results obtained for (5.3) carry over to the ITP (3.2).

Although Rabier and Rheinboldt [37] introduced the space of impulsive-smooth distribution which allow a clean treatment of the ITP (3.2), they did not follow this approach. Instead, they redefine the inhomogeneity to make inconsistent initial values consistent. To this end, let $x^0 \in \mathcal{C}_{\text{imp}}(\mathbb{R}^*)^n$ be a given initial trajectory and $f \in \mathcal{C}_{\text{imp}}(\mathbb{R}^*)^m$ a given inhomogeneity and consider the ITP-DAE

$$\begin{aligned} x_{(-\infty,0)} &= x_{(-\infty,0)}^0, \\ E\dot{x} &= Ax + f_{\text{ITP}}, \end{aligned} \tag{5.4}$$

where

$$f_{\text{ITP}} := E\dot{x}_{(-\infty,0)}^0 - Ax_{(-\infty,0)}^0 + f_{[0,\infty)}.$$

Note that $x_{(-\infty,0)} = x_{(-\infty,0)}^0$ already implies, due to the special choice of f_{ITP} , that

$$(E\dot{x})_{(-\infty,0)} = (Ax + f_{\text{ITP}})_{(-\infty,0)},$$

which shows that (5.4) is in fact equivalent to the ITP (3.2). However, the form of (5.4) has certain disadvantages compared to the ITP formulation (3.2):

1. The second equation of (5.4) suggest that the DAE (1.1) is valid globally (just with a different inhomogeneity), which conflicts with the intuition that an inconsistent initial value is due to the fact that the system description (1.1) is only valid on $[0, \infty)$ and not in the past.
2. In (5.4) the past trajectory of x is formally determined by two equations which could in general be conflicting (depending on the choice of f_{ITP}).
3. When studying an autonomous system (i.e. without the presence of an inhomogeneity), the formulation (5.4) formally leaves the class of autonomous systems.

On the other hand, an interesting advantage of the formulation (5.4) is that, due to Remark 5.1, (5.4) makes sense even when x is an arbitrary distribution and f as well as x^0 are such that f_{ITP} is well defined. In fact, Rabier and Rheinboldt [37, Thm. 4.1] do consider arbitrary distributions $x \in \mathbb{D}^n$ and show that under certain regularity assumptions the solutions are in fact impulsive-smooth.

5.4 Piecewise-Smooth Distributions as Solution Space

Comparing Cobb’s piecewise-continuous distributional solution framework with the impulsive-smooth distributional solution framework the following differences are apparent:

1. $\mathbb{D}_{\text{pw}}\mathcal{C}^0$ is not closed under differentiation.
2. $\mathcal{C}_{\text{imp}}(\mathbb{R}^*)$ does not allow non-smooth inhomogeneities away from $t = 0$.

Rabier and Rheinboldt [37] seem to be aware of the latter problem as they introduce the space $\mathcal{C}_{\text{imp}}(\mathbb{R} \setminus \mathcal{S})$, where $\mathcal{S} = \{t_i \in \mathbb{R} | i \in \mathbb{Z}\}$ is a strictly ordered set with $t_i \rightarrow \pm\infty$ as $i \rightarrow \pm\infty$ and $D \in \mathcal{C}_{\text{imp}}(\mathbb{R} \setminus \mathcal{S})$ is such that $D_{(t_i, t_{i+1})}$ is induced by the corresponding restriction of a smooth function. A similar idea is proposed in [25], however, in both cases the resulting distributional space is not studied in detail. A more detailed treatment can be found in [45, 46] where, in the spirit of Cobb’s definition, the space of piecewise-smooth distributions is defined as follows:

$$\mathbb{D}_{\text{pw}}\mathcal{C}^\infty := \left\{ D = f_{\mathbb{D}} + \sum_{t \in T} D_t \mid \begin{array}{l} f \in \mathcal{C}_{\text{pw}}^\infty, T \subseteq \mathbb{R} \\ \text{locally finite } \forall t \in T : D_t \in \text{span}\{\delta_t, \delta'_t, \delta''_t, \dots\} \end{array} \right\},$$

