12

INSTITUTE FOR COMPUTATIONAL MATHEMATICS AND APPLICATIONS

Technical Report ICMA-93-183 October, 1993

Classical and Generalized Solutions of Time-Dependent Linear Differential Algebraic Equations[1]

BY

PATRICK J. RABIER AND WERNER C. RHEINBOLDT[2]

Department of Mathematics and Statistics

University of Pittsburgh

12

Technical Report ICMA-93-183 October, 1993

Classical and Generalized Solutions of Time-Dependent Linear Differential Algebraic Equations[1]

BY

PATRICK J. RABIER AND WERNER C. RHEINBOLDT[2]

[1] The work was supported in part by ONR-grant N-00014-90-J-1025, and NSF-grant CCR-9203488.

[2] Department of Mathematics and Statistics, University of Pittsburgh, Pittsburgh, PA 15260.

CLASSICAL AND GENERALIZED SOLUTIONS OF TIME-DEPENDENT LINEAR DIFFERENTIAL ALGEBRAIC EQUATIONS[1]

BY

PATRICK J. RABIER AND WERNER C. RHEINBOLDT[2]

ABSTRACT. A reduction procedure is developed for linear time-dependent differential algebraic equations (DAEs) that transforms their solutions into solutions of smaller systems of ordinary differential equations (ODEs). The procedure applies to classical as well as distribution solutions. In the case of analytic coefficients the hypotheses required for the reduction not only are necessary for the validity of the existence and uniqueness results, but they even allow for the presence of singularities. Straightforward extensions including undetermined systems and systems with non-analytic coefficients are also discussed.

1. Introduction.

A time-dependent linear differential-algebraic equation (DAE) is a problem of the form

$$A(t)\dot{x} + B(t)x = b(t), \quad t \in \mathcal{J}, \tag{1.1}$$

where $\mathcal{J} \subset \mathbb{R}$ is an open interval, $A(t), B(t) \in \mathcal{L}(\mathbb{R}^n)$ and $A(t)$ is *not* invertible for any $t \in \mathcal{J}$. In the simplest case, when $A(t)$ is diagonal, this condition implies that - at least locally - some of the scalar equations in the system (1.1) are purely algebraic, which justifies the terminology "differential-algebraic equation". In order to provide a perspective to our work, we begin with a brief review of what is currently known about such problems without, however, including some earlier contributions now subsumed by more general theories.

A significant simplification is the case when $A(t)$ and $B(t)$ in (1.1) are constant. The corresponding existence and uniqueness theory for initial value problems has been known

[1] The work was supported in part by ONR-grant N-00014-90-J-1025, and NSF-grant CCR-9203488
[2] Department of Mathematics and Statistics, University of Pittsburgh, Pittsburgh, PA 15260

for some time: A full account can be found in Griepentrog and Maerz [GrM86] who credit Gantmacher's theory of matrix pencils, [G59]. The book [GrM86] also contains a treatment of the general system (1.1) utilizing a condition of "transferability" which, broadly speaking, corresponds to a generalization of the index 1 case of the constant coefficients theory. Subsequent generalizations to the case of the index 2 and 3 have been obtained by Maerz ([M89a], [M89b]).

A different approach is taken in Campbell and Petzold [CP83] who consider "solvable" systems with analytic coefficients. The solvability assumption requires that an existence and uniqueness theory similar to that for linear ODE's is available. The result of [CP83] is that such systems can be transformed into the form

$$\dot{y} + C(t)y = f(t), \quad t \in \mathcal{J}, \tag{1.2a}$$

$$N(t)\dot{z} + z = g(t), \quad t \in \mathcal{J}, \tag{1.2b}$$

where $C(t)$ and $N(t)$ are analytic functions of t and $N(t)$ is nilpotent upper (or lower) triangular for all $t \in \mathcal{J}$. From the structure of $N(t)$, it follows that the operator $N(t)\frac{d}{dt}$ is nilpotent, so that (1.2b) has the *unique* solution $z = \sum_{k=1}^{n} (-1)^k (N(t)\frac{d}{dt})^k g$, and (1.2a) is an explicit ODE. But no calculable criterion for solvability is given in that paper.

Later, Campbell [C87] proved results of similar type for the case of non-analytic coefficients in (1.1) and provided a calculable necessary and sufficient criterion for solvability. This condition requires, in particular, the characterization of the null-spaces of some augmented systems for each $t \in \mathcal{J}$. He also showed that the solutions of (1.1) can be obtained as the solutions of an explicit ODE in $\mathbb{R}^n$ satisfying consistent initial conditions.

Recently, Kunkel and Mehrmann [KuMe92] have taken a different approach. Under various constant-rank conditions they show that the solutions of (1.1) can be obtained in the form $x = Uy$ where U is a smoothly parametrized family of linear isomorphisms of $\mathbb{R}^n$ and $y = (y_1, y_2, y_3)^T$ solves a simple DAE of the form $\dot{y}_1 = c_1(t)$, $y_2 = c_2(t)$, $0 = c_3(t)$, where $c = (c_1, c_2, c_3)^T$ depends upon b and its derivatives. Solvability requires the dimension of the third variable to be zero, for otherwise y_3 is arbitrary if $c_3 = 0$ or no solution exists when $c_3 \neq 0$, and hence either uniqueness or existence fails to hold. This approach does not require a solvability assumption, for the dimension of the third variable is determined in the process of finding U. On the other hand, when the DAE is solvable, the hypotheses made in [KuMe92] to reach this conclusion are sufficient but not neccesary.

Finally, we mention new theories for general implicit DAE's $F(t, x, \dot{x}) = 0$ developed by the authors ([RRh91], [RRh92]), that can be applied to linear problems. But, of course, these general theories do not capture the special features due to the linear structure.

In all mentioned results only classical solutions are considered; that is, all solutions $x : \mathcal{J} \to \mathbb{R}^n$ of (1.1) are assumed to be (at least) differentiable. As far as we know, the problem of characterizing generalized (distribution) solutions of (1.1), when the right-hand side b is a distribution, has been successfully investigated only for constant A and B and in special cases (see [VLKa81], [Co82], [Ge93]). The difficulty with distribution solutions is that arguments from the "classical" theory involving the numerical values of the solutions can no longer be used. This is the reason why most approaches do not extend to distributions. It appears that the method in [KuMe92] may allow for such an extension, but there is no comment to that extent in that paper.

In this paper, we focus mainly on the case when A and B are analytic. Under computationally verifiable assumptions – weaker than the solvability condition of [CP83] – we prove that both the classical *and* distribution solutions of (1.1) are of the form $x(t) = \Gamma(t)x_\nu(t) + v(t)$ where x_ν solves a smaller system

$$A_\nu \dot{x}_\nu + B_\nu x_\nu = b_\nu, \quad \text{in } \mathcal{J}, \tag{1.3}$$

and $A_\nu(t)$ is invertible for all but isolated values of t. It turns out that, when (1.1) is solvable in the sense of [CP83], (1.3) can also be derived by the method briefly sketched at the end of that paper. But extending this method without the solvability hypothesis requires a much closer investigation of its mechanism, especially in the distribution case.

Equation (1.3) is obtained at the ν^{th} step of a recursive process involving a number of arbitrary choices. We show, however, that both the so-called index ν of (1.1) and any singularities of A_ν are independent of these choices. We also prove that invertibility of $A_\nu(t)$ for each $t \in \mathcal{J}$ is necessary and sufficient for solvability in the sense of [CP83]. But since solvability is not a prerequisite for the validity of our reduction, singularities may exist in (1.3) and the methods of the Fuchs-Frobenius theory are available for the study of the related phenomena.

We also outline a more or less straightforward generalization of our reduction procedure to characterize the classical solutions of *any* analytic system (1.1), although either existence or uniqueness (or both) for corresponding initial value problems is necessarily affected whenever the hypotheses of the main case break down. Distribution solutions can be characterized as well for more general, but perhaps not all non-analytic systems.

The inherent globality of the reduction is especially important to handle boundary value problems, which reduce to boundary value problems for (1.3).

Finally we examine problems (1.1) in the case when A and B are sufficiently smooth but no longer analytic. It turns out that then additional hypotheses are needed (except in the index 1 case) to compensate for the lack of analyticity. Because of this, we are not able to cover fully the "solvable case" of [C87]. On the other hand, our method remains valid with weaker smoothness hypotheses, say C^n, for A, B and b in (1.1), continues to be applicable for distribution solutions, and can be easily generalized to handle undetermined problems satisfying a Fredholm alternative.

This work and [KuMe92] are certainly close in spirit, but major differences exist in the execution of the main idea. Our reduction appears to be technically much simpler and allows for clearer proofs. More importantly, it is this relative simplicity that permitted a deeper analysis under weaker hypotheses.

Our approach makes crucial use of Kato's "transformation functions", and the next section presents a brief review and some applications of this concept. Very closely related but somewhat less precise results can be found in [D64] and [SiB70]. Throughout the paper analytic mappings will be referred to as "mappings of class C^ω".

2. Transformation functions.

A main tool for the proof of the globality of the reduction procedure is the concept of "transformation functions" introduced by T. Kato [K50]. We use only the following simplified version:

Theorem 2.1. *Let $\mathcal{J} \subset \mathbb{R}$ be an open interval and $P \in C^k(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$, $1 \leq k \leq \infty$ or $k = \omega$, a family of orthogonal projections onto $W(t) = \operatorname{rge} P(t)$, $t \in \mathcal{J}$. Then, for any fixed $t_0 \in \mathcal{J}$, there exists a mapping $U \in C^k(\mathcal{J}; O(n))$, where $O(n)$ denotes the orthogonal group of $\mathbb{R}^n$, such that*

$$P(t) = U(t)P(t_0)U(t)^T, \qquad \forall t \in \mathcal{J}. \tag{2.1}$$

Hence $U(t)P(t_0)$ is a linear isomorphism of $W(t_0)$ onto $W(t)$ for $t \in \mathcal{J}$.

The proof of a more general result involving arbitrary projections and a complex variable in place of t may be found in [K82] (p. 113). It will be useful to outline the proof of the somewhat simpler Theorem 2.1. The transformation U is determined as the solution of the linear initial value problem

$$\dot{U} = [\dot{P}, P]U, \tag{2.2}$$

$$U(t_0) = I_n, \tag{2.3}$$

where $[\dot{P}, P] = \dot{P}P - P\dot{P}$ is the commutator of $\dot{P}$ and P. Because of $P^T = P$, the commutator $[\dot{P}, P]$ is skew symmetric whence $\frac{d}{dt}(U^TU) = U^T\dot{U} + \dot{U}^TU = U^T[\dot{P}, P]U - U^T[\dot{P}, P]U = 0$. Therefore, with (2.3), it follows that $U(t)^TU(t) = I_n$ and hence that $U(t) \in O(n)$ for $t \in \mathcal{J}$. For the proof of (2.1) we use $P^2 = P$ which yields $P\dot{P} + \dot{P}P = \dot{P}$ and thus $P\dot{P}P = 0$ and $[\dot{P}, P]P = \dot{P}P$, $P[\dot{P}, P] = -P\dot{P}$. With this we find for $Y = PU$ that $\dot{Y} = \dot{P}U + P\dot{U} = (\dot{P} - P\dot{P})U = \dot{P}PU = [\dot{P}, P]Y$ which shows that Y solves (2.2) with the initial condition $Y(t_0) = P(t_0)$. Since the general solution X of (2.2) is given by $X = UX(t_0)$, it follows that $Y = UP(t_0)$ and, therefore, that $PU = UP(t_0)$ which is (2.1).

For later use we observe that the above relation $\dot{Y} = \dot{P}PU$ also reads $\dot{Y} = \dot{P}Y$ and hence that Y satisfies the differential equation

$$\dot{Y} = \dot{P}Y. \tag{2.4}$$

From $P^T = P$ we infer that $\dot{P}^T = \dot{P}$, so that, by (2.4), Y^T satisfies the differential equation

$$\dot{Y}^T = Y^T\dot{P}. \tag{2.5}$$

Our applications of Theorem 2.1 rest on the following result (see also Lemma 2.2).

Lemma 2.1. *Let $\mathcal{J} \subset \mathbb{R}$ be an open interval, $n > 0$ an arbitrary integer, and $M \in C^k(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$, $1 \le k \le \infty$ or $k = \omega$, such that rank $M(t) = r$, $\forall t \in \mathcal{J}$. Then, the orthogonal projections onto rge $M(t)$ and $\ker M(t)$, respectively, are cf class C^k in $\mathcal{J}$.*

Proof. Let $\hat{t} \in \mathcal{J}$ be given and consider an orthonormal basis $\{\hat{u}_1, \dots, \hat{u}_r\}$ of rge $M(\hat{t})$. Let $v_1, \dots, v_r$ be linearly independent vectors such that $\hat{u}_i = M(\hat{t})v_i$, $1 \le i \le r$. The vectors $u_i(t) = M(t)v_i$ remain linearly independent for t near $\hat{t}$, say, for t in the open interval $\mathcal{J}_{\hat{t}} \subset \mathcal{J}$ containing $\hat{t}$. In other words, since rank $M(t) = r$, they form a basis of rge $M(t)$ for $t \in \mathcal{J}_{\hat{t}}$, and they are of class C^k in $\mathcal{J}_{\hat{t}}$ since this is true for M. By applying the Gramm-Schmidt process to $u_1(t), \dots, u_r(t)$, we obtain an orthonormal basis $\{w_1(t), \dots, w_r(t)\}$ of rge $M(t)$ for $t \in \mathcal{J}_{\hat{t}}$. Since the Gramm–Schmidt process involves only algebraic operations, the vectors $w_1(t), \dots, w_r(t)$ are again C^k functions of $t \in \mathcal{J}_{\hat{t}}$. Because the orthogonal projection of $\mathbb{R}^n$ onto rge $M(t)$ can be expressed as a sum of dyadic products of $w_1(t), \dots, w_r(t)$, it follows that it is of class C^k in $\mathcal{J}_{\hat{t}}$ and, hence also in $\mathcal{J}$ since $\hat{t} \in \mathcal{J}$ was arbitrary. The corresponding result for the orthogonal projection from $\mathbb{R}^n$ onto $\ker M(t)$ follows directly by replacing $M(t)$ with $M(t)^T$ in the above argument. □

Theorem 2.2. *Let $\mathcal{J} \subset \mathbb{R}$ be an open interval, $n \ge 0$ an arbitrary integer, and $M \in C^k(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$, $1 \le k \le \infty$, such that rank $M(t) = r$ for all $t \in \mathcal{J}$. Then the following operators exist:*

(i) $S \in C^k(\mathcal{J}; \mathcal{L}(\mathbb{R}^r, \mathbb{R}^n))$ such that $S(t) \in GL(\mathbb{R}^r,$ rge $M(t))$, $\forall t \in \mathcal{J}$,

(ii) $T \in C^k(\mathcal{J}; \mathcal{L}(\mathbb{R}^{n-r}, \mathbb{R}^n))$ such that $T(t) \in GL(\mathbb{R}^{n-r}, \ker M(t))$, $\forall t \in \mathcal{J}$,

(iii) $V \in C^k(\mathcal{J}; \mathcal{L}(\mathbb{R}^n, \mathbb{R}^r))$ such that $V(t)_{|\text{rge } M(t)} \in GL(\text{rge } M(t), \mathbb{R}^r)$, $\forall t \in \mathcal{J}$, where V can be chosen so that $\ker V(t) = [\text{ rge } M(t)]^\perp$ for all $t \in \mathcal{J}$.

