Index Reduction Methods for Differential-Algebraic Equations

Eberhard Griepentrog

Ingenieurhochschule Berlin, Fachbereich Mathematik 2
Marktstr. 9/13, 0–1134 Berlin

1 Introduction

We consider the differential-algebraic equation (DAE)

$$f(t, x, x') = 0 \ .$$

Gear [Gea88] defines the index of this equation as the smallest number of differentiations applied to (2.1) that is necessary to determine x' uniquely. Using the rank theorem (cf. [Die71, Zor81]) we give a modified version of this definition which is closely related to S. Reich's index idea in the present book. Indeed our index definition is a more analytical notation of Reich's differential-geometric concept.

Based on our index concept we discuss several index reduction methods proposed by Gear [Gea88], Griepentrog [Gri90], Griepentrog&März [GM89], Reich [Rei88] and Rabier&Rheinboldt [RR90].

2 Some Differential-Geometrical Facts

Let $F(t, x, y) : R \times R^n \times R^m \rightarrow R^m$ be a continuously differentiable mapping. Let $DF(t, x, y) : R^m \rightarrow R^m$ be of constant rank $r < m$, where D denotes the differentiation with respect to y. Then the rank theorem provides the representation

$$F(t, x, y) = \phi(t, x, P\psi(t, x, y)) \ , \tag{2.1}$$

where ϕ and ψ are diffeomorphisms and P is the orthogonal projector onto the first r components in R^m. Moreover,

$$H(t, x, y) := DF(t, x, y) = D\phi(t, x, P\psi(t, x, y))PD\psi(t, x, y) \ . \tag{2.2}$$

Now we consider the equation

$$F(t, x, y) = 0 \ . \tag{2.3}$$

Let S be the set of all pairs (t,x) for which a solution y of (2.3) exists. From (2.1) we obtain

$$P\psi(t,x,y) = \phi^{-1}(t,x,0) \ ,$$

that means for $Q := I - P$

$$S = \{(t,x) : Q\phi^{-1}(t,x,0) = 0\} \ . \tag{2.4}$$

We assume S to be non-empty, then S is a differentiable manifold by virtue of (2.4). Moreover, for each $(t,x) \in S$ the equation (2.3) defines a differentiable manifold $M(t,x)$ of the dimension $m-r$ in R^m. Thus we obtain a family SM of manifolds $M(t,x)$. By virtue of (2.1) the equation $P\psi(t,x,y) = \phi^{-1}(t,x,0)$ holds, i. e.

$$M(t,x) = \{y : y = \psi^{-1}(t,x,\phi^{-1}(t,x,0) + Qz), z \in R\} \ . \tag{2.5}$$

The corresponding tangential space and normal space bundles STM and SNM are defined by

$$TM(t,x;y) = \ker(H(t,x,y)) \tag{2.6}$$
$$NM(t,x;y) = \mathrm{im}(H(t,x,y)^T) \ . \tag{2.7}$$

Using (2.2) we obtain

$$ker(H(t,x,y)) = \mathrm{im}([D\psi(t,x,y)]^{-1}Q) \tag{2.8}$$
$$im(H(t,x,y)^T) = \mathrm{im}([D\psi(t,x,y)]^T P) \ . \tag{2.9}$$

The functional $g(t,x,z) = a^T y$ is constant on $M(t,x)$ iff $\mathrm{grad}_x g(t,x,y) = a^T[D\psi(t,x,y(t,x,z))]^{-1}Q = 0$, i. e. iff the vector a belongs to $NM(t,x;y)$ for all $y \in M(t,x)$. In particular the component $e_j^T y$ of $y \in M(t,x)$ depends uniquely on (t,x) iff $e_j \in \mathrm{im}(H(t,x,y)^T)$. In general, the component of y in some subspace U of R^m depends on (t,x) only, iff U is a subspace of $\mathrm{im}(H(t,x,y)^T)$ for all $y \in M(t,x)$.

3 The Index of a Differential-Algebraic Equation

In R^n the general DAE

$$f(t,x,x') = 0 \tag{3.1}$$

is considered, where $f(t,x,y) : R \times R^n \times R^n \rightarrow R^n$ is assumed to be $K+1$ times continuously differentiable with respect to all variables. As a solution of (3.1) we consider each continuously differentiable function $x(t)$ satisfying $f(t,x(t),x'(t)) = 0$ on some t-interval I.

The recursion

$$\begin{aligned} f_0(t,x,y_0) &:= f(t,x,y_0) \\ f_i(t,x,y_0,\ldots,y_i) &:= D_t f_{i-1} + D_x f_{i-1} y_0 + \textstyle\sum_{j=0}^{i-1} D_{y_j} f_{i-1} y_{j+1} \quad \text{for } i = 1,\ldots,K \ , \end{aligned} \tag{3.2}$$

defines the sequence $F_0, F_1, \ldots, F_K$ of compound functions

$$F_k(t,x,\overline{y}_k) := \begin{pmatrix} f_0(t,x,y_0) \\ f_1(t,x,y_0,y_1) \\ \vdots \\ f_k(t,x,y_0,\ldots,y_k) \end{pmatrix}$$

with $\overline{y}_k := (y_0^T, y_1^T, \ldots, y_k^T)^T$. The corresponding Jacobians $H_k(t,x,\overline{y}_k)$ with respect to $\overline{y}_k$ are given by

$$H_k(t,x,\overline{y}_k) := \begin{pmatrix} A & 0 & 0 & \ldots 0 \\ A_{10} & A & 0 & \ldots 0 \\ A_{20} & A_{21} & A & \ldots 0 \\ \vdots & \vdots & \vdots & \ddots \vdots \\ A_{k0} & A_{k1} & A_{k2} & \ldots A \end{pmatrix}$$

where $A_{ij} := \partial f_i / \partial y_j$ and $A := \partial f_0 / \partial y_0$. We assume all $H_k(t,x,\overline{y}_k)$ to have constant ranks r $(k = 0,1,\ldots,K)$.