where $f \in \mathcal{C}_{\text{pw}}^\infty$ is a piecewise-smooth function if, and only if, there exists a strictly ordered locally finite set $\{s_i \in \mathbb{R} | i \in \mathbb{Z}\}$ and $f_i \in \mathcal{C}^\infty, i \in \mathbb{Z}$, such that $f = \sum_{i \in \mathbb{Z}} f_i|_{[s_i, s_{i+1})}$. Clearly,

$$\mathcal{C}_{\text{imp}}(\mathbb{R}^*) \subset \mathbb{D}_{\text{pw}}\mathcal{C}^\infty \subset \mathbb{D}_{\text{pw}}\mathcal{C}^0,$$

and the space of piecewise-smooth distributions resolves each of the above mentioned drawbacks of the piecewise-continuous and impulsive-smooth distributions.

However, the major advantage of considering the space of piecewise-smooth distributions becomes apparent when considering time-varying DAEs:

$$E(t)\dot{x}(t) = A(t)x(t) + f(t). \tag{5.5}$$

If the coefficient matrices $E(\cdot)$ and $A(\cdot)$ are smooth it is no problem to use any of the above distributional solution concepts because the product of a smooth function with any distribution is well defined so that (5.5) makes sense as an equation of distributions. In the discussion of the drawbacks of the Laplace transform approach it was already mentioned that an inconsistent initial value could be seen as the results from the presence of a time-varying system. In fact, the ITP (3.2) can be reformulated as the following time-varying DAE [45, Thm. 3.1.7]:

$$E_{\text{ITP}}(t)\dot{x}(t) = A_{\text{ITP}}(t)x(t) + f_{\text{ITP}}(t),$$

where

$$E_{\text{ITP}}(t) = \begin{cases} 0, & t < 0, \\ E, & t \geq 0, \end{cases} \quad A_{\text{ITP}}(t) = \begin{cases} I, & t < 0, \\ A, & t \geq 0, \end{cases}$$

$$f_{\text{ITP}}(t) = \begin{cases} -x^0(t), & t < 0, \\ f(t), & t \geq 0. \end{cases}$$

The problem is now that the time-varying coefficient matrices are not smooth anymore so that the multiplication with a distribution is not well defined. Rabier and Rheinboldt [37] treated already time-varying DAEs (5.5); however, the interpretation of inconsistent initial values as a time-variant DAE with non-smooth coefficients did not occur to them, maybe because they considered (5.4) where formally the original DAE (with a special choice of the inhomogeneity) with smooth coefficient is considered globally (i.e. in the whole of \mathbb{R} and not only on $[0, \infty)$). Another important motivation for studying time-varying DAEs with non-smooth coefficient matrices is switched DAEs [47]:

$$E_\sigma \dot{x} = A_\sigma x + f,$$

where $\sigma : \mathbb{R} \rightarrow \{1, 2, \dots, P\}$, $P \in \mathbb{N}$, and $(E_1, A_1), \dots, (E_P, A_P)$ are constant matrices.

It turns out that for the space of piecewise-smooth distributions a (non-commutative) multiplication can be defined, named *Fuchssteiner multiplication* after [19, 20], which in particular defines the multiplication of a piecewise-smooth function with a piecewise-smooth distribution. Hence (5.5) makes sense even for coefficient matrices which are only piecewise-smooth.

