Computing Consistent Initial Values for Nonlinear Index-2 Differential-Algebraic Equations

Beate Hansen

Humboldt-Universität zu Berlin, Fachbereich Mathematik,
PSF 1297, O-1086 Berlin, Germany

1 Introduction

Let the nonlinear differential-algebraic equation (DAE)

$$f(x'(t), x(t), t) = 0, \quad t \in [t_0, T], \quad x(t) \in \mathbb{R}^m \tag{1}$$

be given, where the partial Jacobian $\frac{\partial f}{\partial x'} =: f'_{x'}$ is singular on $[t_0, T]$. Treating DAEs some problems concerning initial values arise; a "certain subsystem" of (1) is known to be a consistency condition for the solution of the DAE. Namely, there is not a smooth solution x passing an arbitrarily given point $(t_0, x_0) \in [t_0, T] \times \mathbb{R}^m$ (as it uses to be in case of explicit ODEs). A "reasonable" initial value for a DAE must belong to the underlying constraints manifold, that means it has to be consistent. Inconsistent initial values (t_0, x_0) can cause numerical methods for DAEs to fail, or they can be meaningless from the physical point of view.

DAEs are commonly classified by means of an *index*. There are several definitions of this term, but in each of these index concepts explicit ODEs have the index zero. Furthermore, index-1 equations are known to be treatable essentially easier than *higher index equations* (which have an index ≥ 2). So it seems to be useful to reduce the index of a given DAE before applying numerical solution methods. Such index reduction algorithms were proposed e. g. by Gear [Gea88], Griepentrog [Gri90], and Leimkuhler [Lei88]. Most of these reduction methods differentiate the DAE or parts of it (the equations without derivatives) in order to obtain the remaining x-components' derivatives. Unfortunately, such algorithms have got some disadvantages: Analytical differentiation (using Computer Algebra) or approximation of the derivatives arising during the algorithm's execution is necessary. Furthermore, the reduction algorithm can destroy possibly favourable structure (e. g. matrix sparseness) of the DAE. Another problem is what about initial values while executing such an algorithm: The solution set of a differential-algebraic equation does not change during the index reduction only if the initial value is consistent. But consistency is not a priori to be seen. This is why we make use of another index reduction method introduced by Griepentrog

[Gri90]. Namely, in Sect. 4 we consider index-2 DAEs and realize that consistent initial values for such equations can be obtained via initial values of the corresponding index-1 system which results after Griepentrog's reduction algorithm. But before that, in Sect. 2, we outline the method of Leimkuhler [Lei88] because it seems to be the only handy initialization method for certain classes of nonlinear DAEs so far. Sect. 3 tries to calculate consistent initial values directly for the given index-2 DAE using a special variable transformation which enables us to decouple the DAE into a system of an explicit ODE and two (nonlinear) equations. Each of these three equations yields a component of the solution x. Unfortunately, this method requires symbolic calculation as well as some restrictions concerning the smoothness of $\ker f'_{x'}$ (certain subspaces of $\ker f'_{x'}$ must be continuously differentiable with respect to t, others must be constant).

2 An Initialization Method by Leimkuhler

The author considered the differential-algebraic equation

$$F(x, x', t) = 0, \quad x : [t_0, T] \to \mathrm{IR}^n, \quad F : \mathrm{IR}^{2n+1} \to \mathrm{IR}^n \tag{2}$$

and defined it to be *smoothly solvable on* $[t_0, T]$ if there exists a number $m \geq 0$ such that $F \in C^m$ and the nonlinear system

$$\begin{aligned} F(\xi_0, \xi_1, t) &= 0 \\ DF(\xi_0, \xi_1, \xi_2, t) &= 0 \\ &\vdots \\ D^m F(\xi_0, \cdots, \xi_{m+1}, t) &= 0 \end{aligned} \tag{3}$$

can be solved for ξ_1 in terms of ξ_0 and t: $\xi_1 = \Phi(\xi_0, t)$ where the map Φ is Lipschitz continuous. $DF = (\xi_0, \xi_1, \xi_2, t) := \frac{\partial F}{\partial \xi_0}\xi_1 + \frac{\partial F}{\partial \xi_1}\xi_2 + \frac{\partial F}{\partial t}$ is the *Pseudo derivative* of F, and $D^m F := D(D^{m-1}F)$. Then, the *global index* of a smoothly solvable DAE is said to be the smallest $m \geq 0$ such that the system (3) is solvable for $\xi_1 = \Phi(\xi_0, t)$.

An initial vector (x_0, x'_0) for DAE (2) is called *consistent in* t_0 if the DAE has a C^{m+1} solution passing this vector in t_0.

Leimkuhler assumed a relation $B(x_0, x'_0) = 0$ to be known by the user. The system of the $m + 2$ equations

$$\begin{aligned} &F(x_0, x'_0, t_0) = 0 \\ &D^i F(x_0, x'_0, \ldots, x_0^{i+1}, t_0) = 0 \quad (i = 1, \ldots, m) \\ &B(x_0, x'_0) = 0 \end{aligned}$$

is called *the system* of *consistency equations*. Solving this system for (x_0, x'_0) shall yield the consistent initial vector. It was pointed out that on the one hand the consistency equations can be solved for (x_0, x'_0) in principle, but actually this cannot be done. Namely, computing the derivatives $D^m F$ analytically can be very complicated and, furthermore, the consistency equations system is underdetermined. Hence the author proposed to approximate the first deriative DF by

$$D_h F := \frac{1}{h}\left[\sum_{i=1}^{s} \alpha_i F(z_0 + hc_i z_0') - \sum_{i=1}^{s} \alpha_i F(z_0)\right]$$

and the higher derivatives $D^k F$ by

$$D_h^k F := \frac{k!}{h^k}\left[\sum_{i=1}^{s} \alpha_i F(z_0 + c_i h z_0' + \cdots + \frac{(c_i h)^r}{r!} z_0^{(r)}) - \sum_{i=1}^{s} \alpha_i F(z_0)\right].$$

Then rounding errors, initial displacement h and tolerance τ have to be determined.

The next problem is to solve the (approximated) consistency equations numerically (by the least squares method). There it is advantageous to reduce the equations into a full rank system. This is possible for DAEs given in Hessenberg form. In the case of index-2 and index-3 equations in Hessenberg form the solutions of the (nonsingular) least squares problems converge. Otherwise (in the rank deficient case) there are convergent solution sequences for linear DAEs with constant coefficients as well as for nonlinear index-1 equations

$$E(x,t)x' - f(x,t) = 0$$

with $\operatorname{rank}(E(x,t)) = \text{const}$.

