On linear differential-algebraic equations and linearizations

Roswitha März, Berlin

Abstract

On the background of a careful analysis of linear DAEs, linearizations of nonlinear index-2 systems are considered. Finding appropriate function spaces and their topologies allows to apply the standard Implicit Function Theorem again. Both, solvability statements as well as the local convergence of the Newton-Kantorovich method (quasilinearization) result immediately. In particular, this applies also to fully implicit index 1 systems whose leading nullspace is allowed to vary with all its arguments.

Keywords. Differential algebraic equations, linearization, Newton-Kantorovich method

Introduction

Linearization plays an important standard role in the analysis and numerical treatment of regular differential equations. It is a very nice tool for proving solvability statements, showing asymptotic behaviour, describing the sensitivity with respect to parameters etc. Moreover, iterative linearization methods like the standard Newton-Kantorovich method, which is also well-known as quasilinearization of Bellmann and Kalaba, further damped and regularized versions of that method have proved their value in solving regular boundary value problems for a long time (e.g. Roberts and Shipman (1972), Miele and Iyer (1971), Aktas and Stetter (1977)).
For index 1 differential algebraic equations whose leading nullspace depends on time only, the corresponding linearizations are considered e.g. in März (1984,1986), Griepentrog and März (1986) and Tischendorf (1994). For index 2 DAEs positive results concerning the local solvability of initial value problems and Lyapunov stability via linearizations at consistent values resp. stationary solutions are obtained in März (1992).
The present paper mainly deals with linearizations of index 2 DAEs along given functions that are not necessarily supposed to solve the DAE. Solvability statements for index 2 DAEs are given under low smoothness demands. Further, the local convergence of the Newton-Kantorovich method is proved.
Note that even for the Newton-Kantorovich process we are interested in linearizations along functions not solving the DAE itself and not necessarily satisfying the first and second order constraint. In this context, the geometric approach of transferring the DAE locally to a vector field on the last order constraint manifold will fail to be such a useful tool, as it has been proved on different occasions.

Our main tool is the proposing of appropriate function spaces and operator notions of the DAE problems to obtain Fréchet derivatives that represents homeomorphisms again. Further, standard arguments apply.

The paper is organized as follows:
§ 1 collects some general preliminaries. In § 2 linear index 1 and index 2 results are prepared for being used below. The respective nonlinear index 1 DAEs are shortly mentioned in § 3. § 4 contains the new part for general index 2 DAEs on the background of the explanations in the

linear section. Moreover, the index 2 results are specified in § 5 for application to fully implicit index 1 DAEs whose leading nullspace is allowed to vary with all its arguments.

1 Preliminaries

Given the DAE

$$f(x'(t), x(t), t) = 0, \tag{1.1}$$

where $f : \mathbb{R}^m \times \mathcal{D} \times J \to \mathbb{R}^m$ is continuous and has continuous partial Jacobians $f'_{x'}, f'_x : \mathbb{R}^m \times \mathcal{D} \times J \longrightarrow L(\mathbb{R}^m)$, $\mathcal{D} \subseteq \mathbb{R}^m$ open, J an interval.
The nullspace of the leading Jacobian $f'_{x'}(y, x, t)$ is assumed to be invariant of y, x, that is

$$\ker f'_{x'}(y, x, t) = N(t), \quad (y, x, t) \in \mathbb{R}^m \times \mathcal{D} \times J. \tag{1.2}$$

Moreover, let $N(t)$ vary smoothly with t. This means, N is spanned by a base $n_1, \ldots,$ $n_{m-r} \in C^1(J, \mathbb{R}^m)$, $N(t) = \mathrm{span}\{n_1(t), \ldots, n_{m-r}(t)\}$. Then, $Q := K(K^TK)^{-1}K^T$ has the properties

$$Q \in C^1(J, L(\mathbb{R}^m)), \quad Q(t)^2 = Q(t), \quad \mathrm{im}\, Q(t) = N(t),\ t \in J, \tag{1.3}$$

where $K(t) := [n_1(t), \ldots, n_{m-r}(t)] \in L(\mathbb{R}^{m-r}, \mathbb{R}^m)$, that is, Q represents a C^1 projector function onto N.
On the other hand, if there is any projector function Q having the properties (1.3), the IVPs $n' = Q'n$, $n(t_0) = n_j^0$, $j = 1, \ldots, m-r$, generate an appropriate C^1 base, supposed $n_0^0, \ldots, n_{m-r}^0 \in \mathbb{R}^m$ form a base of $N(t_0)$ (cf. Griepentrog and März (1989)). Hence, the existence of a C^1 base and a C^1 projector function, respectively, are equivalent.
In the following, we denote by Q any C^1 projector function with (1.3), further $P := I - Q$.