Let S_k be the manifold of all pairs (t,x) for which

$$F_k(t,x,\overline{y}_k) = 0 \tag{3.3}$$

is solvable; we call S_k the constraint manifold of the order k. For each $(t,x) \in S_k$ we obtain the manifold $M(t,x) \subseteq R^{(k+1)n}$ of all vectors $\overline{y}_k$ solving (3.3). Its tangential and normal spaces are

$$\begin{aligned} TM_k(t,x;\overline{y}_k) &= \ker(H_k(t,x,\overline{y}_k)) = \mathrm{im}([D\psi_k(t,x,\overline{y}_k)]^{-1} Q_k) \\ NM_k(t,x;\overline{y}_k) &= \mathrm{im}(H_k(t,x,\overline{y}_k)^T) = \mathrm{im}([D\psi_k(t,x,\overline{y}_k)]^T P_k) \end{aligned} \tag{3.4}$$

by means of the decomposition

$$F_k(t,x,\overline{y}_k) = \phi_k(t,x,P_k\psi_k(t,x,\overline{y}_k)) \ . \tag{3.5}$$

Trivially, the definition of the constraint manifolds implies the inclusion chain

$$S_o \supseteq S_1 \supseteq \cdots \supseteq S. \tag{3.6}$$

On the other hand,

$$E_k M_k(t,x) \subseteq M_{k-1}(t,x) \tag{3.7}$$

obviously holds, for $k = 1,\ldots,K$, where the mapping $E_k : R^{(k+1)n} \to R^{kn}$ is defined by

$$E_k(y_0^T, \ldots, y_k^T)^T = (y_0^T, \ldots, y_{k-1}^T)^T \ .$$

Moreover, the inclusion

$$E_k \ker(H_k(t,x,\overline{y}_k)) \subseteq \ker(H_{k-1}(t,x,E_k\overline{y}_k)) \tag{3.8}$$

is valid since $H_k(t,x,\overline{y}_k)\overline{z}_k = 0$ implies

$$H_{k-1}(t,x,E_k\overline{y}_k)E_k\overline{z}_k = 0.$$

Thus we obtain two further inclusion chains for $(t,x) \in S_k$ using $\overline{E}_k := E_1 E_2 \dots E_k$ and $H_i := H_i(t,x,E_{i+1}\dots E_k \overline{y}_k)$:

$$M_0(t,x) \supseteq \overline{E}_1 M_1(t,x) \supseteq \overline{E}_2 M_2(t,x) \supseteq \cdots \supseteq \overline{E}_k M_k(t,x) \tag{3.9}$$

$$\ker(H_0) \supseteq \overline{E}_1 \ker(H_1) \supseteq \overline{E}_2 \ker(H_2) \supseteq \cdots \supseteq \overline{E}_k \ker(H_k) \ . \tag{3.10}$$

Let U_k be the subspace of $R^{(k+1)n}$ spanned by the unit vectors $e_1, \dots, e_n$. Then the partial vectors y_0 of $\overline{y}_k$ are determined uniquely (3.3) iff $U_k \subseteq NM_k(t,x,\overline{y}_k)$ for all $\overline{y}_k \in M_k(t,x)$, i. e. iff $\overline{E}_k \ker(H_k) = \{0\}$. Hence, $\overline{E}_k M_k(t,x)$ is a singleton iff $\overline{E}_k \ker(H_k) = \{0\}$.

Definition. We call the mapping $f(t,x,y)$ an *index-k-mapping* (with $k \le K$) if S_k is non-empty and $\bar{E}_k M_k(t,x)$ is a singleton for all $(t,x) \in S_k$, and if k is the smallest integer with these properties.

Remark. Using the above considerations the singleton requirement for $\bar{E}_k M_k(t,x)$ is equivalent to $\overline{E}_k \ker(H_k) = \{0\}$ or $U_k \subseteq NM_k(t,x;\overline{y}_k)$ for all $\overline{y}_k \in M_k(t,x)$.

As an immediate corollary of our definition we obtain the following fundamental statement.

Theorem 1. *If $f(t,x,y)$ is an index-k mapping, then a vector field $v(t,x)$ satisfying $f(t,x,v(t,x)) = 0$ is defined on S_k. It is the only vector field fulfilling $\overline{E}_k \overline{y}_k = v(t,x)$ with $F_k(t,x,\overline{y}_k) = 0$.*

Using the decomposition (3.5) we can determine the vector field $v(t,x)$ and its definition manifold S_k explicitly:

$$v(t,x) = \overline{E}_k \psi_k^{-1}(t,x,\phi_k^{-1}(t,x,0)) \tag{3.11}$$

$$S_k = \{(t,x) : Q_k \phi_k^{-1}(t,x,0) = 0\} \ . \tag{3.12}$$

A direct consequence of the construction rules for $F_k(t,x,\overline{y}_k)$ is the following theorem, which leads back to the DAEs.

Theorem 2. *If the function $x(t)$ solves $f(t,x,x') = 0$ and if it is $k+1$ times continuously differentiable, then the function pair $(x(t), \overline{y}_k(t))$ with $y_i(t) = x^{(i+1)}(t)$ $(i = 0,1,\dots,k)$ satisfies*

$$F_k(t,x(t),\overline{y}_k(t)) = 0 \ , \tag{3.13}$$

i. e. $x(t)$ proceeds in the constraint manifold S_k.

If a vectorfield $v(t,x)$ with sufficient differentiabiltiy properties is defined, on a manifold $S \subseteq R \times R^n$ we can construct vectorfields $v_1(t,x), v_2(t,x), \dots$ by the recursion

$$\begin{aligned} v_0(t,x) &:= v(t,x), \\ v_i(t,x) &:= D_t v_{i-1}(t,x) + D_x v_{i-1}(t,x) v(t,x) \end{aligned} \tag{3.14}$$

and arrange them as compound vectorfields $\overline{v}_k(t,x) := (v_0(t,x)^T, v_1(t,x)^T, \dots, v_k(t,x)^T)^T$.

Theorem 3. *If a vectorfield $v(t,x)$ satisfying $F_k(t,x,\overline{v}_k(t,x)) = 0$ is uniquely defined, on the constraint manifold S_k then the solutions of all initial value problems $x' = v(t,x)$, $(t_0, x(t_0)) \in S_k$ proceed in S_k and solve $f(t,x,x') = 0$.*

Proof. From $x' = v(t,x)$ and the construction of the $v_i(t,x)$ it follows that $x^{(i)}(t) = v_{i-1}(t,x(t))$, i. e. $F_k(t,x(t),\overline{x}_k(t)) = 0$ (with $\overline{x}_k(t) := (x'(t)^T, \ldots, x^{(k+1)}(t)^T)^T$) implying $f(t,x(t),x'(t)) = 0$ and $(t,x(t)) \in S_k$. □

Due to Theorem 1 the conditions of Theorem 3 are fulfilled if $f(t,x,y)$ is an index-k-mapping. Therefore, it is reasonable to accept Gear's [Gea88] index idea for DAEs in a slightly more precise form.