Proof. (i) Denote by $P(t)$ the orthogonal projection from $\mathbb{R}^n$ onto rge $M(t)$, $t \in \mathcal{J}$. By Lemma 2.1, we have $P \in C^k(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$. Consequently, the mapping $U \in C^k(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$ of Theorem 2.1 satisfies $U(t)P(t_0) \in GL(\text{rge } M(t_0), \text{ rge } M(t))$, $\forall t \in \mathcal{J}$. Now, choose $S_0 \in \mathcal{L}(\mathbb{R}^r, \mathbb{R}^n)$ such that $S_0(\mathbb{R}^r) =$ rge $M(t_0)$ and set $S(t) = U(t)S_0$, $\forall t \in \mathcal{J}$.

(ii) Let $\Pi(t)$ be the orthogonal projection from $\mathbb{R}^n$ onto $\ker M(t)$, $t \in \mathcal{J}$. By Lemma 2.1, we have $\Pi \in C^k(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$. Hence the corresponding mapping $U \in C^k(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$

of Theorem 2.1 satisfies $U(t)\Pi(t_0) \in GL(\ker M(t_0), \ker M(t))$, $\forall t \in \mathcal{J}$. Now, choose $T_0 \in \mathcal{L}(\mathbb{R}^{n-r}, \mathbb{R}^n)$ such that $T_0(\mathbb{R}^{n-r}) = \ker M(t_0)$ and set $T(t) = U(t)T_0$, $\forall t \in \mathcal{J}$.

(iii) With P and U as in the proof of part (i), let $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^r)$ be such that $L(\text{rge } M(t_0)) = \mathbb{R}^r$. Then, set $V(t) = LP(t_0)U(t)^T$, $\forall t \in \mathcal{J}$. It is readily seen that $V(t)|_{\text{rge } M(t)} \in GL(\text{ rge } M(t), \mathbb{R}^r)$. Furthermore, $\ker V(t) = \ker P(t_0)U(t)^T = \{x \in \mathbb{R}^\ell : U(t)^T x \in [\text{rge } M(t_0)]^\perp\} = U(t)\, [\text{rge } M(t_0)]^\perp$. As $U(t) \in O(n)$ maps rge $M(t_0)$ onto rge $M(t)$, it also maps $[\text{rge } M(t_0)]^\perp$ onto $[\text{rge } M(t)]^\perp$, and hence $\ker V(t) = [\text{rge } M(t)]^\perp$, $\forall t \in \mathcal{J}$, as claimed. $\square$

Note that by (2.4) and (2.5) the mapping T of Theorem 2.2 was found to be the unique solutions of the initial value problem

$$\dot{T} = \dot{\Pi}T, \qquad T(t_0) = T_0, \tag{2.6}$$

where $t_0 \in \mathcal{J}$ is some fixed value, $\Pi(t)$ denotes again the orthogonal projection onto $\ker M(t)$, $t \in \mathcal{J}$, and $T_0 \in \mathcal{L}(\mathbb{R}^{n-r}, \mathbb{R}^n)$ is any operator such that $T_0(\mathbb{R}^{n-r}) = \ker M(t_0)$. Analogously, the mapping V is the solution of

$$\dot{V} = V\dot{P}, \qquad V(t_0) = LP(t_0), \tag{2.7}$$

where $P(t)$ is the orthogonal projection onto rge $M(t)$, $t \in \mathcal{J}$ and $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^r)$ is any operator such that $L(\text{rge } M(t_0)) = \mathbb{R}^r$.

In the analytic case, the assumption that rank $M(t) = r$, $\forall t \in \mathcal{J}$, can be dropped without major consequences for Theorem 2.2 because of the following variant of Lemma 2.1:

Lemma 2.2. *Let $\mathcal{J} \subset \mathbb{R}$ be an open interval, $n \geq 1$ any integer, and $M \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$. Set $r = \max_{t \in \mathcal{J}}$ rank $M(t)$. Then, there is a subset $\mathcal{S} \subset \mathcal{J}$ of isolated points such that rank $M(t) = r$ if and only if $t \in \mathcal{J} \setminus \mathcal{S}$. Furthermore, the orthogonal projections onto rge $M(t)$ and $\ker M(t)$, $t \in \mathcal{J} \setminus \mathcal{S}$, which are analytic on $t \in \mathcal{J} \setminus \mathcal{S}$, can be extended as analytic functions over the entire interval $\mathcal{J}$.*

Proof. Since rank $M(t)$ achieves only the discrete values $0, 1, \cdots, n$, there exists a $\bar{t} \in \mathcal{J}$ such that rank $M(\bar{t}) = r$. Identifying $M(t)$ with its matrix in the canonical basis of $\mathbb{R}^n$, we see that some $r \times r$ minor of $M(t)$ does not vanish identically in $\mathcal{J}$, and hence, by analyticity, can vanish only on some subset $\mathcal{S}$ of isolated points of $\mathcal{J}$. Since rge $M(t) =$ rge $M(t)M(t)^T$, $\forall t \in \mathcal{J}$, we may replace M by MM^T without modifying rge $M(t)$ or affecting the analyticity. It then follows from the symmetry of MM^T and Theorem 6.1, p.139 in [K82] that the orthogonal projection $P(t)$ onto rge $M(t)$, $t \in \mathcal{J} \setminus \mathcal{S}$, can be extended as an analytic function in $\mathcal{J}$. (In the notation of [K82], P is the sum of the eigen-projections P_h corresponding to the eigenvalues λ_h of MM^T that do not vanish identically in $\mathcal{J}$, because eigenprojections of symmetric operators are orthogonal projections). Let now $P(t)$ denote the orthogonal projection onto $\ker M(t)$, $t \in \mathcal{J} \setminus \mathcal{S}$. For the proof that P can be extended as an analytic function in $\mathcal{J}$ it suffices to replace $M(t)$ by $M(t)^T M(t)$ and to use once again Theorem 6.1, p.139 in [K 82]. $\square$

Let M, r, and $\mathcal{S}$ be as in Lemma 2.2, and call $P(t)$ the orthogonal projection onto rge $M(t)$, $t \in \mathcal{J} \setminus \mathcal{S}$, so that $P \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$. By the denseness of $\mathcal{J} \setminus \mathcal{S}$ in $\mathcal{J}$ and the relation $PM = M$ in $\mathcal{J} \setminus \mathcal{S}$, we find that $PM = M$ in $\mathcal{J}$ and, in particular, rge $M(t) \subset$ rge $P(t)$ for $t \in \mathcal{S}$. Moreover, $P(t)$ remains an orthogonal projection at points of $\mathcal{S}$ as follow from the relations $P^2 = P$ and $P^T = P$ extending from $\mathcal{J} \setminus \mathcal{S}$ to $\mathcal{J}$. In other words, $P(t)$ is a projection onto a subspace containing rge $M(t)$. Let now $P_1 \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$ be any other family of projections such that rge $P_1(t) =$ rge $M(t)$, $\forall t \in \mathcal{J} \setminus \mathcal{S}$. Obviously, in $\mathcal{J} \setminus \mathcal{S}$, and hence in $\mathcal{J}$, we have $PP_1 = P_1$ so that rge $P_1(t) \subset$ rge $P(t)$, $\forall t \in \mathcal{J}$. By an exchange of the roles of P and P_1, it follows that rge $P_1(t) =$ rge $P(t)$, $\forall t \in \mathcal{J}$. This shows that even at points of $\mathcal{S}$, the range of $P(t)$ is independent of the specific choice of P. For this reason, we call rge $P(t)$ the *extended range* of $M(t)$ to be denoted by ext rge $M(t)$. Then, we have

(2.8a) $$\text{rge } M(t) \subset \text{ ext rge } M(t), \quad \forall t \in \mathcal{J},$$

(2.8b) $$\text{rge } M(t) = \text{ ext rge } M(t), \quad \forall t \in \mathcal{J} \setminus \mathcal{S}.$$

Let $\Pi(t)$ again denote the orthogonal projection onto $\ker M(t)$, $t \in \mathcal{J} \setminus \mathcal{S}$, so that $\Pi \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$. The denseness of $\mathcal{J} \setminus \mathcal{S}$ in $\mathcal{J}$ and the relation $M\Pi = 0$ in $\mathcal{J} \setminus \mathcal{S}$ imply that $M\Pi = 0$ in $\mathcal{J}$. In particular, for $t \in \mathcal{S}$, we have rge $\Pi(t) \subset \ker M(t)$ and hence $\Pi(t)$ projects onto a subspace contained in $\ker M(t)$. Let now $\Pi_1 \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$ be any other family of projections such that rge $\Pi_1(t) = \ker M(t)$, $\forall t \in \mathcal{J} \setminus \mathcal{S}$. Obviously, $\Pi\Pi_1 = \Pi_1$ and $\Pi_1\Pi = \Pi$ in $\mathcal{J} \setminus \mathcal{S}$, and hence in $\mathcal{J}$, so that rge $\Pi_1(t) =$ rge $\Pi(t)$, $\forall t \in \mathcal{J}$. Thus, the range of $\Pi(t)$ remains independent of the specific choice of Π at points of $\mathcal{S}$, and we call rge $\Pi(t)$ the *restricted null-space* of $M(t)$ to be denoted by rest ker $M(t)$. Then

$$\ker M(t) \supset \text{ rest ker } M(t), \qquad \forall t \in \mathcal{J}, \tag{2.9a}$$

$$\ker M(t) = \text{ rest ker } M(t), \qquad \forall t \in \mathcal{J} \setminus \mathcal{S}. \tag{2.9b}$$

By Theorem 2.1, ext rge $M(t)$ and rest ker $M(t)$ are isomorphic to ext rge $M(t_0)$ and rest ker $M(t_0)$, respectively, for any $t, t_0 \in \mathcal{J}$. Hence, by (2.8b) and (2.9b), we have dim ext rge $M(t) = r$, $\forall t \in \mathcal{J}$ and dim rest ker $M(t) = n - r$, $\forall t \in \mathcal{J}$ which shows that the inclusions (2.8a) and (2.9a) are strict for $t \in \mathcal{S}$.

Using the concepts of an extended range and restricted null-space, and by a proof similar to that of Theorem 2.2, we obtain now the following version of that result:

Theorem 2.3. *Let $\mathcal{J} \subset \mathbb{R}$ be an open interval, $n \geq 0$ any integer, and $M \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$. Set $r = \max_{t \in \mathcal{J}}$ rank $M(t)$ and let $\mathcal{S} \subset \mathcal{J}$ be the set of isolated points where rank $M(t) < r$ (see Lemma 2.2). Then the following operators exist:*

(i) $S \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^r, \mathbb{R}^n))$ such that $S(t) \in GL(\mathbb{R}^r, \text{ext rge } M(t))$, $\forall t \in \mathcal{J}$, and, in particular, $S(t) \in GL(\mathbb{R}^r, \text{rge } M(t))$, $\forall t \in \mathcal{J} \setminus \mathcal{S}$.

(ii) $T \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^{n-r}, \mathbb{R}^n))$ such that $T(t) \in GL(\mathbb{R}^{n-r}, \text{rest ker } M(t))$, $\forall t \in \mathcal{J}$, and, in particular, $T(t) \in GL(\mathbb{R}^{n-r}, \ker M(t))$, $\forall t \in \mathcal{J} \setminus \mathcal{S}$.

(iii) $V \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^n, \mathbb{R}^r))$ such that $V(t) \in GL(\text{rge } M(t), \mathbb{R}^r)$, $\forall t \in \mathcal{J}$, and, in particular, $V(t) \in GL(\text{rge } M(t), \mathbb{R}^r)$, $\forall t \in \mathcal{J} \setminus \mathcal{S}$ and V can be chosen so that $\ker V(t) = [\text{ext rge } M(t)]^\perp$, $\forall t \in \mathcal{J}$.

Remark 2.1. For the validity of (2.6) and (2.7), $\Pi(t)$ and $P(t)$ must be chosen as the orthogonal projections onto rest ker $M(t)$ and ext rge $M(t)$, respectively. $\square$

3. The reduction of analytic pairs.

Let $\mathcal{J} \subset \mathbb{R}$ again be an open interval and $A, B \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$. We denote by $A \dotplus B \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}^n))$ the parametrized family of operators

$$(3.1) \qquad (p, x) \in \mathbb{R}^n \times \mathbb{R}^n \longmapsto (A(t) \dotplus B(t))(p, x) \equiv A(t)p + B(t)x, \quad \forall t \in \mathcal{J}.$$

Definition 3.1. *The pair (A, B) is called regular in $\mathcal{J}$ if the rank condition*

$$(3.2) \qquad \operatorname{rank} A(t) \dotplus B(t) = n, \quad \forall t \in \mathcal{J}.$$

holds.

For later use, we note that (3.2) is equivalent with either one of the conditions

$$(3.2') \qquad \operatorname{rank} [A(t)A(t)^T + B(t)B(t)^T] = n, \quad \forall t \in \mathcal{J},$$

$$(3.2'') \qquad \ker A(t)^T \cap \ker B(t)^T = \{0\}, \quad \forall t \in \mathcal{J}.$$

$$(3.2''') \qquad \operatorname{rge} A(t) + \operatorname{rge} B(t) = \mathbb{R}^n, \quad \forall t \in \mathcal{J}.$$

although (3.2''') will not be particularly useful here.