Unfortunately, this initialization method is not applicable to more general nonlinear index-2 differential-algebraic equations. Let us see in the next chapter what about so-called *index-2 tractable* DAEs.

3 Index-2 Tractable DAEs

Let us consider the (DAE)

$$f(x'(t), x(t), t) = 0 \tag{4}$$

where $t \in [t_0, T] \subset \mathbb{R}^1$, $x(t) \in \mathbb{R}^m$, and the partial Jacobian $f_y'(y,x,t)$ is singular everywhere on $[t_0, T]$.

Assume that the following two conditions hold (cf. [GM86], [Mar89]):

1. Let $\mathcal{G} \subset \mathbb{R}^m \times \mathbb{R}^m$ be such a region that $(y,x,t) \in \mathcal{G} \times [t_0,T]$ and $y - \bar{y} \in \ker f_y'(y,x,t)$ imply $((1-a)y + a\bar{y}, x) \in \mathcal{G}$ for all $a \in [0,1]$. Let f be continuous on $\mathcal{G} \times [t_0,T]$. The partial Jacobians $f_y'(y,x,t), f_x'(y,x,t) \in L(\mathbb{R}^m)$ exists everywhere on $\mathcal{G} \times [t_0,T]$, and the corresponding matrix functions $f_y', f_x' : \mathcal{G} \times [t_0,T] \to L(\mathbb{R}^m)$ are continuous.
2. The nullspace of $f_y'(y,x,t)$ depends on t only; let us denote it by $N(t) := \ker f_y'(y,x,t)$. There exists a function $Q \in C^1([t_0,T], L(\mathbb{R}^m))$ such that $Q(t)$ is a projector from $\mathbb{R}^m$ onto $N(t)$ for each $t \in [t_0,T]$. Denote the complementary projector by $P(t) := I - Q(t)$.

These two conditions ("Assumption (A)" in several papers, e. g. [GM86], [Lei88]), imply

$$f(y,x,t) = f(P(t)y,x,t) \text{ on } \mathcal{G}\times[t_0,T]$$

(cf. [GM86], p. 25), i. e. we can understand the DAE (1) as

$$f((Px)'(t) - P'(t)x(t), x(t), t) = 0. \tag{5}$$

Obviously, the function space

$$\mathcal{C} := \{x \in C([t_0,T]); Px \in C^1([t_0,T])\}$$

is an adequately defined space solutions of (1) resp. (2) should belong to. Denote $\mathcal{D} := \{x \in \mathcal{C}; Tr(x) \subset \mathcal{G}\}$ where $Tr(x) := \{((Px)'(t) - P'(t)x(t), x(t), t)t \in [t_0,T]\}$ is the trajectory of x.

A DAE (1) is called *transferable (or tractable with index 1)* locally around $x_* \in \mathcal{D}$ if the linearized in x_* equation is a transferable DAE. In this case we refer to [GM86]. This paper deals with so-called index-2 tractable equations, and, especially, with consistent initial values for such equations.

Definition 1 [Mar89]. Let Assumption (A) be true. The differential-algebraic equation (1) is called *tractable with index 2* locally at $x_* \in \mathcal{D}$ if the linearized in x_* equation is an index-2-tractable DAE.

Theorem 2 ([Mar89], p. 165). *Let (A) be true, $x_* \in \mathcal{D}$. Let the matrix $G_1(y,x,t) := f_y'(y,x,t) + (f_x'(y,x,t) - f_y'(y,x,t)P'(t))Q(t)$ be singular at $Tr(x_*)$. Let $Q_1(y,x,t) \in L(\mathbb{R}^m)$ denote a projection onto* ker $G_1(y,x,t)$*, and let $Q_1(y,x,t)$ depend continuously on (y,x,t) at $Tr(x_0)$. Then the DAE (1) is tractable with index 2 locally at x_* iff the matrix $G_2(y,x,t) := G_1(y,x,t) + f_x'(y,x,t)P)(t)Q_1(y,x,t)$ remains nonsingular at x_*.*

Remark that for special nonlinear DAEs (namely quasi-linear equations and semi-explicit DAEs) index-2 tractability is essentially equivalent to the index-2 definition of various authors (e. g. [GP84]), cf. [Mar89].

An important problem within the investigation of DAEs is how to obtain consistent initial values. We want to deal with this problem in what follows.

Definition 3. A vector $x_0 \in \mathbb{R}^m$ is said to be a *consistent initial value* for the DAE (1) if there is a solution x of this equation with $x(t_0) = x_0$.

We know that determining a consistent initial value is not trivial (as it is a matter of fact for nonsingular ODEs) because x_0 must fulfill, in particular, the "algebraic constraints" within the DAE.

Recall that for index-1 DAEs there is a method which transfers the given equation into its state variable system (cf. [GM86]). This method makes use of the transformation matrix $\binom{P\ Q}{Q\ P}$ which is equal to its inverse. This shall be generalized for index-2 equations in the following.

In order to decouple the given DAE (1) (or (2), respectively) we introduce a choice of variables by the following transformation:

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} := \begin{pmatrix} PP_1 & PQ_1 & Q \\ PQ_1 & Q & PP_1 \\ Q & PP_1 & PQ_1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{6}$$

where $y := (PP_1x)' - (PP_1)'x$ and $z := PQ_1x$. We assume that all submatrices depend only on t and are continuously differentiable with respect to this argument. Let $P_1(t)$ be the complementary projector of $Q_1(t) : P_1(t) := I - Q_1(t)$, and let the following assumption be true:

$$Q_1Q = 0 \quad \text{on } \mathcal{G} \times [t_0, T] \tag{7}$$

Then the projector products PP_1 and PQ_1 can be easily verified to be projectors, too. Furthermore, the relations

$$\begin{aligned} PP_1PQ_1 &= 0 \\ PQ_1Q &= 0, \text{ and} \\ PQ_1PP_1 &= 0 \end{aligned}$$

hold (because of (7)), and we remark that the transformation matrix is nonsingular (for each $t \in [t_0, T]$) and equal to its inverse.

Theorem 4. *Let the DAE (2) be tractable with index 2, and let $x \in \mathcal{C}$ be a solution of (2) satisfying the initial condition*

$$(PP_1)(t_0) \cdot (x(t_0) - x^0) = 0 \ , \tag{8}$$

where x^0 is an arbitrary vector in $\mathbb{R}^m$. Then for the components

$$\begin{aligned} u &= PP_1x \\ v &= PQ_1x, \quad \text{and} \\ Qw &= Qx \end{aligned}$$

the following equations hold:

$$u' - (PP_1)'u = (PP_1)`(v + Qw) + PP_1w \tag{9.1}$$

$$v = PQ_1w \tag{9.2}$$

$$w \text{ isthe(unique)solutionofthenonlinearequation}$$

$$f(w - PQ_1w + (PP_1)'(u + v + Qw) + v' - P'Qw, u + v + Qw, t) = 0 \tag{9.3}$$

Conversely, if these three relations (6) are true then $u = PP_1u$ as well as $v = PQ_1v$ hold, and $x = u + v + Qw$ is a solution of the initial value problem (2), (5).