Definition. The DAE $f(t,x,x') = 0$ is called an *index-k DAE* if $f(t,x,y)$ is an index-k-mapping.

Then we can summarize the above results follows:

Theorem 4. *If $f(t,x,x') = 0$ is an index-k DAE, then all solutions proceed in the differentiable constraint manifold S_k. On S_k a vectorfield $v(t,x)$ is defined with the following properties:*
i. $v(t,x)$ is continuously differentiable,
ii. the solutions $x(t)$ of $f(t,x,x') = 0$ are identical with the solutions of the initial value problems $x'(t) = v(t,x(t))$ with $(t_0,x(t_0)) \in S_k$. S_k and $v(t,x)$ can be determined explicitly using the decomposition

$$\begin{aligned} F_k(t,x,\overline{y}_k) &= \phi_k(t,x,P_k\psi_k(t,x,\overline{y}_k)) \\ S_k &= \{(t,x) : Q_k\phi_k^{-1}(t,x,0) = 0\} \\ v(t,x) &= \overline{E}_k\psi_k^{-1}(t,x,\phi_k^{-1}(t,x,0)) \ . \end{aligned} \tag{3.15}$$

4 Linear Problems

In the linear case, the equation

$$A(t)x' + B(t)x = q(t) \tag{4.1}$$

leads to the system

$$H_k(t)\overline{y}_k + B_k(t)x = q_k(t) \tag{4.2}$$

with $H_k(t) = (A_{ij}(t))_{i,j=0,\ldots,k}$

$$A_{ij}(t) = \binom{i}{j+1} B^{(i-j-1)}(t) + \binom{i}{j} A^{(i-j)}$$

$$B_k(t) := \begin{pmatrix} B(t) \\ B'(t) \\ \vdots \\ B^{(k)}(t) \end{pmatrix}, \quad q_k(t) := \begin{pmatrix} q(t) \\ q'(t) \\ \vdots \\ q^{(k)}(t) \end{pmatrix} .$$

In case of constant coefficients we have $A_{ii} = A$, $A_{ij} = B$ for $i = j+1$ and $A_{ij} = 0$ else. The Kronecker decomposition yields, with $J_k = (e_2 : e_3 : \ldots : e_k : 0)$, $EAF = \operatorname{diag}(I_s, J_{k_1}, \ldots, J_{k_m})$, $EBF = \operatorname{diag}(W, I_{k_1}, \ldots, I_{k_m})$. Using the permutation matrix $L := \operatorname{diag}(I_s, L_{k_1}, \ldots, L_{k_m})$ with $L_k := (e_2 : e_3 : \ldots : e_k : e_1)^T$ we obtain $\overline{A} := LEAF = \operatorname{diag}(I_s, U_{k_1}, \ldots, U_{k_m})$, $\overline{B} := LEBF = \operatorname{diag}(W, L_{k_1}, \ldots, L_{k_m})$, where $U_k = I_k - V_k := \operatorname{diag}(I_{k-1}, 0)$ and the last row of $L_k^j V_k$ is 0 for $j \not\equiv 0 \bmod k$ and e_k^T for $j \equiv 0 \bmod k$. Passing on to the matrix

$$\begin{pmatrix} \overline{A} & 0 & 0 & \ldots & 0 \\ \overline{B} & \overline{A} & 0 & \ldots & 0 \\ 0 & \overline{B} & \overline{A} & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \overline{A} \end{pmatrix}$$

and multiplying it by

$$\begin{pmatrix} I & 0 & 0 & 0 \ldots \\ -L_2 & I & 0 & 0 \ldots \\ L_3 & -L_2 & I & 0 \ldots \\ -L_4 & L_3 & -L_2 & I \ldots \\ \vdots & \vdots & \vdots & \vdots \ddots \end{pmatrix}$$

we receive

$$\overline{H}_k := \begin{pmatrix} \overline{A} & 0 & 0 & \ldots 0 \\ \overline{C} & \overline{A} & 0 & \ldots 0 \\ -L\overline{C} & \overline{C} & \overline{A} & \ldots 0 \\ L^2\overline{C} & -L\overline{C} & \overline{C} & \ldots 0 \\ \vdots & \vdots & \vdots & \ddots \vdots \\ (-L)^{k-1}\overline{C} & (-L)^{k-2}\overline{C} & (-L)^{k-3}\overline{C} & \ldots \overline{A} \end{pmatrix}$$

with $\overline{C} := \overline{B} - L\overline{A} = \operatorname{diag}(W - I_s, L_{k_1}V_{k_1}, \ldots, L_{k_m}V_{k_m})$. Thus we derived the well-known

Theorem 5. *The index k of the DAE (4.1) is the greatest of the integers k_i $(i = 1, \ldots, m)$.*

Proof. Obviously

$$\begin{aligned} e_{s+k_1} &\in \operatorname{im}(((-L)^{k_1-1}\overline{C} : (-L)^{k_1-2}\overline{C} : \ldots)^T) \\ e_{s+k_1+k_2} &\in \operatorname{im}(((-L)^{k_2-1}\overline{C} : (-L)^{k_2-2}\overline{C} : \ldots)^T) \\ &\ldots \\ e_{s+(k+1)n} &\in \operatorname{im}(((-L)^{k_m-1}\overline{C} : (-L)^{k_m-2}\overline{C} : \ldots)^T) \ , \end{aligned}$$

while $e_j \in \operatorname{im}((\overline{A} : 0 : \ldots : 0)^T)$ for all the other unit vectors. □

Returning to the general linear case we obtain, due to the rank theorem, $H_k(t) = V_k(t)P_kW_k(t)$, i. e. (4.2) implies.

$$P_kW_k(t)\overline{y} + V_k(t)^{-1}[B_k(t)x - q_k(t)] = 0.$$

Hence, the constraint manifold S is defined by the equation

$$Q_kV_k(t)^{-1}[B_k(t)x - q_k(t)] = 0 \ , \tag{4.3}$$

therefore it is formed by a differentiable function which assigns to each value of t an affine subspace of R^n. Equation (4.3) is equivalent to

$$\overline{R}_k(t)[B_k(t)x - q_k(t)] = 0 \ , \tag{4.4}$$

where $\overline{R}_k(t) := V_k(t)Q_kV_k(t)^{-1}$ is a continuously differentiable projector along $\mathrm{im}(H_k(t))$.