Suppose that (A, B) is regular in the sense of Definition 3.1. By Lemma 2.2, it follows that

$$(3.3) \qquad \operatorname{rank} A(t) = r, \quad \forall t \in \mathcal{J} \setminus \mathcal{S},$$

where $\mathcal{S} \subset \mathcal{J}$ consists of isolated points, and that there is a family of projections $P \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$ with $\operatorname{rge} P(t) = \operatorname{ext} \operatorname{rge} A(t)$, $\forall t \in \mathcal{J}$. Set $Q = I - P$, so that $Q(t)$ projects onto a complement of $\operatorname{rge} A(t)$, $\forall t \in \mathcal{J} \setminus \mathcal{S}$. Clearly, $\operatorname{rge} A(t) \oplus B(t) = \operatorname{rge} A(t) \oplus Q(t)B(t) = \operatorname{rge} A(t) \dotplus \operatorname{rge} Q(t)B(t)$. Hence, it follows from (3.2) and (3.3) that $\operatorname{rank} Q(t)B(t) = n - r$, $\forall t \in \mathcal{J} \setminus \mathcal{S}$, or, equivalently, that

$$(3.4) \qquad \dim \ker Q(t)B(t) = r, \quad \forall t \in \mathcal{J} \setminus \mathcal{S}.$$

Remark 3.1. Using (3.2) we see easily that (3.4) actually holds for $t \in \mathcal{J}$. Indeed, given $w \in \mathbb{R}^n$, there are $u, v \in \mathbb{R}^n$ such that $A(t)u + B(t)v = w$. Since $QA = 0$ in $\mathcal{J}$, it follows that $Q(t)B(t)v = Q(t)w$. Letting w run over the $(n-r)$-dimensional subspace rge $Q(t)$ we infer that rank $Q(t)B(t) \geq n - r$, $\forall t \in \mathcal{J}$. Thus, equality holds since, by (3.4) and the lower semi-continuity of the rank, rank $Q(t)B(t) \leq n - r$, $\forall t \in \mathcal{J}$. As a result, "rest ker $Q(t)B(t)$" could be replaced by "ker $Q(t)B(t)$" in the subsequent considerations. However, we have chosen to ignore this fact, as this will later enable us to ascertain that the results of this section can be generalized *with the same proofs*, to a case when (3.2) fails to hold. □

Now, let $C \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^r; \mathbb{R}^n))$ be such that

$$C(t) \in GL(\mathbb{R}^r, \text{rest ker } Q(t)B(t)), \quad \forall t \in \mathcal{J}. \tag{3.5}$$

The existence of C is ensured by (3.4) and Theorem 2.3 (ii) with $M = QB$. Analogously, let $D \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^n, \mathbb{R}^r))$ be such that

$$D(t) \in GL(\text{ext rge } A(t), \mathbb{R}^r), \quad \forall t \in \mathcal{J}, \tag{3.6}$$

where now the existence of D is ensured by (3.3) and Theorem 2.3 (iii) with $M = A$.

Remark 3.2. For $t \in \mathcal{J} \setminus \mathcal{S}$ we have $\ker Q(t)B(t) = \{x \in \mathbb{R}^n : B(t)x \in \text{rge } A(t)\}$, which shows that $\ker Q(t)B(t)$ is independent of the particular choice of Q; that is, of P. Consequently, $C(t)$ is independent of the specific choice of P for $t \in \mathcal{J} \setminus \mathcal{S}$ and hence for $t \in \mathcal{J}$. □

Definition 3.2. *For the regular pair (A, B) in $\mathcal{J}$ let $C \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^r, \mathbb{R}^n))$ and $D \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^n, \mathbb{R}^r))$ satisfy (3.5) and (3.6), respectively. Then the pair (A_1, B_1) defined by*

$$A_1, B_1 \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^r)), \quad A_1 = DAC, \quad B_1 = D(BC + A\dot{C}), \tag{3.7}$$

is a reduction of (A, B) in $\mathcal{J}$.

Clearly, a reduction (A_1, B_1) of a regular pair (A, B) is not unique, since it depends upon the choice of C and D. However, any two reductions of (A, B) are equivalent under the following equivalence relation already used in [C87] and rediscovered in [KuMe92].

Definition 3.3. *For any $A, B, \tilde{A}, \tilde{B} \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$ the pairs (A, B) and $(\tilde{A}, \tilde{B})$ are called equivalent in $\mathcal{J}$, to be denoted by $(A, B) \sim (\tilde{A}, \tilde{B})$ in $\mathcal{J}$, if*

$$\tilde{A} = MAN \quad \text{and} \quad \tilde{B} = M(BN + A\dot{N}), \tag{3.8}$$

for some $M, N \in C^\omega(\mathcal{J}; GL(\mathbb{R}^n))$.

Lemma 3.1. *The relation of Definition 3.3 is an equivalence relation.*

Proof. In this proof, the interval $\mathcal{J}$ will be implicit; that is, $(A, B) \sim (\tilde{A}, \tilde{B})$ shall mean that this relation holds in $\mathcal{J}$. Clearly, $(A, B) \sim (\tilde{A}, \tilde{B})$ is trivial. If $(A, B) \sim (\tilde{A}, \tilde{B})$ holds then $A = M^{-1}\tilde{A}N^{-1}$ and we obtain $B = M^{-1}(\tilde{B}N^{-1} + \tilde{A}\frac{d}{dt}N^{-1})$. Indeed, we have $M^{-1}(\tilde{B}N^{-1} + \tilde{A}\frac{d}{dt}N^{-1}) = (BN + A\dot{N})N^{-1} + AN\frac{d}{dt}N^{-1} = B + A(\dot{N}N^{-1} + N\frac{d}{dt}N^{-1}) = B$, since $\dot{N}N^{-1} + N\frac{d}{dt}N^{-1} = \dot{I} = 0$. This shows that $(\tilde{A}, \tilde{B}) \sim (A, B)$.

Now, suppose that $(\hat{A}, \hat{B}) \sim (A, B)$ and $(A, B) \sim (\tilde{A}, \tilde{B})$; that is, $A = \hat{M}\hat{A}\hat{N}$, $B = \hat{M}(\hat{B}\hat{N} + \hat{A}\dot{\hat{N}})$ and $\tilde{A} = MAN$, $\tilde{B} = M(BN + A\dot{N})$ for $\hat{M}, \hat{N}, M, N \in GL(\mathcal{J}; GL(\mathbb{R}^n))$. Then, by a straightforward calculation, we find $\tilde{A} = M\hat{M}\hat{A}\hat{N}N$, $\tilde{B} = M\hat{N}(\hat{B}\hat{N}N + \hat{A}\frac{d}{dt}(\hat{N}N))$ and therefore $(\hat{A}, \hat{B}) \sim (\tilde{A}, \tilde{B})$. $\square$

In the proof of part (iii) of our next theorem and in other places later on we require the following lemma.

Lemma 3.2. *Let $P \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$ be a family of projections such that rge $P(t) =$ ext rge$A(t)$, $\forall t \in \mathcal{J}$ and set $Q = I - P$. Then for any integer $m \geq 0$ and $K \in C^0(\mathcal{J}; \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n))$ we have*

(i) rge $K(t) \subset$ rest ker $Q(t)B(t)$, $\forall t \in \mathcal{J}$,

if and only if

(ii) rge $B(t)K(t) \subset$ ext rge$A(t)$, $\forall t \in \mathcal{J}$,

Note: The equivalence between (i) and (ii) above is *not* true for *fixed* $t \in \mathcal{J}$, although (i) implies (ii) in this case as well.

Proof. Suppose that (i) holds, so that rge $K(t) \subset \ker Q(t)B(t)$, $\forall t \in \mathcal{J}$. This is equivalent with $QBK = 0$; that is, rge $B(t)K(t) \subset \ker Q(t) =$ rge $P(t) =$ ext rge$A(t)$, $\forall t \in \mathcal{J}$, which is (ii). Conversely, assume (ii) to hold, so that $QBK = 0$ and thus rge $K(t) \subset \ker Q(t)B(t)$,

$\forall t \in \mathcal{J}$. Let $\Pi \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$ be the orthogonal projection onto rest ker $Q(t)B(t)$. Since rest ker $Q(t)B(t) = \ker Q(t)B(t)$ except at isolated points of $\mathcal{J}$, we find that $(I - \Pi(t))K(t) = 0$ except at isolated points. But by denseness and continuity, it follows that necessarily $(I - \Pi)K = 0$; that is, $K = \Pi K$, whence rge $K(t) \subset$ rge $\Pi(t) =$ rest ker $Q(t)B(t)$, $\forall t \in \mathcal{J}$ which is (i). □

Theorem 3.1. *Suppose that* $(A, B) \sim (\tilde{A}, \tilde{B})$ *in* $\mathcal{J}$. *Then.*

(i) rank $A(t) =$ *rank* $\tilde{A}(t)$, $\forall t \in \mathcal{J}$.

(ii) (A, B) *is regular in* $\mathcal{J}$ *if and only if* $(\tilde{A}, \tilde{B})$ *is regular in* $\mathcal{J}$.

(iii) If (A, B) *and* $(\tilde{A}, \tilde{B})$ *are regular in* $\mathcal{J}$, *any reduction* (A_1, B_1) *of* (A, B) *in* $\mathcal{J}$ *is equivalent to any reduction* $(\tilde{A}_1, \tilde{B}_1)$ *of* $(\tilde{A}, \tilde{B})$ *in* $\mathcal{J}$.

Proof. By (3.8), (i) is trivial. To show that (3.2), or equivalently (3.2″), either fails or holds simultaneously with A and $\tilde{A}$, it suffices to show that if (3.2″) is valid for the pair (A, B), then it also holds for the pair $(\tilde{A}, \tilde{B})$. Indeed, symmetry of the equivalence relation implies that the converse is true as well. Suppose then that (3.2″) is valid and let $t \in \mathcal{J}$ and $x \in \ker \tilde{A}(t)^T \cap \ker \tilde{B}(t)^T$. By (3.8), this means that $N(t)^T A(t)^T M(t)^T x = 0$ and $N(t)^T B(t)^T M(t)^T x + \dot{N}(t)^T A(t)^T M(t)^T x = 0$. Because of $N^T(t) \in GL(\mathbb{R}^n)$, the first relation provides that $M(t)^T x \in \ker A(t)^T$, and hence the second one reduces to $N(t)^T B(t)^T M(t)^T x = 0$; that is, $M(t)^T x \in \ker B(t)^T$. Thus, $M(t)^T x \in \ker A(t)^T \cap \ker B(t)^T$, whence $M(t)^T x = 0$ by (3.2″). Because of $M(t)^T \in GL(\mathbb{R}^n)$ we find that $x = 0$ which proves (ii).

To prove (iii), we write

$$A_1 = DAC,\ B_1 = D(BC + A\dot{C}),\ \tilde{A}_1 = \tilde{D}\tilde{A}\tilde{C},\ \tilde{B}_1 = \tilde{D}(\tilde{B}\tilde{C} + \tilde{A}\dot{\tilde{C}}) \tag{3.9}$$

where $C, \tilde{C} \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^r, \mathbb{R}^n))$ and $D, \tilde{D} \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^n, \mathbb{R}^r))$ satisfy the conditions of Definition 3.2. Note here that the integer r is the same for both pairs (A, B) and $(\tilde{A}, \tilde{B})$ by part (i) of this theorem. From (3.8), we infer that

$$\tilde{A}_1 = \tilde{D}MAN\tilde{C}, \quad \tilde{B}_1 = \tilde{D}M(BN\tilde{C} + A\frac{d}{dt}(N\tilde{C})). \tag{3.10}$$

Now the proof hinges on the remark that there are mappings $H, L \in C^\omega(\mathcal{J}; GL(\mathbb{R}^r))$ such that

$$NC̃ = CL \tag{3.11}$$

$$\tilde{D}(t)M_{|\text{ext rge } A(t)} = H(t)D_{|\text{ext rge } A(t)}, \quad \forall t \in \mathcal{J}, \tag{3.12}$$

where the existence of H and L will be established further below. For the time being, we show how the validity of (3.10), (3.11) provides for the proof of (iii). From (3.10), (3.11) and (3.12), we obtain at once

$$\tilde{A}_1 = HDACL = HA_1L, \tag{3.13}$$

and $\tilde{B}_1 = \tilde{D}MBCL + HDA\frac{d}{dt}(CL)$. Recall that $C(t)$ maps into rest ker $Q(t)B(t)$ for $t \in \mathcal{J}$, where $Q(t)$ is a projection onto a complement of ext rge$A(t)$. Therefore, by Lemma 3.2 with $K = C$, $B(t)C(t)$ maps into ext rge$A(t)$, $t \in \mathcal{J}$, and hence (3.12) implies that $\tilde{D}MBC = HDBC$. As a result, $\tilde{B}_1 = HDBCL + HDA\frac{d}{dt}(CL) = HD(BC + A\dot{C})L + HDAC\dot{L}$. Using (3.9) we thus get $\tilde{B}_1 = H(B_1L + A_1\dot{L})$ and, together with (3.13), this implies the equivalence of (A_1, B_1) and $(\tilde{A}_1, \tilde{B}_1)$ since $H, L \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^r))$.

To complete the proof, we now verify the validity of (3.11) and (3.12). By hypothesis, $\tilde{B} = M(BN + A\dot{N})$ and, by construction of $\tilde{C}$, $\tilde{B}(t)\tilde{C}(t)$ maps into ext rge $\tilde{A}(t)$ for $t \in \mathcal{J}$. Since ext rge $\tilde{A}(t) =$ ext rge $M(t)A(t)N(t) =$ ext rge $M(t)A(t) = M(t)$(ext rge $A(t)$), this amounts to saying that $M(t)(B(t)N(t)+A(t)\dot{N}(t))\tilde{C}(t)$ maps into $M(t)$(ext rge $A(t)$); that is, $B(t)N(t)\tilde{C}(t)+A(t)\dot{N}(t)\tilde{C}(t)$ maps into ext rge $A(t)$. But then, $B(t)N(t)\tilde{C}(t)$ maps into ext rge $A(t)$. By Lemma 3.2, $N(t)\tilde{C}(t)$ maps into rest ker $Q(t)B(t) =$ rge $C(t)$, $\forall t \in \mathcal{J}$ with $Q \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$ as in (3.4) and (3.5). By rank $\tilde{C}(t) = r$ and $N(t) \in GL(\mathbb{R}^n)$ we have rge $N(t)\tilde{C}(t) =$ rge $C(t)$. On the other hand, injectivity of $C(t)$ implies that $C(t)^TC(t) \in GL(\mathbb{R}^r)$, $\forall t \in \mathcal{J}$. Set $L(t) = (C(t)C(t)^T)^{-1}C(t)^TN(t)\tilde{C}(t)$, for $t \in \mathcal{J}$, so that $L \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^r))$. In fact, $L \in C^\omega(\mathcal{J}; GL(\mathbb{R}^r))$, for as was just seen above, $N(t)\tilde{C}(t)$ maps onto rge $C(t)$ and $C(t)^T$ is one-to-one in rge $C(t)$. Thus, rank $L(t) = r$ as desired. Next, for $\xi \in \mathbb{R}^r$ we have $L(t)\xi = (C(t)^TC(t))^{-1}(C(t)^TN(t)\tilde{C}(t)\xi$, whence

$C(t)^T C(t)L(t)\xi = C(t)^T N(t)\tilde{C}(t)\xi$. Since $C(t)^T$ is one-to-one in rge $C(t) =$ rge $N(t)\tilde{C}(t)$, this gives $C(t)L(t)\xi = N(t)\tilde{C}(t)\xi$; that is, $CL = N\tilde{C}$ which completes the proof of (3.11).