Hence, computing a consistent initial value $x(t_0) = u(t_0) + v(t_0) + Q(t_0)w(t_0)$ we have to solve the nonlinear equation

$$
\begin{aligned}
&\tilde{f}(w, u, v, v', t) := \\
&f\left((PP_1)'(u+v+Qw)+PP_1 w + v' - P'Qw,\, u+PQ_1 w+Qw,\, t\right)=0 \qquad (10)
\end{aligned}
$$

with respect to w. The partial Jacobian $\frac{\partial \tilde{f}}{\partial w}$ is a nonsingular matrix since the DAE (2) is tractable with index 2 (and adding the term $-f'_y PQ_1 + f'_y (PP_1)'Q$ to the matrix G_2 does not change the nonsingularity of this matrix). That means, by the implicit function theorem, (10) is uniquely solvable with respect to w.

Assume we have solved it (e. g. by damped Newton's method using a formula manipulator system) and got a function $w = w(u, v, v', t)$ as a solution. What about computing the component $v = PQ_1 w$, while w depends on v and v'?

Fortunately, one can easily show that there is no problem because $\frac{\partial w}{\partial v}$ as well as $\frac{\partial w}{\partial v'}$ do not belong to the subspace $\mathrm{im}(PQ_1)(t)$:

Case 1. We consider $\tilde{f}$ depending on w, u, v, v' and t as above.

$$
\begin{aligned}
\frac{\partial \tilde{f}}{\partial v} = \frac{\partial \tilde{f}}{\partial w} \cdot \frac{\partial w}{\partial v} \Longrightarrow \frac{\partial w}{\partial v} &= \left(\frac{\partial \tilde{f}}{\partial w}\right)^{-1} \frac{\partial \tilde{f}}{\partial v} \\
&= \left(G_2 - f'_y PQ_1 + f'_y (PP_1)'Q\right)^{-1} f'_y (PP_1)' = \\
&= \left(I - QQ_1 - (PP_1)'QP_1\right) G_2^{-1} f'_y (PP_1)' = \\
&= P_1 P (PP_1)' + (PP_1)'QQ_1 (PP_1)' \\
&\Longrightarrow Q_1 \frac{\partial w}{\partial v} = 0 \ .
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \tilde{f}}{\partial v'} = \frac{\partial \tilde{f}}{\partial w} \cdot \frac{\partial w}{\partial v'} \Longrightarrow \frac{\partial w}{\partial v'} &= \left(\frac{\partial \tilde{f}}{\partial w}\right)^{-1} \frac{\partial \tilde{f}}{\partial v'} \\
&= \left(I - QQ_1 - (PP_1)'QP_1\right) G_2^{-1} f'_y = \\
&= P_1 P + (PP_1)'QQ_1 \\
&\Longrightarrow Q_1 \frac{\partial w}{\partial v'} = 0.
\end{aligned}
$$

Case 2. Now we insert the term $PQ_1 w$ (instead of v) into the arguments of f:

$$
\tilde{f}(w, u, v', t) := f\left((PP_1)'(u + PQ_1 w + Qw) + PP_1 w + v' - P'Qw,\, u + PQ_1 w + Qw,\, t\right)
$$

$$
\frac{\partial \tilde{f}}{\partial v} = \frac{\partial \tilde{f}}{\partial w} \cdot \frac{\partial w}{\partial v} \Longrightarrow \frac{\partial w}{\partial v} = \left(\frac{\partial \tilde{f}}{\partial w}\right)^{-1} \cdot \frac{\partial \tilde{f}}{\partial v} = 0 \text{ because } \tilde{f} \text{ does not depend on } v.
$$

$$\begin{aligned}
\frac{\partial \tilde{f}}{\partial v'} = \frac{\partial \tilde{f}}{\partial w} \cdot \frac{\partial w}{\partial v'} &\Longrightarrow \frac{\partial w}{\partial v'} = \left(\frac{\partial \tilde{f}}{\partial w}\right)^{-1} \frac{\partial \tilde{f}}{\partial v'} \\
&= \left(G_2 - f'_y PQ_1 + f'_y (PP_1)' PQ_1 + f'_y (PP_1)' Q\right) f'_y = \\
&= \left(I - QQ_1 - (PP_1)' PQ_1 - (PP_1)' QP_1\right) G_2^{-1} f'_y = \\
&= P_1 P + (PP_1)' QQ_1 \\
&\Longrightarrow Q_1 \frac{\partial w}{\partial v'} = 0 \ .
\end{aligned}$$

Remark that the nonsingularity of $\frac{\partial \tilde{f}}{\partial w}$, too, does not depend on the choice of v or $PQ_1 w$ within the first argument of the function f.

Applying Newton's method in the points t_0 and $t_0 + h$ (h sufficiently small) we can get an approximate solution for $w(t_0)$ and $w(t_0 + h)$, respectively. Then we compute (approximately) $v(t_0)$ and $v(t_0 + h)$ by premultiplying the approximations of $w(t_0)$ and $w(t_0 + h)$ by the corresponding projector $(PQ_1)(t_0)$ and $(PQ_1)(t_0 + h)$, respectively. Inserting the difference quotient

$$(v(t_0 + h) - v(t_0))\, h^{-1}$$

(approximating $v'(t_0)$) into w yields the complete function w.

Finally, we realize by Theorem 2 that the initial value for the remaining component $u = PP_1 u$ can be chosen arbitrarily in the subspace $\operatorname{im}(PP_1)(t_0)$: $u(t_0) \in \operatorname{im}(PP_1)(t_0)$. But we have to assume that $(PP_1)' PQ_1 = 0$ holds in order to avoid a v-component or a v'-component to occur in the right-hand side of the explicit ODE (9.1).

Obviously, the initial value $x^0 = u(t_0) + v(t_0) + Q(t_0) w(t_0)$ is consistent to the DAE (2) by definition, and its component $u(t_0)$ satisfies the initial condition (8).

Example 1. Let us consider the Hessenberg form DAE

$$f(x'(t), x(t), t) = \begin{pmatrix} x_1'(t) - x_1^2(t) - x_2(t) - t^2 \\ -x_1(t) - t^3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

which has the solution $\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} -t^3 \\ -t^6 - 4t^2 \end{pmatrix}$. Calculate the partial Jacobians $f'_y = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $f'_x = \begin{pmatrix} -2x_1 & -1 \\ -1 & 0 \end{pmatrix}$ as well as the matrix $G_1 = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$. Obviously, $Q_1 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$ is a projector onto $\ker G_1$, namely the canonical one.