Assumption (1.2) simply implies

$$f(y, x, t) - f(P(t)y, x, t) = \int_0^1 f'_{x'}(sy + (1-s)P(t)y, x, t)Q(t)y\,\mathrm{d}s = 0$$

for $(y, x, t) \in \mathbb{R}^m \times \mathcal{D} \times J$, and further

$$f(x'(t), x(t), t) = f(P(t)x'(t), x(t), t) = f((Px)'(t) - P'(t)x(t), x(t), t)$$

for functions $x \in C^1$. This makes clear that the derivative $(Qx)'$ does not appear in (1.1), in fact. The function space

$$C_N^1 := \{x \in C : Px \in C^1\} \tag{1.4}$$

suggests itself as the very natural one for the solutions of (1.1). We should ask for C_N^1 solutions, but not for C^1 solutions in general.
In particular, for semi-explicit equations

$$\left.\begin{aligned} x_1'(t) + \varphi(x_1(t), x_2(t), t) &= 0 \\ \psi(x_1(t), x_2(t), t) &= 0 \end{aligned}\right\} \tag{1.5}$$

we have simply $P = \mathrm{diag}(I, 0)$, $C_N^1 := \{x \in C : x_1 \in C^1\}$.

Higher smoothness of the solution corresponds to higher smoothness demands for the given data, but in most applications one is interested even in lower smoothness.
On this background, (1.1) should be written precisely as

$$f((Px)'(t) - P'(t)x(t), x(t), t) = 0.$$

However, for shortness, we continue to use (1.1) and interpret $P(t)x'(t)$ as an abbreviation of $P(t)((Px)'(t) - P'(t)x(t))$ there.

Next, given a C_N^1 function x_* whose trajectory proceeds in $\mathcal{D}$. For different reasons we might be interested in the linearization of (1.1) along x_*, that is, in the linear equation

$$A(t)z'(t) + B(t)z(t) = q(t), \tag{1.6}$$

the continuous coefficients of which are given by

$$\begin{aligned} A(t) &:= f'_{x'}(y_*(t), x_*(t), t), \\ B(t) &:= f'_x(y_*(t), x_*(t), t), \\ y_*(t) &:= (Px_*)'(t) - P'(t)x_*(t). \end{aligned}$$

Here, x_* is often supposed to be a solution (stationary or nonstationary) of the DAE (1.1). With $x = x_* + z$, equation (1.1) itself may be described approximately by

$$A(t)z'(t) + B(t)z(t) = -f(y_*(t), x_*(t), t), \tag{1.7}$$

supposed z is small enough (in C_N^1) for the Taylor expansion remainder term to be neglected. In particular, starting with a solution x_* and a small perturbation z we arrive at a linear equation (1.6) with a small right-hand side q caused by the small remainder term only. The corresponding equation (1.7) is the homogeneous one.
However, the whole nice game of linearization is to know the opposite: Solving (1.6) for small or vanishing q we should like to have information on how the solutions neighbouring to x_* behave. However, for that the resulting solution z should be small enough in C_N^1.

It seems very natural to measure the size of q in the max-norm $\|\cdot\|_\infty$ of the continuous function space C. On this background, linearizations are shown to work well for the index 1 case (März (1984), (1986)). Unfortunately, for higher index DAEs (1.6), the relations $\|q\|_\infty \to 0$, $z(t_0) = 0$ do not necessarily imply $\|z\|_{C_N^1} \to 0$, or at least $\|z\|_\infty \to 0$ (e.g. Griepentrog and März (1986), p. 21). Thus, from this point of view, it is rather doubtful whether linearization can work well in the index 2 case.
By considering both DAEs (1.1) and (1.6) in further detail, we try to learn more about how to measure the size of q for maintaining the comfort of linearization also in the index 2 case. And, surprisingly, we will succeed!