To prove (3.12) let $S \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^r, \mathbb{R}^n))$ be such that $S(t) \in GL(\mathbb{R}^r, \text{ext rge } A(t))$, $\forall t \in \mathcal{J}$, as provided by Theorem 2.3 (i). Note that $D(t)S(t) \in GL(\mathbb{R}^r)$, $\forall t \in \mathcal{J}$, and set $H(t) = \tilde{D}(t)M(t)S(t)(D(t)S(t))^{-1}$, so that $H \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^r))$. In fact, we obtain $H \in C^\omega(\mathcal{J}; GL(\mathbb{R}^r))$ because $M(t) \in GL(\mathbb{R}^n)$ implies that $M(t)S(t)$ is a bijective mapping into $M(t)[\text{ext rge } A(t)] = \text{ext rge } \tilde{A}(t)$, whereas $\tilde{D}(t)$ is one-to-one in ext rge $\tilde{A}(t)$. Now, for $x \in \text{ext rge } A(t)$, we have $x = S(t)\xi$ for some $\xi \in \mathbb{R}^r$, whence $D(t)x = D(t)S(t)\xi$ and $H(t)D(t)x = H(t)D(t)S(t)\xi$. Thus, $H(t)D(t)x = \tilde{D}(t)M(t)S(t)\xi = \tilde{D}(t)M(t)x$, which proves (3.14). □

Theorem 3.1 makes it easy to define inductively a reduction procedure which, up to equivalence, is independent of any particular choice that must be made at each step. Specifically, suppose that the pair (A, B) is regular in $\mathcal{J}$, and let (A_1, B_1) and $(\tilde{A}_1, \tilde{B}_1)$ be any two reductions of (A, B) in $\mathcal{J}$. By Theorem 3.1, $(A_1, B_1) \sim (\tilde{A}_1, \tilde{B}_1)$ and hence (A_1, B_1) and $(\tilde{A}_1, \tilde{B}_1)$ are simultaneously regular or not regular in $\mathcal{J}$. If they are, then once again by Theorem 3.1, any reduction (A_2, B_2) of (A_1, B_1) is equivalent with any reduction $(\tilde{A}_2, \tilde{B}_2)$ of $(\tilde{A}_1, \tilde{B}_1)$. Therefore, (A_2, B_2) and $(\tilde{A}_2, \tilde{B}_2)$ will be simultaneously regular, and so on. This suggests the following definition:

Definition 3.4. *The pair (A, B) is said to be completely regular in $\mathcal{J}$ if the above inductive reduction procedure can be continued indefinitely. In particular, the complete regularity of a pair (A, B) in $\mathcal{J}$ implies its regularity in $\mathcal{J}$.*

A trivial but important case is when (A, B) is regular and $\max_{t\in\mathcal{J}}$ rank $A(t) = n$ and thus $r = n$. Then we have ext rge $A(t) = \mathbb{R}^n$, $\forall t \in \mathcal{J}$, whence $P(t) = I$ is the only possible choice for a parametrized family of projections onto ext rge $A(t)$. As a result, $Q = I - P = 0$ and therefore $QB = 0$ and rest ker $Q(t)B(t) = \mathbb{R}^n$, $\forall t \in \mathcal{J}$. This shows that $C(t) = D(t) = I$ is a possible choice for the reduction of (A, B) for which then, by (3.7), $(A_1, B_1) = (A, B)$. Hence it follows that (A, B) is completely regular and that all the possible reductions are equivalent to (A, B). In other words, up to equivalence, reduction has no effect on the pair (A, B). The importance of this case is that it will always occur eventually at some stage of the reduction procedure of any completely regular pair. Therefore, despite the fact that the reduction can be theortically continued indefinitely, it ceases to have any effect (up to equivalence) after a finite number of steps. This is the content of the following result:

Theorem 3.2. *For a completely regular pair (A, B) in $\mathcal{J}$ set $A_0 = A, B_0 = B$ and consider a sequence (A_j, B_j), $j = 0, 1, \dots$, where $(A_j B_j)$ is some reduction of (A_{j-1}, B_{j-1}) in $\mathcal{J}$ for each $j > 0$. Then,*

$$r_j = \max_{t \in \mathcal{J}} \operatorname{rank} A_j(t), \qquad \forall j > 0, \tag{3.14}$$

is independent of the specific choice of (A_j, B_j), $j > 0$, and with $r_{-1} = n$ we have

$$A_j(t) \in \mathcal{L}(\mathbb{R}^{r_{j-1}}), \quad \forall t \in \mathcal{J},\ \forall j \geq 0, \tag{3.15}$$

and

$$r_{-1} = n \geq r_0 \geq \cdots \geq r_j \geq \cdots \quad (\geq 0). \tag{3.16}$$

Hence there is a smallest integer $0 \leq \nu \leq n$ such that $r_\nu = r_{\nu-1}$ and ν has the property

$$A_\nu(t) \in GL(\mathbb{R}^{r_{\nu-1}})\ (= GL(\mathbb{R}^{r_\nu})), \quad \forall t \in \mathcal{J} \setminus \mathcal{S}_\nu, \tag{3.17}$$

where $\mathcal{S}_\nu \subset \mathcal{J}$ consists only of isolated points, and ν is also characterized by the fact that it is the smallest integer j for which $A_j(t) \in GL(\mathbb{R}^{r_{j-1}})$ for some $t \in \mathcal{J}$.

The proof is a trivial consequence of Theorem 3.1 and the various definitions.

From the remarks preceeding the theorem, it follows by (3.17) that, up to equivalence, the reduction has no effect on the pair (A_ν, B_ν) and hence the process could as well be stopped at this point.

Definition 3.5. *For a completely regular pair (A, B) in $\mathcal{J}$ the integer $0 \leq \nu \leq n$ of Theorem 3.2 is called the index of the pair in $\mathcal{J}$. Since, ν is independent of the choice of the reduction at each step of the process the index depends only on the pair (A, B).*

4. Global reduction of differential-algebraic equations: classical solutions.

We apply now the reduction procedure of the previous section to the differential algebraic equation

$$A(t)\dot{x} + B(t)x = b(t), \quad t \in \mathcal{J}, \qquad A, B \in C^{\omega}(\mathcal{J}; \mathcal{L}(\mathbb{R}^n)) \tag{4.1}$$

and show that it reduces (4.1) to an explicit linear ODE on the interval $\mathcal{J}$. Naturally A and B in (4.1) shall now be called the *coefficients* of the DAE.

Definition 4.1. *The DAE (4.1) with analytic coefficients is reducible (resp. completely reducible) if the pair (A, B) is regular (resp. completely regular) in $\mathcal{J}$. For completely reducible (4.1) the index of the pair (A, B) in $\mathcal{J}$ is the index of (4.1).*

Theorem 4.1. *For the reducible DAE (4.1) with analytic coefficients let (A_1, B_1) be any reduction of the pair (A, B) in $\mathcal{J}$; that is,*

$$A_1 = DAC, \quad B_1 = D(BC + A\dot{C}), \tag{4.2}$$

where $C \in C^{\omega}(\mathcal{J}; \mathcal{L}(\mathbb{R}^r, \mathbb{R}^n))$, $D \in C^{\omega}(\mathcal{J}; \mathcal{L}(\mathbb{R}^n, \mathbb{R}^r))$ satisfy (3.5) and (3.6), respectively, and $r = \max_{t \in \mathcal{J}}$ rank $A(t)$. Suppose that $b \in C^k(\mathcal{J}; \mathbb{R}^n)$ where $1 \leq k \leq \infty$ or $k = \omega$. Then

(i) There exist $u_0 \in C^k(\mathcal{J}; \mathbb{R}^n)$ such that

$$B(t)u_0(t) - b(t) \in \text{ ext rge } A(t), \quad \forall t \in \mathcal{J}, \tag{4.3}$$

and indeed (4.3) holds for

$$u_0 = B^T(AA^T + BB^T)^{-1}b. \tag{4.4}$$

(ii) A differentiable mapping $x : \mathcal{J} \to \mathbb{R}^n$ solves (4.1) if and only if for given u_0 satisfying (4.3) we have

$$x = Cx_1 + u_0 \tag{4.5}$$

where $x_1 : \mathcal{J} \to \mathbb{R}^r$ *is a differentiable solution of the DAE*

$$(4.6) \qquad A_1(t)\dot{x}_1 + B_1(t)x = b_1(t), \quad t \in \mathcal{J},$$

with b_1 *given by*

$$(4.7) \qquad b_1 = D(b - Bu_0 - A\dot{u}_0).$$

Proof. (i) By the reducibility of (A, B) and (3.2′), we have $A(t)A(t)^T + B(t)B(t)^T \in GL(\mathbb{R}^n)$, $\forall t \in \mathcal{J}$. Thus, with u_0 given by (4.4) and $v_0 = A^T(AA^T + BB^T)^{-1}b$, we obtain $Av_0 + Bu_0 = b$ which shows that (4.3) holds.

(ii) Let $x : \mathcal{J} \to \mathbb{R}^n$ be a differentiable solution of (4.1) so that $B(t)x(t) - b(t) \in$ rge $A(t)$, $\forall t \in \mathcal{J}$. For $u_0 \in C^k(\mathcal{J}; \mathbb{R}^n)$ chosen according to part (i) we have $B(t)(x(t) - u_0(t)) \in$ ext rge $A(t)$, $\forall t \in \mathcal{J}$. Then, by applying Lemma 3.2 with $Q(t)$ as stated there and $K \in C^0(\mathcal{J}; \mathcal{L}(\mathbb{R}^1, \mathbb{R}^n))$ defined by $K(t) = x(t) - u_0(t)$, we obtain $x(t) - u_0(t) \in$ rest ker $Q(t)B(t)$. But now (3.5) implies that $x(t) - u_0(t) = C(t)x_1(t)$ for a (unique) $x_1(t) \in \mathbb{R}^r$, and $x_1(t) = [C(t)^T C(t)]^{-1} C(t)^T (x(t) - u_0(t))$ shows that $x_1 : \mathcal{J} \to \mathbb{R}^r$ is differentiable. By differentiating $x = Cx_1 + u_0$ we obtain $AC\dot{x}_1 + (BC + A\dot{C})x_1 = b - A\dot{u}_0 - Bu_0$ and hence, after multiplication by D, that x_1 solves the DAE (4.6).

Conversely, let $x_1 : \mathcal{J} \to \mathbb{R}^r$ be a differentiable solution of (4.6). Then, by (4.2), we have

$$(4.8) \qquad 0 = D[A(C\dot{x}_1 + \dot{C}x_1 + \dot{u}_0) + B(Cx_1 + u_0) - b].$$

By (3.5) and Lemma 3.2 $B(t)C(t)$ maps into ext rge $A(t)$ while, by (4.3), $B(t)u_0(t) - b(t) \in$ ext rge $A(t)$. This shows that the bracketed term in (4.8) belongs to ext rge $A(t)$ for each $t \in \mathcal{J}$. But, by (3.6) $D(t)$ is an isomorphism of ext rge $A(t)$ onto $\mathbb{R}^r$, $\forall t \in \mathcal{J}$, and hence this bracketed term vanishes. Therefore, x defined by $x = Cx_1 + u_0$ is a solution of (4.1). □

Corollary 4.1. *Let the DAE (4.1) with analytic coefficients be completely reducible on* $\mathcal{J}$ *with index* $\nu \geq 0$ *and assume that* $b \in C^k(\mathcal{J};\mathbb{R}^n)$ *with* $\nu \leq k \leq \infty$ *or* $k = \omega$. *Set* $A_0 = A, B_0 = B$ *and consider a sequence* (A_j, B_j), $j = 0, 1, \ldots$, *where* $(A_j B_j)$ *is some reduction of* (A_{j-1}, B_{j-1}) *in* $\mathcal{J}$ *for each* $j > 0$; *that is,*

$$A_j = D_{j-1}A_{j-1}C_{j-1}, \qquad B_j = D_{j-1}(B_{j-1}C_{j-1} + A_{j-1}\dot{C}_{j-1}) \tag{4.9}$$

where C_{j-1} *and* D_{j-1} *satisfy the condition of the reduction procedure and* r_j, $j \geq 0$ *are defined by (3.16) and* $r_{-1} = n$. *Then the following holds:*

(i) There exist sequences $u_j, b_j \in C^{k-j}(\mathcal{J};\mathbb{R}^{r_{j-1}})$, $0 \leq j \leq \nu$, *such that* $b_0 = b$ *and*

$$B_j(t)u_j(t) - b_j(t) \in \text{ ext rge } A_j(t), \quad \forall t \in \mathcal{J},\ j = 1, \ldots, \nu, \tag{4.10a}$$

$$b_j = D_{j-1}(b_{j-1} - B_{j-1}u_{j-1} - A_{j-1}\dot{u}_{j-1}),\ j = 1, \ldots, \nu. \tag{4.10b}$$

(ii) A differentiable mapping $x_j : \mathcal{J} \to \mathbb{R}^{r_{j-1}}$, $j = 0, \ldots, \nu - 1$, *is a solution of the DAE*

$$A_j(t)\dot{x}_j + B_j(t)x_j = b_j(t), \quad t \in \mathcal{J},\ j = 0, \ldots, \nu - 1, \tag{4.11}$$

if and only if x_j *has the form*

$$x_j = C_j x_{j+1} + u_j \tag{4.12}$$

where $x_{j+1} : \mathcal{J} \to \mathbb{R}^{r_j}$ *is a differentiable solution of the DAE*

$$A_{j+1}(t)\dot{x}_{j+1} + B_{j+1}(t)x_{j+1} = b_{j+1}(t), \quad t \in \mathcal{J}. \tag{4.13}$$