Remark 5.4 (The square of the Dirac impulse) The multiplication of distributions occurs several times in the context of DAEs. The different approaches can be best illustrated by the different treatments of the square of the Dirac impulse:

1. In the context of impulsive-smooth distributions [22, 23, 27] convolution is viewed as a multiplication and the Dirac impulse is the unit element for that multiplication. Hence $\delta^2 = \delta$ in this framework.
2. The Fuchssteiner multiplication for piecewise-smooth distributions yields

$$\delta^2 = 0.$$

3. It is well known that a commutative and associative multiplication which generalizes the multiplication of functions to distributions is not possible in general, but when enlarging the space of distributions the square of the Dirac impulse is well defined (but not a classical distribution). In the context of DAEs this approach was considered in [44], where the square of the Dirac impulse occurs in the analysis of the connection energy (the product of the voltage and current).

Within the framework of piecewise-smooth distributions it is now possible to show [45] that the ITP (3.2) is uniquely solvable for all initial trajectories and all inhomogeneities if, and only if, the matrix pair (E, A) is regular. In particular, the impulses and jumps derived in this framework [47, Thm. 6.5.1] are identical to (4.15) and (4.16) obtained via the Laplace transform approach.

6 Conclusion

The role of the Wong sequences of the matrix pair (E, A) for characterizing the (classical) solutions was highlighted. In particular, explicit solution formulas were given which are similar to the ones obtained for linear ODEs. The quasi-Kronecker form (QKF) and quasi-Weierstraß form (QWF) play a prominent role. For time-varying DAEs with analytical coefficients a time-varying QWF is available, however, time-varying Wong sequences and their connection to a time-varying QWF (or even QKF) have not been studied yet. The problem of inconsistent initial values was discussed and it was shown how the Laplace transform was used to treat this problem. However, it is argued that the Laplace transform approach cannot justify the notion of an inconsistent initial value. With the help of certain distributional solution spaces the notion of inconsistent initial values can be treated in a satisfying way and it also justifies the Laplace transform approach.

References

1. Aplevich, J.D.: *Implicit Linear Systems*. Lecture Notes in Control and Information Sciences, vol. 152. Springer, Berlin (1991)
2. Armentano, V.A.: The pencil $(sE - A)$ and controllability-observability for generalized linear systems: a geometric approach. *SIAM J. Control Optim.* **24**, 616–638 (1986)
3. Berger, T., Trenn, S.: Addition to: “The quasi-Kronecker form for matrix pencils”. *SIAM. J. Matrix Anal. Appl.* **34**(1), 94–101 (2013)