Theorem 6. *The constraint manifold S of the DAE (4.1) is determined by the equation*

$$R_k(t)[B_k(t)x - q_k(t)] = 0 \ , \tag{4.5}$$

where $R_k(t)$ is an arbitrary projector along $\mathrm{im}(H_k(t))$.

Proof. For any two projectors R and $\overline{R}$ along the same subspace the relation $R\overline{R} = R$ holds, hence the multiplication of (4.4) by $R_k(t)$ yields the assertion.

□

The manifolds $M_k(t,x)$ are defined by the equation

$$\begin{aligned}\overline{y}_k &= W_k(t)^{-1}\{V_k(t)^{-1}[q_k(t) - B_k(t)x] + Q_k\overline{z}_k\} \\ &= G_k(t)^{-1}[q_k(t) - B_k(t)x] + \overline{Q}_k(t)\overline{v}_k\end{aligned}$$

using $G_k(t) := V_k(t)W_k(t)$ and the projector $\overline{Q}_k(t) := W_k(t)^{-1}Q_kW_k(t)$ onto $\ker(H_k(t))$.

Theorem 7. *The manifolds $M_k(t,x)$ are affine subspaces parallel to* $\ker(H_k(t))$ *which depend linearly on x:*

$$M_k(t,x) = \overline{p}_k(t,x) + \ker(H_k(t)), \tag{4.6}$$

where $\overline{p}_k(t,x) := G_k(t)^{-1}[q_k(t) - B_k(t)x]$. The index of the DAE (4.1) is the smallest integer k for which $\overline{E}_k\overline{Q}_k(t) = 0$ holds. Hence, the equivalent initial value problem and the corresponding vectorfield are

$$x' = \overline{E}_kG_k(t)^{-1}[q_k(t) - B_k(t)x] \tag{4.7}$$

with $(t_0, x(t_0)) \in S_k$ and $v(t,x) = \overline{E}_k\overline{p}_k(t,x)$, respectively.

For the foundation of reduction methods the following theorem characterizing the manifolds S_1 and $M_1(t,x)$ is important.

Theorem 8. *Together with the equations*

$$\begin{aligned} &A(t)y_0 + B(t)x = q(t) \\ &A(t)y_1 + (A'(t) + B(t))y_0 + B'(t)x = q'(t) \end{aligned} \tag{4.8}$$

defining the manifolds S_1 *and* $M_1(t,x)$ *we consider by means of any differentiable projector* $R(t)$ *along* $\operatorname{im}(A(t))$, *the equation*

$$[A(t) + R(t)B(t)]y_0 + [B(t) + (RB)'(t)]x = q(t) + (Rq)'(t) \tag{4.9}$$

defining the manifolds

$$\overline{S}_1 := \{(t,x) : \text{ (4.9) is solvable}\}$$

and

$$\overline{M}_1(t,x) := \{y_0 : y_0 \text{ solves (4.9)}\} \ .$$

Then $\overline{S}_1 \cap S_0 = S_1$ *and* $\overline{M}_1(t,x) \cap M_0(t,x) = \overline{E}_1 M_1(t,x)$ *hold. Moreover,* $\overline{E}_1 \ker \begin{pmatrix} A & 0 \\ A' + B & A \end{pmatrix} = \ker(A + RB)$ *is valid.*

Proof. $(t,x) \in \overline{S} \cap S_0$ and $y_0 \in \overline{M}_1 \cap M_0$ mean $(A + RB)y_0 = q - Bx + (Rq)' - (RB)'x$ and $RA'y_0 = -R'Ay_0 = R'(-q + Bx)$, i. e. $[A + R(A' + B)]y_0 = q - Bx + R(q' - B'x)$ and, hence, $R(A' + B)y_0 = R(q' - B'x)$ by multiplication by R. Consequently, $Ay_1 + (A' + B)y_0 = q' - B'x$ holds for a suitable y_1, i. e. $(t,x) \in S_1$. On the other hand, $(t,x) \in S_1$ implies $RA'y_0 = -R'Ay_0 = -R'(q - Bx) = -(Rq)' + Rq' + (RB)'x - RB'x$ and $R(A' + B)y_0 = -RB'x + Rq'$, i. e. $RBy_0 = -(RB)'x + (Rq)'$. Addition to the first equation of (4.8) yields (4.9), i. e. $(t,x) \in \overline{S}_1$. Thus the assertion concerning the constraint manifold is true. Moreover, each y_0 solving (4.9), $(y_0^T, y_1^T)^T$ belongs to $M_1(t,x)$ and vice versa, i. e. we have $\overline{E}_1 M_1(t,x) = \overline{M}_1(t,x)$.

In order to prove the third assertion we remark first that $I + RR'$ is a nonsingular matrix since $RR' = R'(I - R)$ and therefore $(I + RR')(I - RR') = I - RR'RR' = I$. Hence the matrix $\begin{pmatrix} R & I - R \\ I - R + R' & R \end{pmatrix}$ is nonsingular by virtue of

$$\begin{pmatrix} R & I - R \\ I - R + R' & R \end{pmatrix} \begin{pmatrix} R & I - R \\ I - R & R \end{pmatrix} = \begin{pmatrix} I & 0 \\ R'R & I + RR' \end{pmatrix}.$$

Hence,

$$\ker \begin{pmatrix} A & 0 \\ A' + B & A \end{pmatrix} = \ker \begin{pmatrix} R & I - R \\ I - R + R' & R \end{pmatrix} \begin{pmatrix} A & 0 \\ A' + B & A \end{pmatrix} =$$

$$\ker \begin{pmatrix} (I - R)(A' + B) & A \\ A + RB & 0 \end{pmatrix}.$$

Thus we have $(A + RB)z = 0$ iff $Az = 0$ and $(A' + B)z + Aw = 0$ for a suitable w, because of $\operatorname{im}((I - R)(A' + B)) \subseteq \operatorname{im}(I - R) = \operatorname{im}(A)$. □

Remark. The third assertion reflects the fact that the tangential spaces of $\overline{M}_1(t,x)$ and $\overline{E}_1 M_1(t,x)$ coincide. The statements of Theorem 8 remain valid if we replace (4.9) by one of the equations

$$[A+RB](t)y_0 + [(I-R)B+(RB)'](t)x = [I-R(t)]q(t)+(Rq)'(t) \tag{4.10}$$

or

$$[(I-R')A+RB](t)y_0 + [(I-R)+RB'](t)x = [I-R(t)]q(t)+Rq'(t)\ , \tag{4.11}$$

since, (4.10) and (4.11) follow from the relations $R(Bx-q)=0$ and $R'(Ay_0+Bx-q)=0$ fulfilled on S_0.