(iii) A differentiable mapping $x : \mathcal{J} \to \mathbb{R}^n$ *is a solution of (4.1) if and only if*

$$x = \Gamma_{\nu-1}x_\nu + v_{\nu-1}, \tag{4.14}$$

where $\Gamma_{\nu-1} \in C^{\omega}(\mathcal{J}; \mathcal{L}(\mathbb{R}^{r_{\nu-1}}, \mathbb{R}^n))$ *and* $v_{\nu-1} \in C^{k-\nu+1}(\mathcal{J}; \mathbb{R}^n)$ *are given by*

$$\Gamma_{\nu-1} = C_0 \cdots C_{\ \,-1}, \tag{4.15}$$

$$v_{\nu-1} = C_0 \cdots C_{\nu-2} u_{\nu-1} + C_0 \cdots C_{\nu-3} u_{\nu-2} + \cdots + C_0 u_1 + u_0, \tag{4.16}$$

and $x_\nu : \mathcal{J} \to \mathbb{R}^{r_{\nu-1}}$ *is a differentiable solution of the linear ODE*

$$A_\nu(t)\dot{x}_\nu + B_\nu(t)x_\nu = b_\nu(t). \tag{4.17}$$

Proof. Parts (i) and (ii) follow from by an inductive application of Theorem 4.1. and (iii) is an immediate consequence of (ii). □

The essential difference between (4.1) and (4.17) is that the latter is almost an explicit ODE since $A_\nu(t) \in GL(\mathbb{R}^{r_{\nu-1}})$ except perhaps at isolated points. Of course, the case when $A_\nu(t) \in GL(\mathbb{R}^{r_{\nu-1}})$ for every $t \in \mathcal{J}$ is of special importance. For this case we have the result:

Corollary 4.2. *Under the same hypotheses and with the notation of Corollary 4.1.*

(i) the condition

$$A_\nu(t) \in GL(\mathbb{R}^{r_{\nu-1}}), \quad \forall t \in \mathcal{J}, \tag{4.18}$$

is independent of the reduction, and

(ii) if (4.18) holds. the initial value problem

$$\begin{cases} A(t)\dot{x} + B(t)x = b(t) \text{ in } \mathcal{J}, \\ x(t_*) = x_*, \quad t_* \in \mathcal{J},\ x_* \in \mathbb{R}^n, \end{cases} \tag{4.19}$$

has a differentiable solution if and only if

$$x_* = \Gamma_{\nu-1}(t_*)x_{\nu_*} + v_{\nu-1}(t_*) \tag{4.20}$$

for a (unique) $x_{\nu_*} \in \mathbb{R}^{r_{\nu-1}}$. *Moreover, if (4.20) holds, the differentiable solution* x *of (4.19) is unique and of class* $C^{k-\nu+1}$.

Proof. (i) If $(\tilde{A}_\nu, \tilde{B}_\nu)$ is another possible choice for (A_ν, B_ν), then $(\tilde{A}_\nu, \tilde{B}_\nu) \sim (A_\nu, B_\nu)$ by Theorem 3.1 (iii), and rank $A_\nu(t) =$ rank $\tilde{A}_\nu(t)$, $\forall t \in \mathcal{J}$ by Theorem 3.1 (i).

(ii) If (4.18) holds, then (4.17) is an explicit linear ODE, and hence the conclusion follows from Corollary 4.1 and standard existence and uniqueness results for linear ODEs – which in particular implies that solutions are defined in all of $\mathcal{J}$. Note that uniqueness of x_{ν_*} in (4.20) follows from the injectivity of $\Gamma_{\nu-1}(t_*)$. □

Remark 4.1. Condition (4.20) expresses the *consistency* of the initial value x_* with the DAE at t_*. □

For $b = 0$ in (4.19) we may choose $u_j = b_j = 0$ in Corollary 4.1 which shows that $v_{\nu-1}(t_*) = 0$ in (4.20). Hence it follows from Corollary 4.2 (ii) that on $\mathcal{J}$ the homogeneous system

$$(4.21) \qquad A\dot{x} + Bx = 0, \quad \text{in } \mathcal{J},$$

has $r_{\nu-1}$ linearly independent (and analytic) solutions which are uniquely determined by their value at any point $t_* \in \mathcal{J}$. We now show that the hypotheses of Corollary 4.1 are *necessary* for the existence of a solution for arbitrary right-hand side in (4.1), and that the same is true of the hypotheses of Corollary 4.2 (ii) if the aforementioned property of the solutions of (4.21) is to be preserved as well.

Theorem 4.2. *(i) In order for the DAE (4.1) with analytic coefficients to have at least one differentiable solution in* $\mathcal{J}$ *for every* $b \in C^\omega(\mathcal{J}; \mathbb{R}^n)$, *it is necessary that (4.1) be completely reducible.*

(ii) Let (4.1) be completely reducible with index $\nu \geq 0$. *If the homogeneous problem (4.21) has* $r_{\nu-1}$ *linearly independent differentiable solutions uniquely determined by their value at any given point* $t_* \in \mathcal{J}$, *then* $A_\nu(t) \in GL(\mathbb{R}^{r_{\nu-1}})$, $\forall t \in \mathcal{J}$ *and rank* $A_{\nu-1}(t) = r_{\nu-1}$, $\forall t \in \mathcal{J}$ *if* $\nu \geq 1$.

Proof. (i) It suffices to prove that if (4.1) has a differentiable solution in $\mathcal{J}$ for every

$b \in C^\omega(\mathcal{J}; \mathbb{R}^n)$, then the pair (A, B) is regular in $\mathcal{J}$ and the reduced DAE

$$A_1(t)\dot{x}_1 + B_1(t)x_1 = b_1(t), \quad t \in \mathcal{J}, \tag{4.22}$$

has at least one differentiable solution in $\mathcal{J}$ for every $b_1 \in C^\omega(\mathcal{J}; \mathbb{R}^r)$ where, of course, $r = \max_{t\in\mathcal{J}}$ rank $A(t)$. Indeed, if this property is established, it may evidently be used to prove inductively the complete reducibility of (4.1).

The regularity of (A, B) in $\mathcal{J}$ is trivial since the value $b(t_0)$ can be arbitrarily chosen for any given t_0, and existence of a differentiable solution of (4.1) ensures that $b(t_0) \in$ rge $A(t_0) \oplus B(t_0)$, whence rank $A(t_0) \oplus B(t_0) = n$.

Let (A_1, B_1) be a reduction of (A, B) so that (3.7) holds with $C \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^r, \mathbb{R}^n))$ and $D \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^n, \mathbb{R}^r))$ satisfying (3.5) and (3.6), respectively. For any $b_1 \in C^\omega(\mathcal{J}; \mathbb{R}^r)$ the existence of $S \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^r, \mathbb{R}^n))$ such that $S(t) \in GL(\mathbb{R}^r,$ ext rge $A(t))$, $\forall t \in \mathcal{J}$ is guaranteed by Theorem 2.3 (i). Clearly, $D(t)S(t) \in GL(\mathbb{R}^r)$, $\forall t \in \mathcal{J}$, and with $b(t) = S(t)(D(t)S(t))^{-1}b_1(t)$, $\forall t \in \mathcal{J}$, we obtain $b \in C^\omega(\mathcal{J}; \mathbb{R}^n)$ such that $b(t) \in$ ext rge $A(t)$ and $D(t)b(t) = b_1(t)$, $\forall t \in \mathcal{J}$.

By hypothesis, the DAE (4.1) has a differentiable solution x for this b, and, in addition, we can choose $u_0 = 0$ in Theorem 4.1 (i). Thus, part(ii) of that theorem ensures that x has the form $x = Cx_1$ where x_1 is a differentiable solution of (4.22) in $\mathcal{J}$. Thus, (4.22) does have a differentiable solution, as claimed.

(ii) The $r_{\nu-1}$ linearly independent solutions of (4.21) can be viewed as the columns of an $n \times r_{\nu-1}$ differentiable matrix $\Phi(t)$ satisfying the equation

$$A(t)\dot{\Phi} + B(t)\Phi = 0, \quad t \in \mathcal{J}. \tag{4.23}$$

Since $\Phi(t)$ is uniquely determined (as a solution of (4.23)) by its value $\Phi(t_*)$ at any $t_* \in \mathcal{J}$, it follows that $\Phi(t)$ has rank $r_{\nu-1}$ for all $t \in \mathcal{J}$. Since we may choose $u_j = b_j = 0$ when $b = 0$ in (4.1), Corollary 4.1 provides that $\Phi(t) = \Gamma_{\nu-1}(t)\Phi_\nu(t)$ where Φ_ν is $r_{\nu-1} \times r_{\nu-1}$ and solves the system $A_\nu(t)\dot{\Phi}_\nu + B_\nu(t)\Phi_\nu = 0$ in $\mathcal{J}$. Because of rank $\Phi(t) = r_{\nu-1}$ we also have rank $\Phi_\nu(t) = r_{\nu-1}$ and therefore $\Phi_\nu(t) \in GL(\mathbb{R}^{r_{\nu-1}})$, $\forall t \in \mathcal{J}$. From the proof of part (i), the existence of a differentiable solution x of (4.1) for every $b \in C^\omega(\mathcal{J};\mathbb{R}^n)$ implies the existence of a differentiable solution x_ν of (4.17) for every $b_\nu \in C^\omega(\mathcal{J};\mathbb{R}^{r_{\nu-1}})$. Clearly $y_\nu = \Phi_\nu^{-1}x_\nu$ is differentiable and solves the system $A_\nu(t)\dot{y}_\nu = b_\nu(t)$ in $\mathcal{J}$ which, in turn, inplies that $b_\nu(t) \in \text{rge } A_\nu(t)$, $\forall t \in \mathcal{J}$. Since $b_\nu(t)$ can be arbitrarily chosen for given $t \in \mathcal{J}$ this shows that rank $A_\nu(t) = r_{\nu-1}$ and therefore $A_\nu(t) \in GL(\mathbb{R}^{r_{\nu-1}})$, $\forall t \in \mathcal{J}$.

To complete the proof, we have to show that rank $A_{\nu-1}(t) = r_{\nu-1}$, $\forall t \in \mathcal{J}$, assuming, of course, $\nu \geq 1$. For this, it is obviously not restrictive to confine attention to the case $\nu = 1$ and hence to show that rank $A(t) = r(= r_0)$, $\forall t \in \mathcal{J}$. The necessary conditions already established in this proof enable us to use Corollary 4.2 with $b = 0$ which shows that Φ in (4.3) is not only differentiable but also analytic in $\mathcal{J}$. By Lemma 2.1 and Theorem 2.1, with $M = \Phi$, we obtain $U \in C^\omega(\mathcal{J};O(n))$ such that $U(t) \in O(n)$ is a linear isomorphisms of $(\text{rge } \Phi(t_0))^\perp$ onto $(\text{rge } \Phi(t))^\perp$ for each $t \in \mathcal{J}$. Thus, if $\{e_1, \dots, e_{n-r}\}$ denotes a basis of $(\text{rge } \Phi(t_0))^\perp$, the matrix $\tilde{\Phi}(t)$, obtained by adding the column vectors $U(t)e_i$, $1 \leq i \leq n-r$, to the matrix $\Phi(t)$, has full rank n for every $t \in \mathcal{J}$, whence $\tilde{\Phi} \in C^\omega(\mathcal{J};GL(\mathbb{R}^n))$.

The change of variables $y = \tilde{\Phi}^{-1}x$ transforms (4.1) into the DAE $A(t)\tilde{\Phi}(t)\dot{y}+(A(t)\dot{\tilde{\Phi}}(t)+ B(t)\tilde{\Phi}(t))y = b(t)$ in $\mathcal{J}$, which therefore has at least one differentiable solution for every $b \in C^\omega(\mathcal{J};\mathbb{R}^n)$. By part (i), this implies that the pair $(A\tilde{\Phi}, A\dot{\tilde{\Phi}} + B\tilde{\Phi})$ is regular. On the other hand, by (4.23) the matrix $A\dot{\tilde{\Phi}} + B\tilde{\Phi}$ has r vanishing columns. Thus, rank $A(t)\dot{\tilde{\Phi}}(t)+ B(t)\tilde{\Phi}(t) \leq n - r$, $\forall t \in \mathcal{J}$. But then, by the regularity, we must have rank $A(t)\tilde{\Phi}(t) \geq r$, $\forall t \in \mathcal{J}$, which implies that rank $A(t) \geq r$ and therefore rank $A(t) = r$, $\forall t \in \mathcal{J}$. □

We note here that some of our arguments in the proof of Part (ii) were borrowed from [CP83] and [C87]. Part (ii) of Theorem 4.2 cannot be improved to guarantee rank $A_j(t) = r_j$, $j = 0, \dots, \nu - 2$, for all $t \in \mathcal{J}$ as the following 2-dimensional counterexample

$$\begin{pmatrix} 0 & t \\ 0 & 0 \end{pmatrix} \dot{x} + x = b(t)$$

shows. This DAE has index 2 and the unique solution $(x_1, x_2) = (b_1 - t\dot{b}_2, b_2)$. In this case we have $r_1 = 0$ and (4.17) is the trivial equation $0 = 0$ in $\mathbb{R}^0 = \{0\}$.

In connection with Theorem 4.2, the setting of Corollary 4.1 is optimal for the existence of solutions of (4.1) for arbitrary (sufficiently smooth) right-hand sides, and the setting of

Corollary 4.2 (ii) is optimal if, in addition, a uniqueness theory similar to that for standard ODE's is to be valid. However, Corollary 4.1 has a broader range of applications since it also allows for the study of *singularities* in (4.1), which may arise if the matrix $A_\nu(t)$ in (4.17) becomes singular at some values $t \in \mathcal{J}$. For related results see, for instance, [H64] or [F65].