4. Berger, T., Trenn, S.: The quasi-Kronecker form for matrix pencils. *SIAM J. Matrix Anal. Appl.* **33**(2), 336–368 (2012)
5. Berger, T., Ilchmann, A., Trenn, S.: The quasi-Weierstraß form for regular matrix pencils. *Linear Algebra Appl.* **436**(10), 4052–4069 (2012). doi:[10.1016/j.laa.2009.12.036](https://doi.org/10.1016/j.laa.2009.12.036)
6. Bernhard, P.: On singular implicit linear dynamical systems. *SIAM J. Control Optim.* **20**(5), 612–633 (1982)
7. Brenan, K.E., Campbell, S.L., Petzold, L.R.: *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*. North-Holland, Amsterdam (1989)
8. Campbell, S.L.: *Singular Systems of Differential Equations I*. Pitman, New York (1980)
9. Campbell, S.L.: *Singular Systems of Differential Equations II*. Pitman, New York (1982)
10. Campbell, S.L., Petzold, L.R.: Canonical forms and solvable singular systems of differential equations. *SIAM J. Algebr. Discrete Methods* **4**, 517–521 (1983)
11. Campbell, S.L., Meyer, C.D. Jr., Rose, N.J.: Applications of the Drazin inverse to linear systems of differential equations with singular constant coefficients. *SIAM J. Appl. Math.* **31**(3), 411–425 (1976). <http://link.aip.org/link/?SMM/31/411/1>. doi:[10.1137/0131035](https://doi.org/10.1137/0131035)
12. Cobb, J.D.: On the solution of linear differential equations with singular coefficients. *J. Differ. Equ.* **46**, 310–323 (1982)
13. Cobb, J.D.: Controllability, observability and duality in singular systems. *IEEE Trans. Autom. Control* **AC-29**, 1076–1082 (1984)
14. Dai, L.: *Singular Control Systems*. Lecture Notes in Control and Information Sciences, vol. 118. Springer, Berlin (1989)
15. Dieudonné, J.: Sur la réduction canonique des couples des matrices. *Bull. Soc. Math. Fr.* **74**, 130–146 (1946)
16. Doetsch, G.: *Introduction to the Theory and Application of the Laplace Transformation*. Springer, Berlin (1974)
17. Drazin, M.P.: Pseudo-inverses in associative rings and semigroups. *Am. Math. Mon.* **65**(7), 506–514 (1958)
18. Frasca, R., Çamlıbel, M.K., Goknar, I.C., Iannelli, L., Vasca, F.: Linear passive networks with ideal switches: consistent initial conditions and state discontinuities. *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.* **57**(12), 3138–3151 (2010)
19. Fuchssteiner, B.: Eine assoziative Algebra über einen Unterraum der Distributionen. *Math. Ann.* **178**, 302–314 (1968)
20. Fuchssteiner, B.: Algebraic foundation of some distribution algebras. *Stud. Math.* **76**, 439–453 (1984)
21. Gantmacher, F.R.: *The Theory of Matrices*, vols. I & II. Chelsea, New York (1959)
22. Geerts, A.H.W.T.: Invariant subspaces and invertibility properties for singular systems: the general case. *Linear Algebra Appl.* **183**, 61–88 (1993). doi:[10.1016/0024-3795\(93\)90424-M](https://doi.org/10.1016/0024-3795(93)90424-M)
23. Geerts, A.H.W.T.: Solvability conditions, consistency and weak consistency for linear differential-algebraic equations and time-invariant linear systems: the general case. *Linear Algebra Appl.* **181**, 111–130 (1993)
24. Geerts, A.H.W.T.: Regularity and singularity in linear-quadratic control subject to implicit continuous-time systems. *IEEE Proc. Circuits Syst. Signal Process.* **13**, 19–30 (1994)
25. Geerts, A.H.W.T., Schumacher, J.M.H.: Impulsive-smooth behavior in multimode systems. Part I: state-space and polynomial representations. *Automatica* **32**(5), 747–758 (1996)
26. Hautus, M.L.J.: The formal Laplace transform for smooth linear systems. In: Marchesini, G., Mitter, S.K. (eds.) *Mathematical Systems Theory. Lecture Notes in Economics and Mathematical Systems*, vol. 131, pp. 29–47. Springer, New York (1976)
27. Hautus, M.L.J., Silverman, L.M.: System structure and singular control. *Linear Algebra Appl.* **50**, 369–402 (1983)
28. Kronecker, L.: Algebraische Reduction der Schaaren bilinearer Formen. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften zu Berlin*, pp. 1225–1237 (1890)
29. Kuijper, M.: *First-Order Representations of Linear Systems*. Birkhäuser, Boston (1994)
30. Kunkel, P., Mehrmann, V.: *Differential-Algebraic Equations. Analysis and Numerical Solution*. EMS Publishing House, Zürich (2006)