At this place it should be mentioned that freezing the time t at, say, $t_* \in J$ and considering the resulting constant coefficient equation

$$A(t_*)z'(t) + B(t_*)z(t) = q(t) \tag{1.8}$$

instead of (1.6) does not make sense in the higher index case in general. This is shown in § 1.3.1 of Griepentrog and März (1986) by different examples. In particular, (1.8) may have index 2, but (1.6) does not, and the opposite may also happen. From this point of view, linearizing the

DAE (1.1) at a given point $(y_0, x_0, t_0) \in I\!R^m \times \mathcal{D} \times J$ seems to be rather useless in the higher index case.
Fortunately, the situation becomes much easier if we start with a certain autonomous DAE (1.1) and linearize at a stationary solution. Then, the linearized equation has constant coefficients arising in a somewhat more natural way, and, in fact, it provides information on how the solutions of (1.1) behave asymptotically (cf. März (1992), Tischendorf (1994)). As a consequence, straightforward generalizations of Lyapunov-Theorems result.

Next, turn to boundary value problems (BVPs) for (1.1). As usually we state the boundary condition by means of a C^1 function $r : I\!R^m \times I\!R^m \to I\!R^m$ to be

$$r(x(t_0), x(T)) = 0. \tag{1.9}$$

Let r_1' and r_2' denote the partial derivatives of r with respect to the first and second components, respectively. The range of the matrix $[r_1'(z_1, z_2), r_2'(z_1, z_2)]$ is supposed to be constant. This matrix has full rank m in case of regular ODEs. However, for DAEs there should be a lower number of independent boundary conditions.
The standard Newton-Kantorovich method (or quasilinearization) and its modifications are approved to work well for BVPs in regular ODEs (e.g. Aktas and Stetter (1977)). But what about the DAE case? Apply the standard Newton-Kantorovich algorithm to our BVP (1.1), (1.9). Starting with an appropriate initial guess function x_0 from C_N^1 we try to form the iterations

$$x_{j+1} = x_j + z_{j+1}, \quad j \geq 0, \tag{1.10}$$

where $z_{j+1} \in C_N^1$ is determined to solve the BVP linearized along x_j, i.e.

$$A_{(j)}(t) z_{j+1}'(t) + B_{(j)}(t) z_{j+1}(t) = -f(x_j'(t), x_j(t), t), \tag{1.11}$$
$$r_1'(x_j(t_0), x_j(T)) z_{j+1}(t_0) + r_2'(x_j(t_0), x_j(T)) z_{j+1}(T) = -r(x_j(t_0), x_j(T)), \tag{1.12}$$

where

$$\begin{aligned} A_{(j)}(t) &:= f_{x'}'(x_j'(t), x_j(t), t), \quad t \in [t_0, T], \\ B_{(j)}(t) &:= f_x'(x_j'(t), x_j(t), t), \quad t \in [t_0, T]. \end{aligned}$$

This gives rise to the following questions: Does the linear DAE have the same index as the nonlinear one has? Does the linear BVP (1.11), (1.12) uniquely determine the correction $z_{j+1} \in C_N^1$? Further, does x_j converge to a solution of the nonlinear BVP (1.1), (1.9). If it does so, in what sense?

In März (1984), BVPs in index 1 DAEs with properly stated boundary conditions are considered. If $x_* \in C_N^1$ denotes the BVP solution to be approximated, we may realize the Newton-Kantorovich iteration process with any initial guess x_0 being close enough to x_* in C_N^1. Then, $x_j \to x_*$ $(j \to \infty)$ in C_N^1 i.e. $\|x_j - x_*\|_\infty + \|Px_j - Px_*\|_\infty \to 0$ $(j \to \infty)$, becomes true. It should be stressed that there is no need for x_0 to satisfy any further constraint.
Below we will show a similar result for the index 2 case. Again the iterations can be realized with an initial guess function x_0, which does not satisfy neither the first nor the second order constraint. It seems that this will turn out to be a very special advantage of quasilinearizations applied to DAEs.