At this stage, it is important to note that the reduction procedure of Section 3 for analytic pairs (A, B) does *not* require regularity of (A, B) and indeed can also be performed under the weaker assumption that $\max_{t\in\mathcal{J}}$ rank $A(t) \oplus B(t) = n$. In particular, the relation (3.4) remains true after enlarging the set $\mathcal{S}$ so as to incorporate the (isolated) points where rank $A(t) \oplus B(t) < n$. All the results of Section 3 remain valid with the same proofs if complete regularity is now replaced by the more general assumption that $\max_{t\in\mathcal{J}}$ rank $A_j(t) \oplus B_j(t) = r_{j-1}$, $(r_{-1} = n)$. But under these weaker assumptions, the DAE (4.1) need no longer be reducible, for the conclusion of Theorem 4.1 (i) does not remain true for arbitrary right-hand sides b. As a consequence, for the validity of Theorem 4.1 (ii) it can no longer be proved, but must be assumed, that there exists a $u_0 \in C^k(\mathcal{J}; \mathbb{R}^n)$ for which $B(t)u_0(t) - b(t) \in$ ext rge $A(t)$. A similar assumption must be made about the b_j's in Corollary 4.1 to arrive at the reduced problem (4.17). In other words, the reduction of (4.1) remains possible only for restricted choices of right-hand sides, independent of the smoothness. To give an idea of the nature of the required restrictions, we mention without proof that if $b \in C^\omega(\mathcal{J}; \mathbb{R}^n)$, then Theorem 4.1 (i) remains valid if the zeroes of the vector $B(t)^T Adj[A(t)A(t)^T + B(t)B(t)^T]b(t)$ – where "Adj" stands for the "Adjugate"; that is,the transpose of the "matrix of cofactors" – all have order greater than or equal to the order of the zeroes of $\det(A(t)A(t)^T + B(t)B(t)^T)$.

Finally, we note that a generalization of the reduction procedure of Section 3 for analytic pairs (A, B) is always possible without any additional assumptions. However, then the reduced pairs (A_j, B_j) are such that $A_j, B_j \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^{n_j}, \mathbb{R}^{r_{j-1}}))$ with $n_j \geq r_{j-1}$; in other words, A_j and B_j need no longer be square matrices. Once again, the reduction stabilizes after a finite number $\nu \geq 0$ of steps, up to a straightforward generalization of the equivalence of Section 3. But, in the case when $n_j > r_{j-1}$ for at least one index j, it can be shown that local uniqueness for the solutions of the initial value problem (4.19) is *not* true. This is due to the fact that, when possible, the reduction to the form (4.17) leads to a system with more unknowns than equations and a surjective $A_\nu(t)$ except at isolated points. On the other hand, the condition $n_j = r_{j-1}$, $\forall j \geq 0$, amounts exactly to $\max_{t\in\mathcal{J}}\text{rank } A_j(t) \oplus B_j(t) = r_{j-1}$, $\forall j \geq 0$, the case discussed above. For some further comments along this line see also Section 7.

5. Global reduction of differential-algebraic equations: generalized solutions.

Let $\mathcal{D}(\mathcal{J})$ denote the space of infinitely differentiable real valued functions with compact support in $\mathcal{J}$ and $\mathcal{D}'(\mathcal{J})$ its dual, the space of distributions in $\mathcal{J}$. For $x = (x_i) \in (\mathcal{D}'(\mathcal{J}))^n$ and $\varphi = (\varphi_i) \in (\mathcal{D}(\mathcal{J}))^n$, we set

$$\langle x, \varphi\rangle = \sum_{i=1}^{n} \langle x_i, \varphi_i\rangle,$$

where, on the right, $\langle\cdot,\cdot\rangle$ denotes the duality pairing between $\mathcal{D}'(\mathcal{J})$ and $\mathcal{D}(\mathcal{J})$. For $M \in C^\infty(\mathcal{J}; \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))$ and $x \in (\mathcal{D}'(\mathcal{J}))^n$, the product $Mx \in (\mathcal{D}'(\mathcal{J}))^m$ is defined by

$$\langle Mx, \varphi\rangle = \langle x, M^T\varphi\rangle, \quad \forall\varphi \in (\mathcal{D}(\mathcal{J}))^m.$$

We consider now the generalized DAE

$$A(t)\dot{x} + B(t)x = b, \quad b \in (\mathcal{D}'(\mathcal{J}))^n, \tag{5.1}$$

where, as before, the coefficients A and B are analytic in $\mathcal{J}$. A solution of (5.1) is any $x \in (\mathcal{D}'(\mathcal{J}))^n$ such that $\langle A\dot{x} + Bx, \varphi\rangle = \langle b, \varphi\rangle$, for all $\varphi \in (\mathcal{D}(\mathcal{J}))^n$, where, of course, $\dot{x}$ denotes the derivative of x in the sense of distributions.

For the extension of the reduction of Section 4 to (5.1) we need a preliminary lemma.

Lemma 5.1. *Suppose that the pair (A, B) is regular and set $r = \max_{t\in\mathcal{J}}\text{rank } A(t)$. Let $P \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$ be such that $P(t)$ is a projection onto ext rge $A(t)$, $\forall t \in \mathcal{J}$, and denote by $\Pi(t)$ the orthogonal projection onto $\ker Q(t)B(t)$, $\forall t \in \mathcal{J}$, where $Q = I - P$. Then:*

(i) For every $\varphi \in (\mathcal{D}(\mathcal{J}))^n$ there exists a $\psi \in (\mathcal{D}(\mathcal{J}))^n$ such that $(I - \Pi)\varphi = B^TQ^T\psi$.

(ii) Let $C \in C^\omega(\mathcal{J};\mathcal{L}(\mathbb{R}^r,\mathbb{R}^n))$ satisfy (3.5). Then, for $x \in (\mathcal{D}'(\mathcal{J}))^n$, there exists a $x_1 \in (\mathcal{D}'(\mathcal{J}))^r$ such that $x = Cx_1$ if and only if $(I-\Pi)x = 0$.

(iii) Let $D \in C^\omega(\mathcal{J};\mathcal{L}(\mathbb{R}^n,\mathbb{R}^r))$ satisfy (3.6). Then, for $\psi \in (\mathcal{D}(\mathcal{J}))^n$, there exists a $\varphi \in (\mathcal{D}(\mathcal{J}))^r$ such that $P^T\psi = P^TD^T\varphi$.

Proof. (i) In line with Remark 3.1 we have $\dim\ker Q(t)B(t) = r$ for $t \in \mathcal{J}$ and hence it follows that rge $B(t)^TQ(t)^T = [\ker Q(t)B(t)]^\perp =$ rge $(I-\Pi(t))$ is $(n-r)$-dimensional, and, because of rank $Q(t)^T =$ rank $Q(t) = n-r$, that $B(t)^T \in GL(\text{rge } Q(t)^T, [\ker Q(t)B(t)]^\perp)$. By Theorem 2.2 (i) and (iii), there are mappings $S \in C^\omega(\mathcal{J};\mathcal{L}(\mathbb{R}^{n-r},\mathbb{R}^n))$ and $V \in C^\omega(\mathcal{J};\mathcal{L}(\mathbb{R}^n,\mathbb{R}^{n-r}))$ such that

$$S(t) \in GL(\mathbb{R}^{n-r}, \text{rge } Q(t)^T), \quad V(t)|_{\text{rge } B(t)^TQ(t)^T} \in GL(\text{rge } B(t)^TQ(t)^T, \mathbb{R}^{n-r})$$

and $\ker V(t) = \ker Q(t)B(t)$, $\forall t \in \mathcal{J}$. This implies that $VB^TS \in C^\omega(\mathcal{J};GL(\mathbb{R}^{n-r}))$. Thus, for $\varphi \in (\mathcal{D}(\mathcal{J}))^n$, the choice $\psi = S(VB^TS)^{-1}V\varphi$, provides that $Q^T\psi = \psi$ and $VB^T\psi = V\varphi$; that is, $(I-\Pi)(B^TQ^T\psi - \varphi) = 0$. But since $(I-\Pi)B^TQ^T = B^TQ^T$, this means that $B^TQ^T\psi = (I-\Pi)\varphi$.

(ii) Suppose first that $x = Cx_1$. Since, again in view of Remark 3.1, $\ker C(t)^T = [\text{rge } C(t)]^\perp = [\ker Q(t)B(t)]^\perp$ we have $C^T(I-\Pi) = 0$ and therefore, $\langle x,(I-\Pi)\varphi\rangle = \langle x_1, C^T(I-\Pi)\varphi\rangle = 0$, $\forall \varphi \in (\mathcal{D}(\mathcal{J}))^n$; that is, $(I-\Pi)x = 0$.

Conversely, suppose that $(I-\Pi)x = 0$ and set $x_1 = (C^TC)^{-1}C^Tx$. Then, for $\varphi \in (\mathcal{D}(\mathcal{J}))^n$, we have $\langle Cx_1,\varphi\rangle = \langle x, C(C^TC)^{-1}C^T\varphi\rangle$. But, $C(t)(C(t)^TC(t))^{-1}C(t)^T$ is the orthogonal projection $\Pi(t)$ onto rge $C(t) = \ker Q(t)B(t)$ (see once again Remark 3.1). Thus, because $\langle x,(I-\Pi)\varphi\rangle = 0$ by hypothesis, we find that $\langle Cx_1,\varphi\rangle = \langle x,\Pi\varphi\rangle = \langle x,\varphi\rangle$ which gives $x = Cx_1$.

(iii) If we can show that $P(t)^T \in GL(\text{rge } D(t)^T, \text{rge } P(t)^T)$, $\forall t \in \mathcal{J}$, the method used in the proof of (i) above will provide the desired result. Since rank $D(t)^T =$ rank $D(t) = r$, and rank $P(t)^T =$ rank $P(t) = r$, $\forall t \in \mathcal{J}$, it suffices to show that $P(t)^Tu = 0$ for some $u \in$ rge $D(t)^T$ implies that $u = 0$. Such a u is an element of $[\ker D(t)]^\perp \cap [\text{ext rge } A(t)]^\perp = [\ker D(t) + \text{ext rge } A(t)]^\perp$. But, since $\ker D(t) \cap \text{ext rge } A(t) = \{0\}$, we have $\ker D(t) \oplus \text{ext rge } A(t) = \mathbb{R}^n$, so that, indeed, u must be 0. □

We are now in a position to prove the "distribution" version of Theorem 4.1.

Theorem 5.1. *Suppose that the DAE (5.1) with analytic coefficients and* $r = \max_{t\in\mathcal{J}} A(t)$ *is reducible, and let* (A_1, B_1) *be any reduction (3.7) of the pair* (A, B) *in* $\mathcal{J}$ *with mappings* $C \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^r, \mathbb{R}^n))$, *and* $D \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^n, \mathbb{R}^r))$ *satisfying (3.5) and (3.6), respectively. Then:*

(i) There exists a $u_0 \in (\mathcal{D}'(\mathcal{J}))^n$ *such that*

$$Q(Bu_0 - b) = 0, \tag{5.2}$$

where $Q = I - P$ *and* $P \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$ *is such that* $P(t)$ *is a projection onto ext rge* $A(t)$, $\forall t \in \mathcal{J}$. *In particular, (5.2) holds for*

$$u_0 = B^T(AA^T + BB^T)^{-1}b. \tag{5.3}$$

(ii) A distribution $x \in (\mathcal{D}'(\mathcal{J}))^n$ *solves (5.1) if and only if for any* u_0 *satisfying (5.2)*

$$x = Cx_1 + u_0, \tag{5.4}$$

where $x_1 \in (\mathcal{D}'(\mathcal{J}))^r$ *solves the DAE*

$$A_1\dot{x}_1 + B_1x_1 = b_1, \tag{5.5}$$

where b_1 *is given by*

$$b_1 = D(b - Bu_0 - A\dot{u}_0). \tag{5.6}$$

Proof. (i) Let u_0 be defined by (5.4) and set $v_0 = A^T(AA^T + BB^T)^{-1}b$. Then, $Av_0 + Bu_0 = b$ and for $\varphi \in (\mathcal{D}(\mathcal{J}))^n$ we have $\langle Q(Bu_0 - b), \varphi\rangle = \langle Bu_0 - b, Q^T\varphi\rangle = -\langle Av_0, Q^T\varphi\rangle = -\langle v_0, A^TQ^T\varphi\rangle = 0$ since $A^TQ^T = (QA)^T = 0$.

(ii) Let $x \in (\mathcal{D}(\mathcal{J}))^n$ be a solution of (5.1), so that for any $\varphi \in (\mathcal{D}(\mathcal{J}))^n$ we have $\langle Q(Bx - b), \varphi\rangle = \langle Q(A\dot{x} + Bx - b), \varphi\rangle = \langle A\dot{x} + Bx - b, Q\varphi\rangle = \langle 0, Q\varphi\rangle = 0$ and therefore

$Q(Bx - b) = 0$ which, because of $Q(Bu_0 - b) = 0$, implies that $QB(x - u_0) = 0$. Thus $\langle x - u_0, B^T Q^T \psi\rangle = 0$, for all $\psi \in (\mathcal{D}(\mathcal{J}))^n$ and, hence, by Lemma 5.1 (i), $\langle x - u_0, (I - \Pi)\varphi\rangle = 0$, $\forall \varphi \in (\mathcal{D}(\mathcal{J}))^n$; that is, $(I - \Pi)(x - u_0) = 0$. Thus, Lemma 5.1 (ii) ensures that $x = Cx_1 + u_0$ for some $x_1 \in (\mathcal{D}'(\mathcal{J}))^r$. Substituting this into (5.1), we find that $AC\dot{x}_1 + (BC + A\dot{C})x_1 = b - Bu_0 - A\dot{u}_0$, and multiplying both sides by D we obtain that x_1 solves (5.5).

Conversely, let $x_1 \in (\mathcal{D}'(\mathcal{J}))^r$ solve the DAE (5.5) and set $x = Cx_1 + u_0$, whence $D(A\dot{x} + Bx - b) = 0$. By construction of x, we have $Q(Bx - b) = Q(BCx_1 + Bu_0 - b) = 0$ since $QBC = 0$ (see Remark 3.1) and $Q(Bu_0 - b) = 0$. Since $QA = 0$, this implies that

$$Q(A\dot{x} + Bx - b) = 0, \tag{5.7}$$

and therefore $DP(A\dot{x} + Bx - b) = D(A\dot{x} + Bx - b) = 0$ which is equivalent with

$$\langle A\dot{x} + Bx - b,\ P^T D^T \varphi\rangle = 0, \quad \forall \varphi \in (\mathcal{D}(\mathcal{J}))^r. \tag{5.8}$$

Let $\psi \in (\mathcal{D}(\mathcal{J}))^n$, so that $\psi = P^T\psi + Q^T\psi$. By Lemma 5.1 (iii), there exists a $\varphi \in (\mathcal{D}(\mathcal{J}))^r$ such that $P^T\psi = P^T D^T \varphi$. Thus, $\psi = P^T D^T \varphi + Q^T \psi$, and $\langle A\dot{x} + Bx - b, \psi\rangle = \langle A\dot{x} + Bx - b, P^T D^T \varphi\rangle + \langle A\dot{x} + Bx - b, Q^T \psi\rangle = 0$ by (5.7) and (5.8). This shows that x solves (5.1). $\square$

By recursive application of Theorem 5.1, we see that when the DAE (5.1) is completely reducible, its solutions are of the form $x = \Gamma_{\nu-1} x_\nu + v_{\nu-1}$ where $x_\nu \in (\mathcal{D}'(\mathcal{J}))^{r_{\nu-1}}$ solves

$$A_\nu \dot{x}_\nu + B_\nu x_\nu = b_\nu, \tag{5.9}$$

and $A_\nu(t)$ is invertible for all but isolated values of $t \in \mathcal{J}$. Here, A_ν, B_ν, $\Gamma_{\nu-1}$ and $v_{\nu-1}$ are obtained as in Corollary 4.1 with the obvious modifications that now $u_j, b_j \in (\mathcal{D}'(\mathcal{J}))^{r_{j-1}}$ (hence $v_{\nu-1} \in (\mathcal{D}'(\mathcal{J}))^n$) and that, instead of (4.10a), u_j is required to satisfy the condition $Q_j(B_j u_j - b_j) = 0$, where $Q_j = I - P_j$ and $P_j \in C^\omega(\mathcal{J}; \mathcal{L}(\mathbb{R}^{r_{j-1}}))$ is such that $P_j(t)$ is a projection onto ext rge $A_j(t)$, $\forall t \in \mathcal{J}$.