31. Lamour, R., März, R., Tischendorf, C.: Differential Algebraic Equations: A Projector Based Analysis. *Differential-Algebraic Equations Forum*, vol. 1. Springer, Heidelberg (2013)
32. Lewis, F.L.: A survey of linear singular systems. *IEEE Proc. Circuits Syst. Signal Process.* **5**(1), 3–36 (1986)
33. Liberzon, D., Trenn, S.: Switched nonlinear differential algebraic equations: solution theory, Lyapunov functions, and stability. *Automatica* **48**(5), 954–963 (2012). doi:[10.1016/j.automatica.2012.02.041](https://doi.org/10.1016/j.automatica.2012.02.041)
34. Lundberg, K.H., Miller, H.R., Trumper, D.L.: Initial conditions, generalized functions, and the Laplace transform. *IEEE Control Syst. Mag.* **27**(1), 22–35 (2007). doi:[10.1109/MCS.2007.284506](https://doi.org/10.1109/MCS.2007.284506)
35. Opal, A., Vlach, J.: Consistent initial conditions of linear switched networks. *IEEE Trans. Circuits Syst.* **37**(3), 364–372 (1990)
36. Owens, D.H., Debeljkovic, D.L.: Consistency and Liapunov stability of linear descriptor systems: a geometric analysis. *IMA J. Math. Control Inf.* **2**, 139–151 (1985)
37. Rabier, P.J., Rheinboldt, W.C.: Time-dependent linear DAEs with discontinuous inputs. *Linear Algebra Appl.* **247**, 1–29 (1996)
38. Rabier, P.J., Rheinboldt, W.C.: Theoretical and numerical analysis of differential-algebraic equations. In: Ciarlet, P.G., Lions, J.L. (eds.) *Handbook of Numerical Analysis*, vol. VIII, pp. 183–537. Elsevier, Amsterdam (2002)
39. Reißig, G., Boche, H., Barton, P.I.: On inconsistent initial conditions for linear time-invariant differential-algebraic equations. *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.* **49**(11), 1646–1648 (2002)
40. Rosenbrock, H.H.: *State Space and Multivariable Theory*. Wiley, New York (1970)
41. Schwartz, L.: *Théorie des Distributions*. Hermann, Paris (1957, 1959)
42. Sincovec, R.F., Erisman, A.M., Yip, E.L., Epton, M.A.: Analysis of descriptor systems using numerical algorithms. *IEEE Trans. Autom. Control* **AC-26**, 139–147 (1981)
43. Tanwani, A., Trenn, S.: On observability of switched differential-algebraic equations. In: *Proc. 49th IEEE Conf. Decis. Control*, Atlanta, USA, pp. 5656–5661 (2010)
44. Tolsa, J., Salichs, M.: Analysis of linear networks with inconsistent initial conditions. *IEEE Trans. Circuits Syst.* **40**(12), 885–894 (1993). doi:[10.1109/81.269029](https://doi.org/10.1109/81.269029)
45. Trenn, S.: *Distributional differential algebraic equations*. Ph.D. thesis, Institut für Mathematik, Technische Universität Ilmenau, Universitätsverlag Ilmenau, Ilmenau, Germany (2009). <http://www.db-thueringen.de/servlets/DocumentServlet?id=13581>
46. Trenn, S.: A normal form for pure differential algebraic systems. *Linear Algebra Appl.* **430**(4), 1070–1084 (2009). doi:[10.1016/j.laa.2008.10.004](https://doi.org/10.1016/j.laa.2008.10.004)
47. Trenn, S.: Switched differential algebraic equations. In: Vasca, F., Iannelli, L. (eds.) *Dynamics and Control of Switched Electronic Systems—Advanced Perspectives for Modeling, Simulation and Control of Power Converters*, pp. 189–216. Springer, London (2012). Chap. 6
48. Verghese, G.C., Levy, B.C., Kailath, T.: A generalized state-space for singular systems. *IEEE Trans. Autom. Control* **AC-26**(4), 811–831 (1981)
49. Weierstraß, K.: *Zur Theorie der bilinearen und quadratischen Formen*. Berl. Monatsb., pp. 310–338 (1868)
50. Wilkinson, J.H.: Linear differential equations and Kronecker’s canonical form. In: de Boor, C., Golub, G.H. (eds.) *Recent Advances in Numerical Analysis*, pp. 231–265. Academic Press, New York (1978)
51. Wong, K.T.: The eigenvalue problem $\lambda T x + S x$. *J. Differ. Equ.* **16**, 270–280 (1974)
52. Yip, E.L., Sincovec, R.F.: Solvability, controllability and observability of continuous descriptor systems. *IEEE Trans. Autom. Control* **AC-26**, 702–707 (1981)