Now we are able to compute the next matrix G_2 within the matrix chain and to verify that it is nonsingular on $\mathbb{R}^2 \times [t_0, T]$:

$$G_2 = \begin{pmatrix} 1 - 2x_1 & -1 \\ -1 & 0 \end{pmatrix} .$$

Furthermore, we realize that

$$PP_1 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad PQ_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad QP_1 = \begin{pmatrix} 0 & 0 \\ -1 & 1 \end{pmatrix}, \quad QQ_1 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Remark that the projector functions Q_1 and P_1 do not depend on (y, x) and are (trivially) continuously differentiable with respect to t, i. e. the additional suppositions in Assumption (A) are satisfied.

Now compute the x-components

$$u = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad v = \begin{pmatrix} x_1 \\ 0 \end{pmatrix} \text{ and } w = \begin{pmatrix} 0 \\ x_2 \end{pmatrix} + \begin{pmatrix} x_1 \\ 0 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

* Let us use Newton's method for solving the nonlinear equation

$$f(PP_1 w + (PP_1)'(u + v + Qw) + v' - P'Qw, u + v + Qw, t) = 0 \ .$$

This can be done by the formula manipulator system MAPLE which yields the following iteration results. The iteration matrix is $\frac{\partial f}{\partial w} = \begin{pmatrix} -2w_1 & -1 \\ -1 & 0 \end{pmatrix}$ which is obviously nonsingular on $\mathbb{R}^2 \times [t_0, T]$. Using the starting point $w^{(0)} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ we carry out Newton's method and obtain after two steps of iteration the solutions

$$w^{(2)} = \begin{pmatrix} -t_0^3 \\ y_1' - t_0^6 - t_0^2 \end{pmatrix}.$$

Let us denote $w^{(2)} = w^*$.

Analogously, solving the system

$$\tilde{f}(w, u, v, v', t_0 + h) = 0$$

(choosing the same starting point $w^{(0)}$) we realize

$$w_h^{(2)} = \begin{pmatrix} -(t_0 + h)^3 \\ v_1' - (t_0 + h)^6 - (t_0 + h)^2 \end{pmatrix} =: w_h^*$$

to be a solution of it.

* Now compute (approximately) the vector functions

$$v = (PQ_1)(t_0)w^* = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} -t_0^3 \\ v_1' - t_0^6 - t_0^2 \end{pmatrix} = \begin{pmatrix} -t_0^3 \\ 0 \end{pmatrix} \quad \text{and}$$

$$\begin{aligned} v_h &= (PQ_1)(t_0 + h)w_h^* = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} -(t_0 + h)^3 \\ v_1' - (t_0 + h)^6 - (t_0 + h)^2 \end{pmatrix} \\ &= \begin{pmatrix} -(t_0 + h)^3 \\ 0 \end{pmatrix} \end{aligned}$$

and approximate

$$\frac{v_h - v}{h} = \frac{1}{h} \cdot \begin{pmatrix} -3t_0^2 h - 3t_0 h^2 - h^3 \\ 0 \end{pmatrix} \underset{h \to 0}{\longrightarrow} \begin{pmatrix} -3t_0^2 \\ 0 \end{pmatrix}.$$

* In this way we obtain the complete solution function w evaluated at t_0:

$$w(t_0) = \begin{pmatrix} -t_0^3 \\ -3t_0^2 - t_0^6 - t_0^2 \end{pmatrix} = \begin{pmatrix} -t_0^3 \\ -t_0^6 - 4t_0^2 \end{pmatrix}.$$

* Finally, the consistent initial value can be composed:

$$x(t_0) = u(t_0) + v(t_0) + Q(t_0)w(t_0) = \begin{pmatrix} -t_0^3 \\ -t_0^6 - 4t_0^2 \end{pmatrix}.$$

Note that performing Newton's method requires to calculate with real numbers as well as with symbols (namely v_1'). That means, we are enforced to use a computer algebra system in order to solve the nonlinear equation $\tilde{f}(w, u, v, v', t) = 0$ with respect to w which is somewhat unsatisfactory from the numerical point of view. This is why we propose another possibility for computing consistent initial values for index-2 tractable differential-algebraic equations in the next section.

4 Index Reduction

Griepentrog [Gri90] introduced a method for index reduction by means of adding certain equations to the original DAE. This method for linear time- varying differential-algebraic equations is easily applicable to a certain class of nonlinear index-2 DAEs, too. We want to use it for determining consistent initial values for such equations. We denote the x-components by $\mathbf{u} := Px$ and $\mathbf{w}' := Qx$, respectively. Moreover, we introduce the denotations $\mathbf{x} := \begin{pmatrix} \mathbf{u} \\ \mathbf{w} \end{pmatrix} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}$, $\mathbf{y} := \begin{pmatrix} \mathbf{u}' \\ \mathbf{w}' \end{pmatrix} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} \in \mathbb{R}^{2m}$. Then a new differential-algebraic equation is given by

$$\mathbf{f}(\mathbf{x}', \mathbf{x}, t) = 0 \tag{11}$$

where

$$\mathbf{f}(\mathbf{y}, \mathbf{x}, t) := \begin{pmatrix} f(\mathbf{y}_1 - P'\mathbf{y}_2, \mathbf{x}_1 + \mathbf{y}_2, t) \\ Q\mathbf{y}_1 - P\mathbf{y}_2 - P'\mathbf{x}_1 \end{pmatrix}.$$

In order to verify that this DAE is tractable with the index one we calculate the corresponding matrices of the matrix chain (e. g. [Mar89]):

$$\frac{\partial \mathbf{f}}{\partial \mathbf{y}} := \begin{pmatrix} f_y' & f_x' - f_y P' \\ Q & P \end{pmatrix}, \quad \frac{\partial \mathbf{f}}{\partial \mathbf{x}} := \begin{pmatrix} f_x' & 0 \\ -P' & 0 \end{pmatrix}.$$

Obviously, a projector onto $\ker \frac{\partial \mathbf{f}}{\partial \mathbf{y}}$ is given by the matrix $\mathbf{Q} = \begin{pmatrix} PQ_1 & 0 \\ QQ_1 & 0 \end{pmatrix}$ where Q_1 denotes a projector onto $\ker G_1$ with $G_1 := f_y' + (f_x' - f_y' P')$.