2 Linear index 1 and index 2 equations

Consider the linear equation

$$A(t)x'(t) + B(t)x(t) = q(t), \quad t \in J, \tag{2.1}$$

with continuous matrix coefficients. Introduce the basic subspaces

$$\begin{aligned} N(t) &:= \ker A(t) \subset \mathbb{R}^m, \\ S(t) &:= \{z \in \mathbb{R}^m : B(t)z \in \operatorname{im} A(t)\} \end{aligned}$$

and assume $N(t)$ to vary smoothly with t. Obviously, $S(t)$ is the subspace where the homogeneous equation solutions proceed.
Again, let $Q \in C^1(J, L(\mathbb{R}^m))$ denote a projector function such that

$$Q(t)^2 = Q(t), \quad \operatorname{im} Q(t) = N(t), \quad t \in J,$$

further $P(t) := I - Q(t)$.

Definition (Griepentrog and März (1986)): The DAE (2.1) is said to be index 1 (or transferable) on J if

$$N(t) \oplus S(t) = \mathbb{R}^m, \quad t \in J, \tag{2.2}$$

becomes true.

Condition (2.2) implies the matrices

$$A_1(t) := A(t) + (B(t) - A(t)P'(t))Q(t), \quad t \in J, \tag{2.3}$$

to be nonsingular. The matrices

$$G_1(t) := A(t) + B(t)Q(t), \quad t \in J, \tag{2.4}$$

are nonsingular simultaneously and, further, $A_1 = G_1(I - PP'Q)$.
Multiplying (2.1) by PA_1^{-1} and QA_1^{-1} we decouple this equation into the system

$$\left.\begin{aligned} (Px)' - P'Px + PA_1^{-1}BPx &= PA_1^{-1}q \\ Qx + QA_1^{-1}BPx &= QA_1^{-1}q \end{aligned}\right\}. \tag{2.5}$$

Now, a solution expression results immediately. In fact we have

$$x = Px + Qx = (I - QA_1^{-1}B)u + QA_1^{-1}q \in C_N^1,$$

where $u \in C^1$ solves the inherent regular ODE

$$u' - P'u + PA_1^{-1}Bu = PA_1^{-1}q$$

and starts at $u(t_0) \in \operatorname{im} P(t_0)$ for some $t_0 \in J$. The matrix

$$I - Q(t)A_1(t)^{-1}B(t) =: P_{\mathrm{can}}(t) \tag{2.6}$$

may easily be shown to represent the projector onto $S(t)$ along $N(t)$. This is why it is called the canonical projector for the index 1 case.

Geometrically, the index 1 case means that the subspace $S(t) = \mathrm{im}\, P_{\mathrm{can}}(t)$ is filled by the solutions of the homogeneous equation. More precisely, for each given $t_0 \in J$, $x_0 \in S(t_0)$, there is exactly one solution of the homogeneous DAE, passing through x_0 at time t_0.
Obviously, the DAE (2.1) is solvable for each $q \in C$, and the solution is given on the whole interval J. Moreover, the solution depends continuously on the inhomogeneity.
Recall further that (2.2) is equivalent for the matrix pencils $\{A(t), B(t)\}$ and $\{A(t), B(t) - A(t)P'(t)\}$, $t \in J$, to be regular with index 1.

For higher index DAEs, in particular for those having index 2, the situation becomes more distinct. Geometrically, only a certain subspace of $S(t)$ is filled by the homogeneous DAE solutions. The inhomogeneous DAE is no more solvable for all continuous q, but only for those q having certain smoother components additionally. In the consequence, $\|q\|_\infty \to 0$ does not imply $\|x\|_\infty \to 0$ for the DAE solution satisfying homogeneous initial conditions, that is, the DAE solution does not depend continuously of the source q (in the given topologies). Moreover, the local matrix pencil $\{A(t), B(t)\}$ makes no sense for the DAE in general.

To be more precise, we have to deal with certain additional subspaces. Introduce

$$\begin{aligned} N_1(t) &:= \ker A_1(t) \subset \mathbb{R}^m \\ S_1(t) &:= \{z \in \mathbb{R}^m : B(t)P(t)z \in \mathrm{im}\, A_1(t)\}. \end{aligned}$$

The nullspace $N_1(t)$ has the same dimension as $N(t) \cap S(t)$.