If $A_\nu(t)$ is invertible for every $t \in \mathcal{J}$, then (5.9) is simply the explicit ODE with analytic coefficients

$$\dot{x}_\nu + A_\nu^{-1} B_\nu x_\nu = A_\nu^{-1} b_\nu,$$

whose solutions form an $r_{\nu-1}$-dimensional affine subspace of $(\mathcal{D}'(\mathcal{J}))^{r_{\nu-1}}$ (see e.g. Schwartz [S66]). The solutions of (5.1) thus form an $r_{\nu-1}$-dimensional affine subspace of $(\mathcal{D}'(\mathcal{J}))^n$. Of course, in this framework, initial value problems make no sense, in general, since distributions need not have pointwise values. However, Corollary 4.2 can be extended for special choices of b, as, for example, $b \in W^{k,1}_{\text{loc}}(\mathcal{J})$ for some $k \geq \nu$. In this case we note, without proof, that Corollary 4.2 has an obvious generalization due to the fact that the solutions of (4.1) are in $W^{k-\nu+1,1}_{\text{loc}}(\mathcal{J}) \subset C^0(\mathcal{J})$ and hence have numerical values at all the points of $\mathcal{J}$.

Initial value problems also make sense when the right-hand side of (5.1) is a discontinuous function, provided that the initial value is not chosen at a point of discontinuity. It turns out that this result yields rather complete answers to the problem of solving (4.1) with "inconsistent" initial condition; that is, with an initial condition x_* that does *not* satisfy the requirement (4.20). This problem has been discussed at some length in the constant coefficient case because of its relevance in problems from various applications. Details will be presented elsewhere.

Remark 5.1 At the end of Section 4, we observed that the reduction procedure for the DAE (4.1) could be extended to any system with analytic coefficients, provided that suitable limitations are placed on the right-hand sides b_j. It is interesting to note that this is no longer true if the problem is understood in the sense of distributions. Indeed, if the pair (A, B) is not regular, part (i) of Lemma 5.1 fails to hold even in the case when $\max_{t\in\mathcal{J}} \text{rank } A(t) \oplus B(t) = n$ and hence when $\ker Q(t)B(t)$ must be replaced by rest ker $Q(t)B(t)$. Indeed, at a point $t_0 \in \mathcal{J}$ where rest $\ker Q(t_0)B(t_0) \subsetneq \ker Q(t_0)B(t_0)$, we have rge $B(t_0)^T Q(t_0)^T \subsetneq$ rge $(I - \Pi)(t_0)$, and if $\varphi \in (\mathcal{D}(\mathcal{J}))^n$ is chosen such that $(I - \Pi)(t_0)\varphi(t_0) \notin$ rge $B(t_0)^T Q(t_0)^T$, it is clear that no $\psi \in (\mathcal{D}(\mathcal{J}))^n$ exists such that $(I - \Pi)\varphi = B^T Q^T \psi$. But, without Lemma 5.1 (i), it is no longer possible to ascertain that every solution of (5.1) must have the form (5.5), even if existence of u_0, as in (5.2), is assumed. The converse, however, remains true: that is, if x_1 solves (5.6), then x given by (5.5) solves (5.10). This discrepancy may be explained by the fact that the crucial implication (ii) $\Rightarrow$ (i) in Lemma 3.2 relies heavily upon a continuity argument for K (meaningless when $K = x - u_0$ is a distribution) if rest ker $Q(t)B(t)$ and $\ker Q(t)B(t)$ do not coincide for *all* $t \in \mathcal{J}$. $\square$

6. Boundary value problems.

The global reduction result of Corollary 4.1 (iii) is especially convenient for handling boundary value problems. Indeed, suppose that we associate the completely reducible DAE (4.1) with index ν with a condition of the form

$$g(a, b, x(a), x(b), \dot{x}(a), \dot{x}(b)) = 0, \tag{6.1}$$

where $g : \mathcal{J} \times \mathcal{J} \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$, and $a, b \in \mathcal{J}$. By Corollary 4.1 (iii) (and using the notation of that corollary) we know that $x : \mathcal{J} \to \mathbb{R}^n$ solves (4.1) if and only if

$$x = \Gamma_{\nu-1} x_\nu + v_{\nu-1}, \tag{6.2}$$

where $x_\nu : \mathcal{J} \to \mathbb{R}^{r_{\nu-1}}$ solves the ODE

$$\dot{x}_\nu + A_\nu(t)^{-1} B_\nu(t) x_\nu = A_\nu(t)^{-1} b_\nu(t), \quad t \in \mathcal{J}, \tag{6.3}$$

assuming, of course, $A_\nu(t) \in GL(\mathbb{R}^{r_{\nu-1}})$, $\forall t \in \mathcal{J}$. Thus, by substituting (6.2) into (6.1), we see that $x : \mathcal{J} \to \mathbb{R}^n$ solves (4.1) and satisfies the boundary condition (6.1) if and only if x has the form (6.2) with a solution x_ν of the ODE (6.3) that satisfies the boundary condition

$$g_\nu(a, b, x_\nu(a), x_\nu(b), \dot{x}_\nu(a), \dot{x}_\nu(b)) = 0, \tag{6.4}$$

where $g_\nu : \mathcal{J} \times \mathcal{J} \times \mathbb{R}^{r_{\nu-1}} \times \mathbb{R}^{r_{\nu-1}} \times \mathbb{R}^{r_{\nu-1}} \times \mathbb{R}^{r_{\nu-1}} \to \mathbb{R}^m$ is given by

$$\begin{aligned} g_\nu(t, \tau, x_\nu, p_\nu, \xi_\nu, \eta_\nu) = g(t, \tau, &\Gamma_{\nu-1}(t) x_\nu + v_{\nu-1}(t), \Gamma_{\nu-1}(t) p_\nu + \dot{\Gamma}_{\nu-1}(t) x_\nu + \dot{v}_{\nu-1}(t), \\ &\Gamma_{\nu-1}(\tau) \xi_\nu + v_{\nu-1}(\tau), \Gamma_{\nu-1}(\tau) \eta_\nu + \dot{\Gamma}_{\nu-1}(\tau) \xi_\nu + \dot{v}_{\nu-1}(\tau)). \end{aligned} \tag{6.5}$$

But in fact, since by (6.3), $\dot{x}_\nu(a)$ and $\dot{x}_\nu(b)$ are given in terms of a, $x_\nu(a)$ and b, $x_\nu(b)$, respectively, the boundary condition (6.4) takes the form

$$h_\nu(a, b, x_\nu(a), x_\nu(b)) = 0, \tag{6.6}$$

where $h_\nu : \mathcal{J} \times \mathcal{J} \times \mathbb{R}^{r_{\nu-1}} \times \mathbb{R}^{r_{\nu-1}} \to \mathbb{R}^m$ is obtained by substituting

$$p_\nu = A_\nu(t)^{-1}[b_\nu(t) - B_\nu(t)x_\nu], \quad \eta_\nu = A_\nu(\tau)^{-1}[b_\nu(\tau) - B_\nu(\tau)\xi_\nu],$$

into the arguments of g_ν in (6.5).

The problem (6.3) with the boundary condition (6.6) is a boundary value problem for an explicit linear ODE which therefore can be handled by the standard theory. Furthermore, if $g = g(t, \tau, x, p, \xi, \eta)$ in (6.1) is an affine function of (x, p, ξ, η), then $h_\nu = h_\nu(t, \tau, x_\nu, \xi_\nu)$ in (6.6) is an affine function of (x_ν, ξ_ν).

The exact same considerations apply to the distribution case if $b \in W^{\nu,1}_{\text{loc}}(\mathcal{J})$. Indeed, if so, we may guarantee that the sequence (b_j) of Corollary 4.1 satisfies $b_j \in W^{\nu-j,1}_{\text{loc}}(\mathcal{J})$ for $0 \leq j \leq \nu$, whence $x_\nu \in W^{1,1}_{\text{loc}}(\mathcal{J}) \subset C^0(\mathcal{J})$ has a numerical value at each point of $\mathcal{J}$.

7. The non-analytic case.

For $A, B \in C^\infty(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$, the regularity condition (3.2) is no longer sufficient to repeat the arguments of Section 3 and define reductions (A_1, B_1) of (A, B). This is due to the fact that the concepts of extended range and restricted null-space do not necessarily make sense, and that rank $A(t)$ need no longer equal its maximum value in $\mathcal{J}$ at all but isolated points. However, if we strengthen the regularity condition as follows

$$\text{rank } A(t) \oplus B(t) = n, \ \text{rank } A(t) = r, \quad \forall t \in \mathcal{J}, \tag{7.1}$$

then reductions (A_1, B_1) of (A, B) can be obtained by the method of Section 3. To see this, it suffices to replace "ext rge" and "rest ker" by "rge" and "ker", respectively, and to use Theorem 2.2 in lieu of Theorem 2.3. If (7.1) again holds for the pair (A_1, B_1) with n replaced by r, and r replaced by a new integer $r_1 \leq r$, then we obtain a new pair (A_2, B_2) by reduction of (A_1, B_1), and so on. The proof of Theorem 3.1 can be repeated verbatim, provided that "C^ω" is replaced by "C^∞" everywhere, including in the equivalence definition of pairs (Definition 3.1). Likewise, Theorem 3.2 has an obvious extension if complete regularity of the pair (A, B) is replaced by the stronger requirement that for $j \geq 0$ the pair (A_j, B_j) satisfies the appropriate analog of condition (7.1). In particular, an index $\nu \geq 0$ can be assigned to the pair (A, B) in this case.

In line with this, it should be evident how the results of Section 4 up to and including Corollary 4.2 and those of Sections 5 and 6 can be extended when, with $A_0 = A, B_0 = B$, we have

$$\text{rank } A_j(t) \oplus B_j(t) = r_{j-1}, \quad \forall t \in \mathcal{J},\ \forall j \geq 0, \tag{7.2}$$

$$\text{rank } A_j(t) = r_j, \quad \forall t \in \mathcal{J},\ \forall j \geq 0, \tag{7.3}$$

where $r_{-1} = n \geq r_0 \geq \cdots \geq r_j \geq \cdots \geq 0$, and for each $j \geq 1$, (A_j, B_j) denotes a reduction of (A_{j-1}, B_{j-1}). On the other hand, Theorem 4.2 has no interesting generalization to this case, and most of the discussion concluding Section 4 becomes immaterial with one notable exception: The reduction of pairs (A, B) remains possible with A_j, B_j being $n_j \times r_{j-1}$, $n_j \geq r_{j-1}$, if (7.3) is unchanged, and (7.2) is replaced by the weaker condition

$$\text{rank } A_j(t) \oplus B_j(t) = \rho_j, \quad \forall t \in \mathcal{J},\ r_j \leq \rho_j \leq n,\ \forall j \geq 0. \tag{7.4}$$

Once again there is a smallest integer $\nu \geq 0$ such that $r_{\nu-1} = r_\nu$ and $A_\nu(t)$ is onto $\mathbb{R}^{r_{\nu-1}}$ for all $t \in \mathcal{J}$. As in Section 4, if $n_j > r_{j-1}$ for at least one j, local uniqueness of the solutions of (4.19) is lost. Furthermore, because of (7.4) it is easy to characterize those right-hand sides b for which the sequences u_j, b_j are defined, and hence for which the reduction of (4.1) to the form (4.17) is possible. In fact, it is both necessary and sufficient that (with $b_0 = b$)

$$b_j(t) \in \text{rge } A_j(t) \oplus B_j(t) = \text{rge } (A_j(t)A_j(t)^T + B_j(t)B_j(t)^T),\ t \in \mathcal{J},\ 0 \leq j \leq \nu - 1, \tag{7.5}$$

where the surjectivity of $A_\nu(t)$ ensures that this condition holds for $j = \nu$. If so, there exists a w_j as smooth as b_j such that $b_j = (A_j(t)A_j(t)^T + B_j(t)B_j(t)^T)w_j$, and we may choose

$u_j = B_j^T w_j$. Obviously, the condition (7.5) is also necessary for existence of a solution. If b is a distribution, the condition (7.5) is replaced by $b_j = (A_j(t)A_j(t)^T + B_j(t)B_j(t)^T)w_j$ where $w_j \in (\mathcal{D}'(\mathbb{R}))^n$: we shall not enter into the details.

Similar conclusions regarding this "Fredholm Alternative" for undetermined system ($n_j > r_{j-1}$ for at least one j) are obtained in [KuMe92] assuming three, instead of our two, constant rank conditions at each step of their reduction.

Remark 7.1. When A and B are analytic and the DAE (4.1) has index 1, it follows from Theorem 4.2 that *both* conditions in (7.1) as well as rank $A_1(t) = r$, $\forall t \in \mathcal{J}$, are necessary for the validity of an existence and uniqueness theory similar to that for ODE's. Thus, the hypotheses of this section are identical to those made in the analytic case for index 1 problems "without singularities". This is no longer true for higher index problems. □

Extending the results about classical solutions from $A, B \in C^\infty(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$ to $A, B \in C^m(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$, $m \geq n$, requires nothing more than replacing "C^∞" by "C^m" everywhere, and $\nu \leq k \leq \infty$ by $\nu \leq k \leq m$ in Theorem 4.1 as well as Corollaries 4.1 and 4.2. In fact, even C^n regularity is mostly superfluous and only C^ν regularity is needed where ν is the index of (A, B). But since ν (and even the existence of ν) is not known a priori, the hypothesis $A, B \in C^\nu(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$ is ambiguous and should be replaced by "A, B are sufficiently smooth".