Now we compute the matrix

$$\mathbf{G}_1 := \frac{\partial \mathbf{f}}{\partial \mathbf{y}} + \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \cdot \mathbf{Q} = \begin{pmatrix} f_y' + f_x' PQ_1 & f_x' - f_y' P' \\ Q - P'PQ_1 & -P \end{pmatrix}$$

and realize that it is nonsingular on $[t_0, T]$ because the matrix $G_2 := G_1 + f_y' P Q_1$ is so (by assumption of index-2 tractability).

Let us verify whether the assumption (A) is satisfied for the DAE (8):

1. Let the required property for the region $\mathcal{G} \in \mathbb{R}^{2m} \times \mathbb{R}^{2m}$ hold.
   * Obviously, $\mathbf{f} \in C(\mathcal{G} \times [t_0, T], \mathbb{R}^{2m})$.
   * The Jacobians $\mathbf{f_y}'(\mathbf{y}, \mathbf{x}, t)$ and $\mathbf{f_x}'(\mathbf{y}, \mathbf{x}, t) \in L(\mathbb{R}^{2m})$ exist and are continuous with respect to all arguments since the Assumption (A) holds for the original DAE (1).

2. In order to satisfy the second part of assumption (A) we have to make some additional suppositions:
   * The nullspace $\ker \mathbf{f_y}'(\mathbf{y}, \mathbf{x}, t) =: \mathbf{R}(t)$ is independent of $(\mathbf{y}, \mathbf{x})$ if and only if $PQ_1 = (PQ_1)(t)$ and $QQ_1 = (QQ_1)(t)$ hold.
   * $\mathbf{R}$ is a smooth nullspace, i. e. there exists a $\mathbf{Q} \in C^1([t_0, T], L(\mathbb{R}^{2m}))$ with $\mathbf{Q}(t)$ a projector onto $\mathbf{R}(t)$ for all $t \in [t_0, T]$ if and only if $PQ_1 \in C^1([t_0, T], L(\mathbb{R}^m))$ and $QQ_1 \in C^1([t_0, T], L(\mathbb{R}^m))$.

   (That means we have to demand that the projector-function Q_1 (and consequently, the complementary projector P_1, too) depends on t and is continuously differentiable with respect to this argument. Remark that this assumption is the same one made in Chap. 3, p. 5.).

Again, Assumption (A) implies that

$$\mathbf{f}(\mathbf{y}, \mathbf{x}, t) = \mathbf{f}(\mathbf{P}\mathbf{y}, \mathbf{x}, t)$$

holds for all $t \in [t_0, T]$ and for all $(\mathbf{y}, \mathbf{x}) \in \mathcal{G}$. In this way we formulate the DAE (8) as

$$\mathbf{f}(\mathbf{x}', \mathbf{x}, t) = \mathbf{f}((\mathbf{P}\mathbf{x})' - \mathbf{P}'\mathbf{x}, \mathbf{x}, t) = 0 \tag{12}$$

and define the function space

$$\mathcal{C} := \{\mathbf{x} \in C([t_0, T], \mathbb{R}^{2m}) | \mathbf{P}\mathbf{x} \in C^1([t_0, T], \mathbb{R}^{2m})\}$$

as we have done in the case of DAE (1).

Thus, Griepentrog's method yields a differential-algebraic equation (11) (or (12), respectively) being tractable with the index 1 and satisfying Assumption (A). Consequently, we can apply the following theorem (cf. [GM86]).

Theorem 5. *Let equation (11) be a transferable (i. e. index-1 tractable) DAE, let Assumption (A) be satisfied. Then the system*

$$\mathbf{F}(\mathbf{y}_0, \mathbf{x}_0, t_0) := \begin{pmatrix} \mathbf{f}(\mathbf{y}_0, \mathbf{x}_0, t_0) \\ \mathbf{Q}(t_0)\mathbf{y}_0 + \mathbf{P}(t_0)(\mathbf{x}_0 - \mathbf{x}^0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{13}$$

has exactly one solution $(\mathbf{y}_0, \mathbf{x}_0) \in \mathbb{R}^{4m}$ and $\mathbf{x}_0$ is a consistent initial value for the DAE (8). There $\mathbf{x}^0 \in \mathbb{R}^{2m}$ is an arbitrary vector.

The following section investigates how consistent initial values for the given DAE (1) can be computed using this theorem.

First we introduce the denotation $\mathbf{z} := (\mathbf{y}, \mathbf{x})$ and calculate the Jacobian $\frac{\partial \mathbf{f}}{\partial \mathbf{z}}$: Since we have already determined the matrices $\mathbf{f}'_{\mathbf{y}}, \mathbf{f}'_{\mathbf{x}}$ and $\mathbf{Q}$ it is easy to realize that $\mathbf{P} = \begin{pmatrix} I - PQ_1 & 0 \\ -QQ_1 & I \end{pmatrix}$ and, therefore, we obtain

$$\frac{\partial \mathbf{f}}{\partial \mathbf{z}} = \begin{pmatrix} f'_y & f'_x - f_y P' & f'_x & 0 \\ Q & -P & -P' & 0 \\ PQ_1 & 0 & I - PQ_1 & 0 \\ QQ_1 & 0 & -QQ_1 & I \end{pmatrix} .$$

Using this Jacobian matrix within Newton's method requires it to be nonsingular on $\mathcal{G} \times [t_0, T]$. We verify that this is a matter of fact. Let $\frac{\partial \mathbf{f}}{\partial \mathbf{z}} \cdot (a, b, c, d)^T = 0$. This yields a $4m \times 4m$-system of equations. Premultiplying the second of them by P and Q, respectively, we get

$$Pb + PP'c = 0 \quad \text{and} \quad Qa - QP'c = 0 \ .$$

The third equation multiplied by the projectors PQ_1 and $(I - PQ_1)$ yields

$$PQ_1 a = (I - PQ_1)c = 0 \ .$$

So we realize $c = PQ_1 c$ and $b = Qb$. Moreover, $a = Qa + PP_1 a$ holds, which means that $Q_1 a = 0$. Finally, the last equation of the system provides $d = QQ_1 c$.

Now we add the term $(f'_x - f'_y P')Qa + f'_x PQ_1 a$ to both sides of the first equation and premultiply the new one by the inverse of the (nonsingular) matrix G_2. In this way we obtain

$$a = -b - Q_1 c + Qa \ .$$

Multiplying this equation by Q_1 yields

$$0 = -Q_1 b - Q_1 c, \quad \text{i. e.} \quad Q_1 c = 0.$$

Hence $c = 0$ and therefore $b = -QQ_1 c = 0$.

$$Pa = -c = 0, \quad \text{i. e.} \quad a = 0$$

and $d = QQ_1 c = 0$.