Lemma 9. *Let $g(t)$ and $R(t)$ be differentiable vector and projector functions, respectively. Then $g(t)=0$ iff $R(t_0)g(t_0)=0$ for some t_0 and one of the following relations holds:*
1. $[g+(Rg)'](t)=0$,
2. $[(I-R)g+(Rg)'](t)=0$,
3. $[g+Rg'](t)=0$,
4. $[(I-R)g+Rg'](t)=0$.

Proof. One direction is obvious, for the other one we use $R(Rg)' = RR'g + Rg' = R'(I-R)g+Rg' = (Rg)'-R'(Rg)$ or $Rg' = (Rg)'-R'g = (Rg)'-R'(Rg)$ regarding $(I-R)g=0$. Thus the function $z(t) := (Rg)(t)$ satisfies the homogeneous explicit differential equation $z' = (R'-I)z$ in the first case and $z' = R'z$ in the other cases. The initial condition $z(t_0)$ then implies $z(t)\equiv 0$, leading to the assertion $g(t)\equiv 0$. □

Theorem 10. *The function $x(t)$ solves $A(t)x'+B(t)x=q(t)$ iff it satisfies, for any projector $R(t)$ along $\mathrm{im}(A(t))$ and for some t_0, the initial condition*

$$R(t_0)[B(t_0)x(t_0)-q(t_0)]=0 \tag{4.12}$$

and it is a solution of one of the following DAEs:

$$[A+RB]x' + [B+(RB)']x = q+(Rq)' \tag{4.13}$$
$$[A+RB]x' + [(I-R)B+(RB)']x = (I-R)q+(Rq)' \tag{4.14}$$
$$[(I-R')A+RB]x' + [B+RB']x = q+Rq' \tag{4.15}$$
$$[(I-R')A+RB]x' + [(I-R)B+RB']x = (I-R)q+Rq'\ . \tag{4.16}$$

Proof. We have to apply the four relations of Lemma 9 to the function $g(t) := A(t)x'(t)+B(t)x(t)-q(t)$ and to regard $RA' = -R'A$ only. □

Thus, if we guarantee an initial pair $(t_0, x(t_0))$ on the constraint manifold S_0, we can solve one of the equations (4.13), (4.14), (4.15), (4.16) instead of the original problem (4.1), and the calculated solution proceeds on S_0 automatically. The new problem has the constraint manifold $\overline{S}_1$, i. e. regarding the initial restriction (4.12) for some t_0 we obtain the actual constraint manifold $\overline{S}_1 \cap S_0 = S_1$. Since, moreover, $\overline{M}_1(t,x) = \overline{E}_1 M_1(t,x) \cap M_0(t,x)$ the index of the new problem is reduced by 1. Therefore the transfer from the DAE (4.1) to one of the DAEs (4.13), (4.14), (4.15), (4.16) with the initial restriction (4.12) is called an index reduction method. In particular, if the original problem is an index-1 DAE, the reduced one is practically an explicit ordinary differential equation, because of $\ker(A + RB) = \ker((I - R')A + RB) = 0$ according to Theorem 8.

The reduction methods via (4.13), (4.14) are proposed by Gear [Gea88], Griepentrog [Gri90] and Griepentrog&März [GM89], while the equation (4.16) is used by Rheinboldt [RR90]. Another, essentially different method for the index reduction is suggested by Reich [Rei88] (cf. also his article in this book). Reich's ideas are presented in a strongly differential geometric language, we will formulate them in a more analytical form suitable to our preceding considerations.

The constraint manifold S_0 of (4.1) is determined by the equation $(RB)(t)x - (Rq)(t) = 0$, and its tangential space is $TS_0(t,x) = \{\begin{pmatrix} u \\ v \end{pmatrix} : u \in R,$ $v \in R^n, u[(RB)'x - (Rq)'] + (RB)v = 0\}$. Hence, we have for all $(t,x) \in S_0$

$$(RB)'(t)x - (Rq)'(t) \in \mathrm{im}(R(t)B(t)) \ ,$$

i. e. there is a $z(t,x)$ with

$$(RB)'(t)x - (Rq)'(t) = -(RB)(t)z \ . \tag{4.17}$$

z is not defined uniquely, because of $RBz = RB\overline{z}$ for $z - \overline{z} \in \ker(RB)$. However, if $Q(t)$ is any projector onto $\ker(R(t)B(t))$, then the equations (4.17) and $Q(t)z = 0$ determine $z(t,x)$ uniquely. We obtain

$$z(t,x) = -[(RB)^T RB + Q^T Q]^{-1}(RB)^T[(RB)'x - (Rq)'] = Z(t)x + w(t) \ . \tag{4.18}$$

For the solution $x(t)$ of (4.1) the differentiation of $(RBx - Rq)(t) = 0$ yields $(RB)(t)x'(t) + (RB)'(t)x(t) - (Rq)'(t) = 0$, hence we obtain $(RB)(t)[x'(t) - z(t,x)] = 0$, i. e. $Q(t)x'(t) = Q(t)[x'(t) - z(t,x)] = x'(t) - z(t,x)$ due to $Q(t)z(t,x) = 0$. Now, as a reduced equation we consider the DAE (cf. [Gri90])

$$A(t)Q(t)x' + [B(t) + A(t)Z(t)]x = q(t) - A(t)w(t) \ . \tag{4.19}$$

In order to justify the notion "reduced" we have to show that (4.19) has the constraint manifold $\overline{S} = S_1$ and that $\overline{M}(t,x) = \overline{E}_1 M_1(t,x)$. Here $\overline{S}$ and $\overline{M}(t,x)$ are defined by

$$\overline{S} := \{(t,x) : [A(Qy + z) + Bx - q](t) = 0 \text{ is solvable}\}$$

and

$$\overline{M}(t,x) := \{y : [A(Qy + z) + Bx - q](t) = 0 \text{ and } (RB)(t)(y - z) = 0\} \ .$$

We use the following lemma.