A little more care must be exercised regarding generalized solutions. Recall that, given an integer $k \geq 0$, a distribution $x \in (\mathcal{D}'(\mathcal{J}))^n$ is said to be of order k if $\langle x, \cdot \rangle$ extends by continuity to elements of $(\mathcal{D}^k(\mathcal{J}))^n$, where $\mathcal{D}^k(\mathcal{J})$ is the space of real-valued, k-times continuously differentiable functions in $\mathcal{J}$ with compact support (see [S66] for details). Distributions of order k can be multiplied by C^ℓ (matrix) functions if $\ell \geq k$ but not if $\ell \leq k$. As a result, if $A, B \in C^m(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$ has index ν in the sense of this section (so that necessarily $m \geq \nu$), then A_ν and B_ν are of class $C^{m-\nu}$, and b_ν is of order $k + \nu$. In order to transform (5.10) into an explicit ODE by inverting A_ν, we must be able to multiply b_ν by A_ν^{-1}. Since A_ν^{-1} is $C^{m-\nu}$, this requires $m - \nu \geq k + \nu$ and hence $k \leq m - 2\nu$. In other words, in general, the reduction of Section 5 can be performed only if $m \geq 2\nu$, and only when the right-hand side b is a distribution of order $m - 2\nu$ or less. On the other hand, if $b \in W^{-\nu,1}_{\mathrm{loc}}(\mathcal{J})$ it is easily seen that $m \geq \nu$ suffices as in the case of classical solutions.

Problems with index 1 have been studied extensively, at least when no singularities are present and when attention is confined to classical solutions. Since the hypotheses for index 1 used in this section prohibit existence of singularities, it may be of interest to check how they relate to others made in the literature. Throughout this discussion the stronger hypotheses of this section are used under which "ext rge" and "rest ker" become "rge" and "ker", respectively.

Theorem 7.1. *For $A, B \in C^1(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$ the pair (A, B) has index 1 in $\mathcal{J}$ in the sense of this section if and only if*

(i) rank $A(t) = r, \quad \forall t \in \mathcal{J}$

and one of the following four equivalent conditions hold:

(ii-a) $\{x \in \ker A(t),\ B(t)x \in \text{rge } A(t)\} \implies x = 0, \quad \forall t \in \mathcal{J}$,

(ii-b) $A + QB \in C^1(\mathcal{J}; GL(\mathbb{R}^n))$, where $Q = I - P$ and $P \in C^1(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$ is such that $P(t)$ is a projection onto rge $A(t)$, $\forall t \in \mathcal{J}$.

(ii-c) $A + Q(B + \dot{A}) \in C^1(\mathcal{J}; GL(\mathbb{R}^n))$ with Q as in (ii-b).

(ii-d) $A + BK \in C^1(\mathcal{J}; GL(\mathbb{R}^n))$, where $K \in C^1(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$ is such that $K(t)$ is a projection onto $\ker A(t)$, $\forall t \in \mathcal{J}$.

Proof. (a) Suppose that (A, B) has index 1 in $\mathcal{J}$, so that (i) holds. Using the previous notation, we have $A_1(t) \in GL(\mathbb{R}^r)$, $\forall t \in \mathcal{J}$, whence by (3.6) and (3.7), $A(t)C(t)$ has rank r for $t \in \mathcal{J}$. By (i) this occurs if and only if rge $C(t) \cap \ker A(t) = \{0\}$. But rge $C(t) = \ker Q(t)B(t)$ by (3.5), and hence we have $\ker Q(t)B(t) \cap \ker A(t) = \{0\}$, $\forall t \in \mathcal{J}$. Since $Q(t)$ projects onto a complement of rge $A(t)$, it follows that $x \in \ker Q(t)B(t) \cap \ker A(t)$ if and only if $x \in \ker A(t)$ and $B(t)x \in$ rge $A(t)$ which proves (ii-a). Conversely, assume that (i) and (ii-a) hold. Condition (ii-a) and the second part of (7.1) coincide. By (i) and (ii-a), for each $t \in \mathcal{J}$ the null-space of the mapping $(p, x) \in \mathbb{R}^n \times \mathbb{R}^n \mapsto (A(t)x, A(t)p + B(t)x) \in \mathbb{R}^n \times \mathbb{R}^n$ has dimension dim ker $A(t) = n - r$. Hence, its rank equals $\rho = n + r$. But clearly $\rho \leq$ rank $A(t)$ + rank $A(t) \oplus B(t) = r +$ rank $A(t) \oplus B(t)$. This implies that rank $A(t) \oplus B(t) = n$, $\forall t \in \mathcal{J}$ which proves the first part of (7.1) and allows us to speak now of a reduction (A_1, B_1) of (A, B) in $\mathcal{J}$. Then, reversing the steps in the first part of the proof, we arrive at rank $A_1(t) = r$, $\forall t \in \mathcal{J}$; that is, the index of (A, B) in $\mathcal{J}$ is 1.

(b) By Lemma 2.1, condition (i) ensures the existence of $P \in C^1(\mathcal{J}; \mathcal{L}(\mathbb{R}^n))$ such that $P(t)$ is a projection onto rge $A(t)$, $\forall t \in \mathcal{J}$, (actually it is the orthogonal projection). With $Q = I - P$ it is obvious that (ii-a) is equivalent with (ii-b).

(c) In order to show that conditions (ii-b) and (ii-c) are equivalent, suppose first that (ii-b) holds and let $u \in \mathbb{R}^n$ be such that $A(t)u + Q(t)(B(t) + \dot{A}(t))u = 0$ which, obviously, requires that $u \in \ker A(t)$. Now, using $QA \equiv 0$ we see that $Q\dot{A} = -\dot{Q}A$, and hence that $Q(t)B(t)u = 0$. Thus, $A(t)u + Q(t)B(t)u = 0$, whence $u = 0$ and (ii-c) holds. Conversely, if (ii-c) is valid then for $u \in \ker(A(t) + Q(t)B(t))$ we have $u \in \ker A(t) \cap \ker Q(t)B(t)$ and, because of $u \in \ker A(t)$, it follows that $u \in \ker \dot{Q}(t)A(t) = \ker Q(t)\dot{A}(t)$. Thus, $u \in \ker\,(A(t) + Q(t)B(t) + Q(t)\dot{A}(t))$ and (ii-c) implies that $u = 0$ which proves the validity of (ii-b).

(d) For the equivalence of (ii-b) and (ii-d) assume first that (ii-b) is valid and let $u \in \mathbb{R}^n$ be such that $A(t)u + B(t)K(t)u = 0$ and hence $Q(t)B(t)K(t)u = 0$. Then we have $[A(t) + Q(t)B(t)]K(t)u = 0$, since $AK \equiv 0$, and therefore $K(t)u = 0$ and $A(t)u = 0$; that is, $u \in \ker A(t)$. But then $u = K(t)u = 0$ shows that (ii-d) holds. Conversely, suppose that (ii-d) holds so that $A(t)^T + K(t)^T B(t)^T \in GL(\mathbb{R}^n)$. Since $K(t)^T$ projects onto a complement of rge $A(t)^T$, it plays the same role as $Q(t)$ above, but with $A(t)$, $B(t)$ replaced by $A(t)^T$, $B(t)^T$, respectively. Likewise, $Q(t)^T$ plays the role of $K(t)$ in this case. Thus, from the first part, we have $A(t)^T + B(t)^T Q(t)^T \in GL(\mathbb{R}^n)$, and hence $A(t) + Q(t)B(t) \in GL(\mathbb{R}^n)$ by transposition. This shows that (ii-b) holds. $\square$

Conditions (i) and (ii-a) are exactly the index 1 conditions in the sense of [RRh92] for the DAE (1.1). The simplest way to see this is to use the remark in [RRh92] that "index 1" in the sense of that paper is the same as "index 1" in the sense of [RRh91], and that the conditions given in the latter paper are equally valid without the second-order derivative terms. With this simplification, the constant rank condition and condition (2.22) of [RRh91] – which are necessary and sufficient for index 1 – become rank $D_pF(t,x,p) = r$ and $\{u \in \ker D_pF(t,x,p),\ D_xF(t,x,p)u \in \text{rge } D_pF(t,x,p)\} \Longrightarrow u = 0$, respectively, where $F(t,x,p) = A(t)p + B(t)x - b(t)$, so that they coincide with (i) and (ii-a) of Theorem 7.1, respectively.

Remark 7.2. More generally, it can be shown that the index of this section always coincides with the index of [RRh92] for the DAE (1.1). The proof of this, although conceptually simple, involves notational complications and the details will not be given here. But we note that the independence of the index of this paper from the specific choices made at each stage of the reduction can be viewed as a consequence of the fact that different choices merely express the geometric, and hence intrinsic, procedure of [RRh92] in different charts. □

Condition (ii-b) is useful for the study of numerical algorithms: see [RRh93].

As a straightforward verification reveals, conditions (i) and (ii-c) are exactly those of the case "$j_0 = 1$" in Theorem 3.3 of [C87].

Conditions (i) and (ii-d) of Corollary 7.3 are used in [GrM86] as a criterion for "transferability", a concept therefore equivalent to the index 1 condition of this section.

Acknowledgement: We thank Steve Campbell for pointing out to us several contributions related to this work and for helpful discussions.

References

[C87] S.L. Campbell, *A general form for solvable linear time varying singular systems of differential equations*, SIAM J. Math. Anal. **18** (1987), 1101-1115.

[Co82] D. Cobb, *On the solutions of linear differential equations with singular coefficients*, J. Diff. Equations **46** (1982), 311-323.

[CP83] S.L. Campbell and L.R. Petzold, *Canonical forms and solvable singular systems of differential equations*, SIAM J. Alg. Disc. Meth. **4** (1983), 517-521.

[D64] V. Dolezal, *The existence of a continuous basis of certain linear subspaces of E_r which depends on a parameter*, Cas. Pro. Pest. Mat. **89** (1964), 466-468.

[F65] K.O. Friedrichs *Advanced Ordinary Differential Equations*, Gordon and Breach, New York (1965).

[G59] F. R. Gantmacher. *The Theory of Matrices, Vol 1 and 2*, Translation from the Russian, Chelsea Publ Co., New York, NY, 1959.

[Ge93] T. Geerts, *Solvability for singular systems*, Lin. Alg. Appl. **181** (1993), 111-130.

[GrM86] E. Griepentrog and R. Maerz, *Differential–Algebraic Equations and Their Numerical Treatment*, vol. 88, Teubner-Texte zur Mathematik, Leipzig, 1986.

[H64] P. Hartman, *Ordinary Differential Equations*, J. Wiley and Sons, New York (1964).

[K50] T. Kato, *On the adiabatic theorem of quantum mechanics*, J. Phys. Soc. Jap. **5** (1950), 435-439.

[K82] T. Kato, *A Short Introduction to Perturbation Theory for Linear Operators*, Springer-Verlag, New York NY, 1982.

[KuMe92] P. Kunkel and V. Mehrmann, *Canonical forms for linear differential-algebraic equations with variable coefficients*, Technical Report 69, Inst. f. Geometrie und Prakt. Mathematik, RWTH Aachen, Germany, 1992.

[M89a] R. Maerz, *Index 2 differential-algebraic equations*, Results in Math. **15** (1989), 149-171.

[M89b] R. Maerz, *Some new results concerning index-3 differential-algebraic equations*, J. Math. Anal. Appl. **140** (1989), 177-199.

[RRh91] P.J. Rabier and W.C. Rheinboldt, *A general existence and uniqueness theory for implicit differential-algebraic equations*, Diff. and Integral Equations **4** (1991), 563-582.

[RRh92] P.J. Rabier and W.C. Rheinboldt, *A geometric treatment of implicit differential-algebraic equations*, J. Diff. Equations (to appear).

[RRh93] P.J. Rabier and W.C. Rheinboldt, *Finite differences for time dependent linear DAE's*, Appl. Math. Letters (submitted).

[S66] L. Schwartz, *Theorie des Distributions*, Hermann, Paris, France, 1966.

[SiB70] L.M. Silverman and R.S. Bucy, *Generalizations of a theorem of Dolezal*, Math. Systems Theory **4** (1970), 334-339.

[VLKa81] G.C. Verghese, B.C. Levy and T. Kailath, *A generalized state-space for singular systems*, IEEE Trans. Automat. Control AC- **26** (1981), 811-831.

# REPORT DOCUMENTATION PAGE	Form Approved OMB No 0704-0188

1. AGENCY USE ONLY (*Leave blank*)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	10-15-93	TECHNICAL REPORT

4. TITLE AND SUBTITLE
CLASSICAL AND GENERALIZED SOLUTIONS OF TIME-DEPENDENT LINEAR DIFFERENTIAL ALGEBRAIC EQUATIONS

5. FUNDING NUMBERS
ONR-N-00014-90-J-1025
NSF-CCR-9203488

6. AUTHOR(S)
Patrick J. Rabier
Werner C. Rheinboldt

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Department of Mathematics and Statistics
University of Pittsburgh

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
ONR
NSF

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)
A reduction procedure is developed for linear time-dependent differential algebraic equations (DAEs) that transforms their solutions into solutions of smaller systems of ordinary differential equations (ODEs). The procedure applies to classical as well as distribution solutions. In the case of analytic coefficients the hypotheses required for the reduction not only are necessary for the validity of the existence and uniqueness results, but they even allow for the presence of singularities. Straightforward extensions including undetermined systems and systems with non-anal non-analytic coefficients are also discussed.

14. SUBJECT TERMS
linear DAE, classical solut ons, distribution solutions boundary value problems

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
unclassified	unclassified	unclassified	

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to ***stay within the lines*** to meet ***optical scanning requirements***.

Block 1. Agency Use Only *(Leave blank)*.

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C	- Contract	**PR**	- Project
G	- Grant	**TA**	- Task
PE	- Program Element	**WU**	- Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. *(If known)*

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

DOD - See DoDD 5230.24, "Distribution Statements on Technical Documents."
DOE - See authorities.
NASA - See Handbook NHB 2200.2.
NTIS - Leave blank.

Block 12b. Distribution Code.

DOD - Leave blank.
DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
NASA - Leave blank.
NTIS - Leave blank.

Block 13. Abstract. Include a brief *(Maximum 200 words)* factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code *(NTIS only)*.

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.