Before investigating feasibility and behaviour of Newton's method we recall that we can insert

$$(\mathbf{Px})' - \mathbf{P}'\mathbf{x} = \begin{pmatrix} (I - PQ_1)\mathbf{u}' \\ -QQ_1 \mathbf{u}' + \mathbf{w}' \end{pmatrix}$$

instead of $\mathbf{y} = \mathbf{x}'$ into the first argument of $\mathbf{f}$. This implies that $PQ_1 \mathbf{u}' = -QQ_1 \mathbf{u}' = 0$, i. e. $Q_1 \mathbf{u}' = 0$ holds for the first argument of $\mathbf{f}$. Denote $z := (\mathbf{y}_1 - P'\mathbf{y}_2, \mathbf{x}_1 + \mathbf{y}_2)$.

Now consider Newton's method

$$\frac{\partial \mathbf{f}}{\partial \mathbf{z}}(\mathbf{z}^{(l)})\mathbf{a}^{(l)} = -\mathbf{F}(\mathbf{z}^{(l)}, t_0), \quad \mathbf{z}^{(l+1)} = \mathbf{a}^{(l)} + \mathbf{z}^{(l)} \quad (l = 0, 1, 2, \ldots) \tag{14}$$

with $\mathbf{a}^l = (\mathbf{a}_1^l, \mathbf{a}_2^l, \mathbf{a}_3^l, \mathbf{a}_4^l)^T \in \mathbb{R}^{4m}$. Let the starting point be $\mathbf{z}^{(0)} = \begin{pmatrix} \mathbf{u}_0' \\ \mathbf{w}_0' \\ \mathbf{u}_0 \\ \mathbf{w}_0 \end{pmatrix}$.

Obviously, the matrix $\frac{\partial \mathbf{f}}{\partial \mathbf{z}}(\mathbf{z}^{(l)})$ depends on the iteration index l. But because of our additional supposition (in Assumption (A)) that P_1 and Q_1 depend on t only we remark that the only submatrices of $\mathbf{f_z}'(\mathbf{z}^l)$ depending on the iteration index are f_y', $f_x' - f_y'P'$ and f_x'. Denote them by $[f_y']^l$, $[f_x' - f_y'P]^l$ and $[f_x']^l$, respectively.

Let us consider the first Newton step, i. e. let l be equal to 0. We have to solve the system

$$[f_y']^0\mathbf{a}_1^{(0)} + [f_x' - f_y'P']^0\mathbf{a}_2^{(0)} + [f_x']^0\mathbf{a}_3^{(0)} = -f(\mathbf{y}_1^{(0)} - P'\mathbf{y}_2^{(0)}, \mathbf{x}_1^{(0)} + \mathbf{y}_2^{(0)}, t_0) =$$
$$-f(z^{(0)}, t_0)$$
$$Q\mathbf{a}_1^{(0)} - P\mathbf{a}_2^{(0)} - P'\mathbf{a}_3^{(0)} = 0$$
$$PQ_1\mathbf{a}_1^{(0)} + (I - PQ_1)\mathbf{a}_3^{(0)} = -(I - PQ_1)(\mathbf{u}_0 - \mathbf{u}^0) = -PP_1(\mathbf{u}_0 - \mathbf{u}^0) = 0$$
$$QQ_1\mathbf{a}_1^{(0)} - QQ_1\mathbf{a}_3^{(0)} + \mathbf{a}_4^{(0)} = QQ_1(\mathbf{u}_0 - \mathbf{u}^0) - (\mathbf{w}_0 - \mathbf{w}^0)$$

(all submatrices evaluated at t_0).

Remark. The solution component $\tilde{x}^0 = (PP_1)(t_0)x^0 = (PP_1x)(t_0)$ must satisfy the relation $(PP_1)(t_0)(\mathbf{u}(t) - \mathbf{u}^0) = 0$ for all $t \in [t_0, T]$. Hence, choose the starting point $\mathbf{z}^0 = (\mathbf{y}^0, \mathbf{x}^0)$ in such a way that

$$(PP_1)(t_0)(\mathbf{u}_0 - \mathbf{u}^0) = 0$$

holds.

The system above is uniquely solvable, and we get

$$\mathbf{a}^{(0)} = \begin{pmatrix} (-PP_1 - P'PQ_1)[G_2^{-1}]^0 f(\mathbf{u}_0' - P'\mathbf{w}_0', \mathbf{u}_0 + \mathbf{w}_0', t_0) \\ -QP_1[G_2^{-1}]^0 f(\mathbf{u}_0' - P'\mathbf{w}_0', \mathbf{u}_0 + \mathbf{w}_0', t_0) \\ -PQ_1[G_2^{-1}]^0 f(\mathbf{u}_0' - P'\mathbf{w}_0', \mathbf{u}_0 + \mathbf{w}_0', t_0) \\ -QQ_1[G_2^{-1}]^0 f(\mathbf{u}_0' - P'\mathbf{w}_0', \mathbf{u}_0 + \mathbf{w}_0', t_0) + QQ_1(\mathbf{u}_0 - \mathbf{u}^0) - (\mathbf{w}_0 - \mathbf{w}^0) \end{pmatrix}$$

$$=: \begin{pmatrix} (-PP_1 - P'PQ_1)[G_2^{-1}]^0 f(z^{(0)}, t_0) \\ -QP_1[G_2^{-1}]^0 f(z^{(0)}, t_0) \\ -PQ_1[G_2^{-1}]^0 f(z^{(0)}, t_0) \\ -QQ_1[G_2^{-1}]^0 f(z^{(0)}, t_0) + QQ_1(\mathbf{u}_0 - \mathbf{u}^0) - (\mathbf{w}_0 - \mathbf{w}^0) \end{pmatrix}$$

as its solution vector. According to (11) we compute

$$\mathbf{z}^{(1)} = \mathbf{a}^{(0)} + \mathbf{z}^{(0)} =$$

$$\begin{pmatrix} (-PP_1 - P'PQ_1)[G_2^{-1}]^0 f(z^{(0)}, t_0) + \mathbf{u}_0' \\ -QP_1[G_2^{-1}]^0 f(z^{(0)}, t_0) + \mathbf{w}_0' \\ -PQ_1[G_2^{-1}]^0 f(z^{(0)}, t_0) + \mathbf{u}_0 \\ -QQ_1[G_2^{-1}]^0 f(z^{(0)}, t_0) + QQ_1(\mathbf{u}_0 - \mathbf{u}^0) - (\mathbf{w}_0 - \mathbf{w}^0) + \mathbf{w}_0 \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{y}_1^{(1)} \\ \mathbf{y}_2^{(1)} \\ \mathbf{x}_1^{(1)} \\ \mathbf{x}_2^{(1)} \end{pmatrix}$$

which completes the first step of Newton's method.