Lemma 11. *Assume*

$$(RB)z + (RB)'x = (Rq)' \text{ and } Qz = 0. \tag{4.20}$$

Then the equation systems

$$AQy + Az + Bx - q = 0, \quad RB(y - z) = 0 \tag{4.21}$$

and

$$Ay + Bx - q = 0, \quad R\{(A' + B)y + B'x - q'\} = 0 \tag{4.22}$$

for the determination of y are equivalent.

Proof. Let y satisfy (4.21). Then $y = z + Qw$ and $A(Qy + z) = A(Qw + z) = Ay$, i. e. $Ay + Bx - q = 0$. (4.20) implies $0 = RBz + R'(Bx - q) + R(B'x - q') = RBz - R'Ay + R(B'x - q') = RBy - RBQw + RA'y + RB'x - Rq' = R\{(A' + B)y + B'x - q'\}$. Now let y solve (4.22), then $RBy + (RB)'x = -RA'y - RB'x + Rq' + (RB)'x = R'\{Ay + Bx - q\} + (Rq)' = (Rq)'$ yields $RB(y - z) = 0$. Hence $y - z = Q(y - z)$, i. e. $y = z + Qy$ and $AQy + Az + Bx - q = 0$. □

Clearly, the lemma verifies $S_1 = \overline{S}$ since the solvability of (4.21) and (4.22) are equivalent and $R\{(A' + B)y + B'x - q'\} = 0$ means $(A' + B)y + B'x - q' \in \mathrm{im}(A)$. Moreover, $\overline{M}(t, x) = \overline{E}_1 M_1(t, x)$ holds. We summarize our investigations in

Theorem 12. *The transfer from the DAE (4.1) to the DAE (4.19) reduces the index by 1.*

The index reduction based on Theorem 12 has an essential advantage compared with that of Theorem 10: Since each solution $x(t)$ of (4.1) has the property $x'(t) = Q(t)x'(t) + z(t, x)$, the variety of the solutions of (4.19) is the same as that of (4.1), while the DAEs (4.13), (4.14), (4.15), (4.16) have an enlarged solution variety, which must be restricted by an initial condition (4.12). The advantage of the reduction using Theorem 10 is that it leads to an explicit ordinary differential equation if the original problem (4.1) is an index-1-DAE, since then $(A + RB)y + (B + (RB)')x = q + (Rq)'$ defines y uniquely. In (4.21), also in the index-1 case the equation $RB(y - z) = 0$ is necessary for the uniqueness. Finally we remark that the reduction approach using the function $z(t, x) = Z(t)x + w(t)$ is more convenient than the original one proposed by Reich, because it leads to a linear DAE (4.19) again.

5 Reduction Methods for Nonlinear Problems

In this chapter we consider the generalization of the investigated index reduction methods to nonlinear problems. For quasilinear DAEs

$$A(t, x)x' + q(t, x) = 0 \tag{5.1}$$

the first two equations of Lemma 9 generate the DAEs

$$[A + D_x(Rq)](t, x)x' + [q + D_t(Rq)](t, x) = 0 \tag{5.2}$$
$$[A + D_x(Rq)](t, x)x' + [(I - R)q + D_t(Rq)](t, x) = 0 \tag{5.3}$$

for any projector $R(t, x)$ along $\mathrm{im}(A(t, x))$. The application of the other two equations in Lemma 9 is not reasonable because we lose the quasilinearity in the arising DAEs.

Theorem 13. *The transfer from the DAE (5.1) to one of the DAEs (5.2) or (5.3) with the initial restriction*

$$(Rq)(t_0, x(t_0)) = 0 \tag{5.4}$$

for some t_0 reduces the index by 1.

Proof. The application of Lemma 9 to the function $g(t) := A(t, x(t))x'(t) + q(t, x(t))$ shows that (5.2), (5.4) or (5.3), (5.4) imply $(Rq)(t, x(t)) = 0$, i. e. $(t, x(t)) \in S_0$. If we define $\overline{S}_1 := \{(t, x) \ : \exists y_0([A + D_x(Rq)](t, x)y_0 + [q + D_t(Rq)](t, x) = 0)\}$, then we have to prove $S_1 = \overline{S}_1 \cap S_0$, where S_1 consists of all pairs (t, x) for which $Ay_0 + q = 0$, $Ay_1 + [D_x(Ay_0 + q) + D_tA]y_0 + D_tq = 0$ is solvable. These equations can be replaced by

$$Ay_0 + q = 0, R[D_x(Ay_0 + q) + D_tA]y_0 + RD_tq = 0 \ . \tag{5.5}$$

From (5.5) it follows, for $y_0 \in \overline{E}_1M_1(t, x)$, $0 = [D_t(RA) + D_x(RAy_0 + Rq) - (D_tR)A]y_0 - (D_xR)y_0(Ay_0 + q) + D_t(Rq) - (D_tR)q = -(D_tR)(Ay_0 + q) + D_x(Rq)y_0 + D_t(Rq) = [A + D_x(Rq)]y_0 + q + D_t(Rq)$, i. e. $(t, x) \in \overline{S}_1 \cap S_0$ and $y \in \overline{M}_1(t, x) \cap M_0(t, x)$.