$l = 1$: First we have to compute the modified matrices $[f'_y]^1$, $[f'_x - f'_y P']^1$ and $[f'_x]^1$ according to the new imputs $(\mathbf{y}^{(1)}, \mathbf{x}^{(1)})$. Then we perform the second Newton step: $\frac{\partial \mathbf{F}}{\partial \mathbf{z}}(\mathbf{z}^{(1)})\mathbf{a}^{(1)} = -\mathbf{F}(\mathbf{z}^{(1)}, t_0)$, i. e.

$$\begin{pmatrix} [f'_y]^1 & [f'_x - f'_y P']^1 & [f'_x]^1 & 0 \\ Q & -P & -P' & 0 \\ PQ_1 & 0 & I - PQ_1 & 0 \\ QQ_1 & 0 & -QQ_1 & I \end{pmatrix} \begin{pmatrix} \mathbf{a}_1^{(1)} \\ \mathbf{a}_2^{(1)} \\ \mathbf{a}_3^{(1)} \\ \mathbf{a}_4^{(1)} \end{pmatrix} =$$

$$\begin{pmatrix} -f(\mathbf{y}_1^{(1)} - P'\mathbf{y}_2^{(2)}, \mathbf{x}_1^{(1)} + \mathbf{y}_2^{(1)}, t_0) \\ -Q\mathbf{y}_1^{(1)} + P\mathbf{y}_2^{(1)} + P'\mathbf{x}_1^{(1)} \\ -PQ_1\mathbf{y}_1^{(1)} - (I - PQ_1)(\mathbf{x}_1^{(1)} - \mathbf{x}_1^{0(1)}) \\ -QQ_1\mathbf{y}_1^{(1)} + QQ_1(\mathbf{x}_1^{(1)} - \mathbf{x}_1^{0(1)}) - (\mathbf{x}_2^{(1)} - \mathbf{x}_2^{0(1)}) \end{pmatrix}$$

Using some relations between the projectors and its derivatives we realize the right-hand side to be equal to

$$\begin{pmatrix} -f(z^{(1)}, t_0) \\ 0 \\ -PP_1(\mathbf{u}_0 - \mathbf{u}^0) \\ -\mathbf{w}^0 \end{pmatrix} = \begin{pmatrix} -f(z^{(1)}, t_0) \\ 0 \\ 0 \\ -\mathbf{w}^0 \end{pmatrix}$$

where $z^{(1)} := (\mathbf{y}_1^{(1)} - P'\mathbf{y}_2^{(1)}, \mathbf{x}_1^{(1)} + \mathbf{y}_2^{(1)})$. The (unique) solution of this system of equations is $\mathbf{a}^{(1)}$:

$$\mathbf{a}^{(1)} = \begin{pmatrix} (-PP_1 - P'PQ_1)[G_2^{-1}]^1 f(z^{(1)}, t_0) \\ -QP_1[G_2^{-1}]^1 f(z^{(1)}, t_0) \\ -PQ_1[G_2^{-1}]^1 f(z^{(1)}, t_0) \\ -\mathbf{w}^0 - QQ_1[G_2^{-1}]^1 f(z^{(1)}, t_0) \end{pmatrix}.$$

Further we calculate the second iteration vector

$$\begin{aligned} \mathbf{z}^{(2)} &= \mathbf{a}^{(1)} + \mathbf{z}^{(1)} \\ &= \begin{pmatrix} \mathbf{u}'_0 - (PP_1 + P'PQ_1)([G_2^{-1}]^0 f(z^{(0)}, t_0) + [G_2^{-1}]^1 f(z^{(1)}, t_0)) \\ \mathbf{w}'_0 - QP_1([G_2^{-1}]^0 f(z^{(0)}, t_0) + [G_2^{-1}]^1 f(z^{(1)}, t_0)) \\ \mathbf{u}_0 - PQ_1([G_2^{-1}]^0 f(z^{(0)}, t_0) + [G_2^{-1}]^1 f(z^{(1)}, t_0)) \\ -QQ_1([G_2^{-1}]^0 f(z^{(0)}, t_0) + [G_2^{-1}]^1 f(z^{(1)}, t_0)) + QQ_1(\mathbf{u}_0 - \mathbf{u}^0) \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{y}_1^{(2)} \\ \mathbf{y}_2^{(2)} \\ \mathbf{x}_1^{(2)} \\ \mathbf{x}_2^{(2)} \end{pmatrix}. \end{aligned}$$

Analogously, we have to perform the next step ($l=2$), i. e. to solve the system

$$\begin{pmatrix} [f_y']^2 & [f_x'-f_y'P']^2 & [f_x']^2 & 0 \\ Q & -P & -P' & 0 \\ PQ_1 & 0 & I-PQ_1 & 0 \\ QQ_1 & 0 & -QQ_1 & I \end{pmatrix} \mathbf{a}^{(2)} = \begin{pmatrix} -f(z^{(2)},t_0) \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Its solution is the vector function

$$\mathbf{a}^{(2)} = \begin{pmatrix} -(PP_1+P'PQ_1)[G_2^{-1}]^2 f(z^{(2)},t_0) \\ -QP_1[G_2^{-1}]^2 f(z^{(2)},t_0) \\ -PQ_1[G_2^{-1}]^2 f(z^{(2)},t_0) \\ -QQ_1[G_2^{-1}]^2 f(z^{(2)},t_0) \end{pmatrix}.$$

This yields the third iteration value

$$\mathbf{z}^{(3)} = \mathbf{a}^{(2)} + \mathbf{z}^{(2)}$$

$$= \begin{pmatrix} \mathbf{u}_0' - (PP_1+P'PQ_1)([G_2^{-1}]^0 f(z^{(0)},t_0) + [G_2^{-1}]^1 f(z^{(1)},t_0) + [G_2^{-1}]^2 f(z^{(2)},t_0)) \\ \mathbf{w}_0' - QP_1([G_2^{-1}]^0 f(z^{(0)},t_0) + [G_2^{-1}]^1 f(z^{(1)},t_0) + [G_2^{-1}]^2 f(z^{(2)},t_0)) \\ \mathbf{u}_0 - PQ_1([G_2^{-1}]^0 f(z^{(0)},t_0) + [G_2^{-1}]^1 f(z^{(1)},t_0) + [G_2^{-1}]^2 f(z^{(2)},t_0)) \\ QQ_1(\mathbf{u}_0 - \mathbf{u}^0) - QQ_1([G_2^{-1}]^0 f(z^{(0)},t_0) + \ldots + [G_2^{-1}]^2 f(z^{(2)},t_0)) \end{pmatrix}.$$

Clearly, the lth step must be:

$$\begin{pmatrix} [f_y']^l & [f_x'-f_y'P']^l & [f_x']^l & 0 \\ Q & -P & -P' & 0 \\ PQ_1 & 0 & I-PQ_1 & 0 \\ QQ_1 & 0 & -QQ_1 & I \end{pmatrix} \mathbf{a}^{(l)} = \begin{pmatrix} -f(z^{(l)},t_0) \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Longrightarrow \mathbf{a}^{(l)} = \begin{pmatrix} -(PP_1+P'PQ_1)[G_2^{-1}]^l f(z^{(l)},t_0) \\ -QP_1[G_2^{-1}]^l f(z^{(l)},t_0) \\ -PQ_1[G_2^{-1}]^l f(z^{(l)},t_0) \\ -QQ_1[G_2^{-1}]^l f(z^{(l)},t_0) \end{pmatrix}$$

$$\Longrightarrow \mathbf{z}^{(l+1)} = \mathbf{a}^{(l)} + \mathbf{z}^{(l)}$$

$$= \begin{pmatrix} \mathbf{u}_0' - (PP_1+P'PQ_1)\sum_{i=0}^{l}[G_2^{-1}]^i f(z^{(i)},t_0) \\ \mathbf{w}_0' - QP_1\sum_{i=0}^{l}[G_2^{-1}]^i f(z^{(i)},t_0) \\ \mathbf{u}_0 - PQ_1\sum_{l=0}^{i}[G_2^{-1}]^i f(z^{(i)},t_0) \\ QQ_1(\mathbf{u}_0-\mathbf{u}^0) - QQ_1\sum_{i=0}^{l}[G_2^{-1}]^i f(z^{(i)},t_0) \end{pmatrix} = \begin{pmatrix} \mathbf{y}_1^{(l+1)} \\ \mathbf{y}_2^{(l+1)} \\ \mathbf{x}_1^{(l+1)} \\ \mathbf{x}_2^{(l+1)} \end{pmatrix}$$

where $[G_2^{-1}]^i$ is the matrix G_2^{-1} evaluated in $\mathbf{z}^{(i)} := (\mathbf{y}_1^{(i)} - P'\mathbf{y}_2^{(i)}, \mathbf{x}_1^{(i)} + \mathbf{y}_2^{(i)})$.

Suppose $f(z^{(l)},t_0)=0$, i. e. we have found a solution of $\mathbf{F}(\mathbf{z},t_0)=0$ after l steps. But it is not sufficient to solve this equation in t_0. In order to substitute the occuring derivatives we have to consider the analogous system of equations in the point t_0+h, too. A simple example will explain this fact.

Example 2. We investigate the same DAE as in Example 1, i. e.

$$f(x', x, t) = \begin{pmatrix} x_1' - x_1^2 - x_2 - t^2 \\ -x_1 - t^3 \end{pmatrix}.$$

Choose the initial vectors $\mathbf{z}^{(0)} = \begin{pmatrix} \mathbf{u}^{(0)'} \\ \mathbf{w}^{(0)'} \\ \mathbf{u}^{(0)} \\ \mathbf{w}^{(0)} \end{pmatrix}$ and $\mathbf{z}_h^{(0)} = \begin{pmatrix} \mathbf{u}_h^{(0)'} \\ \mathbf{w}_h^{(0)'} \\ \mathbf{u}_h^{(0)} \\ \mathbf{w}_h^{(0)} \end{pmatrix}$, respectively, for the iteration start in t_0 resp. t_0+h. (Let all components of $\mathbf{z}_h^{(0)}$ be sufficiently close to those of $\mathbf{z}^{(0)}$.)

Having executed two steps of Newton's method we obtain the solutions

$$\mathbf{z}^{(2)} = \begin{pmatrix} \mathbf{u}_1^{(1)'} \\ \mathbf{u}_2^{(1)'} \\ 0 \\ -t_0^6 - 4t_0^2 - h(3t_0 + h) \\ -t_0^3 \\ 0 \\ 0 \\ -t_0^3 \end{pmatrix} =: \begin{pmatrix} \mathbf{u}_1^{(2)'} \\ \mathbf{u}_2^{(2)'} \\ \mathbf{w}_1^{(2)'} \\ \mathbf{w}_2^{(2)'} \\ \mathbf{u}_1^{(2)} \\ \mathbf{u}_2^{(2)} \\ \mathbf{w}_1^{(2)} \\ \mathbf{w}_2^{(2)} \end{pmatrix}$$

and

$$\mathbf{z}_h^{(2)} = \begin{pmatrix} \mathbf{u}_{h1}^{(0)'} \\ \mathbf{u}_{h2}^{(0)'} \\ 0 \\ -(t_0 + h)^6 - 4(t_0 + h)^2 + \mathbf{u}_{h1}^{(0)'} \\ -(t_0 + h)^3 \\ 0 \\ 0 \\ -(t_0 + h)^3 \end{pmatrix} =: \begin{pmatrix} \mathbf{u}_{h1}^{(2)'} \\ \mathbf{u}_{h2}^{(2)'} \\ \mathbf{w}_{h1}^{(2)'} \\ \mathbf{w}_{h2}^{(2)'} \\ \mathbf{u}_{h1}^{(2)} \\ \mathbf{u}_{h2}^{(2)} \\ \mathbf{w}_{h1}^{(2)} \\ \mathbf{w}_{h2}^{(2)} \end{pmatrix}.$$

Compute $\frac{\mathbf{u}_{h1}^{(2)} - \mathbf{u}_1^{(2)}}{h} = -3t_0^2 - h(3t_0 + h) = -3t_0^2 + 0(h)$. Now it is easy to verify that $\mathbf{F}(\mathbf{z}^{(2)}, t_0) = 0$ holds. Hence, $\mathbf{z}^{(2)} = \mathbf{z}^*$, and we can compute the complete initial value

$$x = \mathbf{u} + \mathbf{w}' = \begin{pmatrix} -t_0^3 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ -t_0^6 - 4t_0^2 - 0(h) \end{pmatrix} = \begin{pmatrix} -t_0^3 \\ -t_0^6 - 4t_0^2 - 0(h) \end{pmatrix}$$

which is obviously consistent.

References

[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.

[GP84] C. W. Gear and L. R. Petzold. Ode methods for the solution of differential algebraic systems. *SIAM J. Numer. Anal.*, 21:716–728, 1984.

[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.

[Lei88] B. J. Leimkuhler. *Approximation methods for the consistent initialization of differential-algebraic equations.* PhD thesis, Univ. of Illinois, Dept. of Computer Science, 1988. Report, UIUC DCS-R-88-1450.

[Mar89] R. März. Index-2 differential-algebraic equations. *Results in Math.*, 15:149–171, 1989.