On the other hand, $[A + D_x(Rq)]y_0 + q + D_t(Rq) = 0$ implies, for $Ay_0 + q = 0$ the equations (5.5) and therefore $(t, x) \in S_1$ and $y \in \overline{E}_1M_1(t, x)$. Consequently, $S_1 = \overline{S}_1 \cap S_0$ and $M_0(t, x) \cap \overline{M}_1(t, x) = \overline{E}_1M_1(t, x)$ hold. □

The reduction method from Theorem 12 can also be generalized easily to quasilinear problems. Then $z(t, x)$ is defined by the equations $D_x(Rq)(t, x)z = -D_t(Rq)(t, x)$, $Q(t, x)z = 0$, where $Q(t, x)$ is any projector onto $\ker(D_x(Rq)(t, x))$, and we obtain, because of $D_x(Rq)x' + D_t(Rq) = 0$ the reduced DAE

$$A(t, x)[Q(t, x)x' + z(t, x)] + q(t, x) = 0 \ , \tag{5.6}$$

which is quasilinear again. $z(t, x)$ can be calculated as

$$z(t, x) = -\{[D_x(Rq)^T D_x(Rq) + Q^T Q]^{-1} D_x(Rq)^T D_t(Rq)\}(t, x) \ . \tag{5.7}$$

Lemma 14. *Assume*

$$[D_x(Rq)z + D_t(Rq)](t, x) = 0 \textit{ and } (Qz)(t, x) = 0 \ . \tag{5.8}$$

Then the equation systems

$$\begin{aligned} A(t, x)[Q(t, x)y + z(t, x)] + q(t, x) = 0 \\ D_x(Rq)(t, x)(y - z(t, x)) = 0 \end{aligned} \tag{5.9}$$

and

$$\begin{aligned} A(t,x)y+q(t,x)=0, \\ R(t,x)\{[D_tA(t,x)+D_xA(t,x)y+D_xq(t,x)]y+D_tq(t,x)\}=0 \end{aligned} \tag{5.10}$$

for the determination of y are equivalent.

Proof. Let y satisfy (5.9), then $y = z + Qw$ for some w and $A(Qy+z) = A(Qw+z) = Ay$, i. e. $Ay+q=0$. (5.8) implies $0 = D_x(Rq)z + D_t(Rq) = D_x(Rq)(y-Qw)+(D_tR)q+RD_tq = D_x(Rq)y+R(D_tA)y+RD_tq = R(D_xAy)y+R(D_xq)y+R(D_tA)y+RD_tq$ by virtue of $R(D_xAy) = D_x(RA)y-(D_xR)Ay = (D_xR)q = D_x(Rq)-R(D_xq)$. Hence (5.9) implies (5.10), now we assume y to solve (5.10). Then, by means of $0 = R\{[D_tA+D_xAy+D_xq]y+D_tq\} = D_x(Rq)y+R(D_tA)y+R(D_tq) = D_x(Rq)y+D_t(RA)y-(D_tR)Ay+D_t(Rq)-(D_tR)q = D_x(Rq)y+D_t(Rq)$, we obtain the relation $D_x(Rq)(y-z)=0$. From this fact it follows immediately that $y = Qy+z$, i. e. $A(Qy+z)+q=0$. □

Again, the lemma guarantees the following statement.

Theorem 15. *The transfer from the DAE (5.1) to the DAE (5.6) reduces the index by 1.*

Remark. The condition $D_x(Rq)(x'-z)=0$ is fulfilled automatically since $(Rq)(t,x(t))=0$ implies $D_x(Rq)x'+D_t(Rq)=0$.

Finally we investigate the general case

$$f(t,x,x')=0. \tag{5.11}$$

For this problem, Rabier & Rheinboldt, [RR90] propose the transfer to

$$P(x,x')f(x,x')+[I-P(x,x')]D_tf(x,x')x'=0 \tag{5.12}$$

in the autonomous case $f(x,x')=0$ where $P(x,y)$ is the orthoprojector onto $\mathrm{im}(D_xf(x,y))$. The DAE (5.12) is equivalent to $f(x,x')=0$ if an initial restriction $f(x(t_0),x'(t_0))=0$ for some t_0 is fulfilled. Namely, the application of Eq. 4 in Lemma 9 to $g(t) := f(x(t),x'(t))$ provides immediately (5.12) for $P(x,y) = I-R(x,y)$ because of $(I-P)D_yf=0$.

We generalize the suggestion of [RR90] to the non-autonomous case and arbitrary projectors $R(t,x,y)$ along $\mathrm{im}(D_yf(t,x,y))$. Then the reduced equation corresponding to (5.11) is proposed to be

$$[I-R(t,x,x')]f(t,x,x')+R(t,x,x')[D_xf(t,x,x')x'+D_tf(t,x,x')]=0 \tag{5.13}$$

under the initial restriction

$$f(t_0,x(t_0),x'(t_0))=0 \tag{5.14}$$

for some t_0. Clearly Lemma 9 guarantees the equivalence of (5.11) and (5.13), (5.14) too.

The constraint manifolds S_0, S_1 of (5.11) and $\overline{S}$ of (5.13) are given by

$$\begin{aligned} S_0 &= \{(t,x) : \exists y(f(t,x,y) = 0)\} \\ S_1 &= \{(t,x) : \exists y, y_1(f(t,x,y) = 0 \text{ and } D_t f + (D_x f)y + (D_y f)y_1 = 0)\} \\ \overline{S} &= \{(t,x) : \exists y([(I-R)f + RD_t f](t,x,y) + [RD_x f](t,x,y)y = 0)\} \ , \end{aligned}$$

and the manifolds $M_1(t,x)$ and $\overline{M}(t,x)$ are formed by all vectors y, y_1 solving the defining equations for S_1 and $\overline{S}$. For the reduction property we have to prove $S_1 = S_0 \cap \overline{S}$ and $M_0(t,x) \cap \overline{M}(t,x) = \overline{E}_1 M_1(t,x)$. First we remark that $S_1 = \{(t,x) \ : \exists y(f(t,x,y) = 0 \text{ and } (RD_t f)(t,x,y) + (RD_x f)(t,x,y)y = 0)\}$. Assume $(t,x) \in S_0 \cap \overline{S}$ and $y \in M_0(t,x) \cap \overline{M}(t,x)$. Then

$$R(D_t f)(t,x,y) + R(D_x f)(t,x,y)y = 0 \tag{5.15}$$

implies the existence of a y_1 satisfying $(D_y f)y_1 + (D_x f)y + D_t f = 0$, i. e. $y \in \overline{E}_1 M_1(t,x)$. Conversely, $(t,x) \in S_1$ and $y \in \overline{E}_1 M_1(t,x)$ yield (5.15) due to $RD_y f = 0$, i. e. $y \in M_0(t,x) \cap \overline{M}(t,x)$. Thus, the reduction property is verified.

Theorem 16. *The transfer from the DAE (5.11) to the DAE (5.13) with the initial restriction (5.14) reduces the index by 1.*

As an example we consider the application of the reduction methods to the autonomous equations of motions of mass point systems:

$$\dot{v} = f(r) + D_r g(r)^T z, \quad \dot{r} = v, \quad 0 = g(r) \ , \tag{5.16}$$

where r, v, z denote the position, velocity and Lagrange vectors, respectively, and the constraints in $g(r) = 0$ are assumed to be independent. It is well-known that (5.16) is an index-3-problem. By the substitution $z = \dot{w}$ we obtain the quasilinear index-2-DAE (cf. [Gea88] or [Gri90]) with $G(r) := D_r g(r)$

$$\dot{v} - G(r)^T \dot{w} = f(r), \quad \dot{r} = v, \quad 0 = g(r) \ . \tag{5.17}$$

Applying one reduction step of Theorem 13 we obtain

$$\dot{v} - G(r)^T \dot{w} = f(r), \quad \dot{r} = v, \quad G(r)\dot{r} = g(r) \text{ or } G(r)\dot{r} = 0 \ . \tag{5.18}$$

The second reduction step yields the matrix $A(r) = \begin{pmatrix} I & 0 & -G^T \\ 0 & I & 0 \\ G & H & 0 \end{pmatrix}$ with $H = -D_r G v + G$ or $H = -D_r G v$. In each case $A(r)$ is nonsingular and we obtain actually an explicit ordinary differential equation.

The use of one reduction step of Theorem 15 generates

$$\dot{v} - G(r)^T \dot{w} = f(r), \quad Q(r)\dot{r} = v, \quad 0 = g(r) \ , \tag{5.19}$$

where $Q(r)$ is any projector onto $TS_0(r) = \ker(G(r))$. This is an index-1-DAE: We verify this statement applying one reduction step of Theorem 13. Then, in the arising DAE, the matrix $A(r) = \begin{pmatrix} I & 0 & -G^T \\ I-Q & Q & 0 \\ 0 & G & 0 \end{pmatrix}$ is nonsingular, because of $G(I-Q) = G$ i. e. (5.19) is really an index-1-problem.

Finally, we have to consider an important special case of (5.11) which is characterized by $\ker(D_y f(t,x,y)) = N(t)$. Then, for any projector $Q_N(t) = I - P_N(t)$ onto $N(t)$ the DAE (5.11) is equivalent to $0 = f(t,x(t),P_N(t)x'(t)) = f(t,x(t),(P_N x)'(t) - P_N'(t)x(t))$ since $f(t,x,y) - f(t,x,P_N y) = \int_0^1 D_y f(t,x,sy + (1-s)P_N y) ds Q_N y = 0$. By the substitution $u(t) := (P_N x)(t), v(t) := (Q_N x)(t)$ we obtain

$$\begin{aligned} f(t, u+v, u' - P_N' v) &= 0 \\ Q_N u + P_N v &= 0 \;, \end{aligned} \tag{5.20}$$

which is equivalent to (5.11) in our special case. Furthermore the equation $Q_N u + P_N v = 0$ is equivalent to

$$Q_N u' + Q_N' u + P_N v = 0 \tag{5.21}$$

with the initial restriction $(Q_N u)(t_0) = 0$ for some t_0. Namely, $Q_N u + P_N v = 0$ yields $Q_N u = 0$ and $P_N v = 0$, i. e. $(Q_N u)' = 0$ and therefore (5.21). On the other hand, (5.21) implies $0 = Q_N u' + Q_N Q_N' u = (Q_N u)' - Q_N' u + Q_N' P_N u = (Q_N u)' - Q_N'(Q_N u)$. This is a homogeneous linear differential equation for $Q_N u$, hence the initial restriction yields $Q_N u = 0$ and then $P_N v = 0$. Therefore we can replace (5.20) by

$$\begin{aligned} f(t, u+v, u' - P_N' v) &= 0 \\ Q_N u' + Q_N' u + P_N v &= 0 \end{aligned} \tag{5.22}$$

with the initial restriction $(Q_N u)(t_0) = 0$.

In [Gri90] we proposed an index reduction by the substitution $v(t) = w'(t)$. There we proved (but only in the linear case) the following

Theorem 17. *The transfer from (5.11) to the DAE*

$$\begin{aligned} f(t, u+w', u' - P_N' w') &= 0 \\ Q_N u' + P_N w' + Q_N' u &= 0 \end{aligned} \tag{5.23}$$

under the initial restrictions

$$(Q_N x)(t_0) = 0, w(t_0) = 0 \textit{ for some } t_0 \tag{5.24}$$

reduces the index by 1.

In many practical problems the transfer from (5.11) to (5.23), (5.24) is very natural and useful because only the derivative of the component $u(t)$ of $x(t)$ occurs in (5.23). Therefore we can extend the solution space requiring only the differentiability of $P_N x$ (cf. [GM86]).

Remark. Until now Theorem 17 has not been proved in the general nonlinear case, but it is well-known that (5.23) is an index-0 DAE if (5.11) is an index-1 DAE, because the Jacobian of (5.23) is nonsingular under these conditions (cf. [Gri90]).

References

[Die71] J. Dieudonné. *Grundzüge der Modernen Analysis.* Volume I, Deutscher Verlag der Wissenschaften, Berlin, 1971.

[Gea88] C. W. Gear. Differential-algebraic equation index transformation. *SIAM J. Sci. & Statist. Comp.*, 9:39–47, 1988.

[GM86] E. Griepentrog and R. März. *Differential-Algebraic Equations And Their Numerical Treatment. Teubner-Texte zur Mathematik; 88*, Teubner, Leipzig, 1986.

[GM89] E. Griepentrog and R. März. Basic properties of some differential-algebraic equations. *Z. für Anal. u. ihre Anwendungen*, 8:25–40, 1989.

[Gri90] E. Griepentrog. The index of differential-algebraic equations and its significiance for the circuit simulation. In R. E. Bank, R. Bulirsch, and K. Merten, editors, *Mathem. Modelling and Simulation of Electrical Circuits and Semiconductor Devices*, pages 11–26, Birkhäuser, Basel, 1990. ISNM 93.

[Rei88] S. Reich. *Differential-algebraic equations and vector fields on manifolds.* Preprint 09-02-88, Techn. Univ., Sektion Informationstechnik, Dresden, 1988.

[RR90] P.J. Rabier and W.C. Rheinboldt. *A general existence and uniqueness theorem for implicit differential–algebraic equations.* Technical Report ICMA-90-145, Univ. Pittsburgh, 1990.

[Zor81] V.A. Zorič. *Mathematical Analysis.* Volume I, Nauka, Moskva, 1981. in Russian.