Definition (März (1989)): The DAE (2.1) is said to be index-2 tractable (shortly index 2) on J if the conditions

$$\left.\begin{aligned} &\dim N_1(t) = \mathrm{const} > 0, \\ &N_1(t) \oplus S_1(t) = \mathbb{R}^m, \quad t \in J, \end{aligned}\right\} \tag{2.7}$$

are valid.

Supposing that (2.1) has index 2 we introduce the projector $Q_1(t)$ onto $N_1(t)$ along $S_1(t)$, $P_1(t) := I - Q_1(t)$, $t \in J$. Now, the matrix

$$G_2(t) := A_1(t) + B(t)P(t)Q_1(t), \quad t \in J,$$

is known to be nonsingular. Further, for the projector $Q_1(t)$, the relations

$$Q_1(t) = Q_1(t)G_2(t)^{-1}B(t)P(t), \qquad Q_1(t)Q(t) = 0 \tag{2.8}$$

become true.
If, additionally, Q_1 belongs to the class C^1, we form

$$\begin{aligned} A_2 &:= A_1 + B_1Q_1, \qquad B_1 := (B - A_1(PP_1)')P, \\ A_2 &= G_2(I - P_1(PP_1)'PQ_1). \end{aligned}$$

Obviously, $A_2(t)$ is nonsingular since $G_2(t)$ is so. Further, we have

$$Q_1 = Q_1G_2^{-1}BP = Q_1A_2^{-1}BP = Q_1A_2^{-1}B_1.$$

Next we decompose the unknown solution into $x = Qx + Px = Qx + PP_1x + PQ_1x =: w + u + Pv$ and multiply (2.1) by $PP_1A_2^{-1}$, $QP_1A_2^{-1}$ and $Q_1A_2^{-1}$, respectively. After carrying out a few technical computations we obtain the decoupled system

$$u' - (PP_1)'u + PP_1A_2^{-1}Bu = PP_1A_2^{-1}q, \tag{2.9}$$

$$-(Qv)' + (QQ_1)'(u + Pv) + w + QP_1A_2^{-1}Bu = QP_1A_2^{-1}q, \tag{2.10}$$

$$v = Q_1A_2^{-1}q, \tag{2.11}$$

where $u(t_0) \in \mathrm{im}\, P(t_0)P_1(t_0)$ at some $t_0 \in J$ implies $u = PP_1u$.

In particular, in case $q(t)$ vanishes identically, the solution component $Q_1(t)x(t) = v(t)$ does so, too. Hence, the homogeneous equation solution is given by

$$x = u + w = (I - (QQ_1)' - QP_1A_2^{-1}B)u = (I - (QQ_1)' - QP_1A_2^{-1}B)PP_1u. \tag{2.12}$$

Denote $\pi_{\text{can}} := (I - (QQ_1)' - QP_1A_2^{-1}B)PP_1$. It may be checked immediately that

$$PP_1\pi_{\text{can}} = PP_1, \quad \pi_{\text{can}}^2 = \pi_{\text{can}}, \quad \ker \pi_{\text{can}}(t) = \ker P(t)P_1(t) = N(t) \oplus N_1(t)$$

hold true. The next assertion makes clear why π_{can} is said to be the canonical projector for the index 2 case. At this point it should be noticed that, in the constant coefficient case, π_{can} represents nothing else but the spectral projection onto the (relative) finite eigenspace of the matrix pencil $\{A, B\}$ along the infinite one (cf. Lewis (1986), März (1993)). Hence, $\pi_{\text{can}}(t)$ may be understood as the spectral projector for the timevarying case.
Obviously, the canonical projector $\pi_{\text{can}}(t)$ is much more complicate than the projector $P(t)P_1(t)$. Fortunately, a lot of things can already be achieved by using the easier projection only. However, we should always keep in mind the strongly close relationship of both projectors.

Theorem 2.1 *Let* (2.1) *be an index* 2 *DAE with continuously differentiable* Q_1. *Then the subspace* $\mathrm{im}\, \pi_{\text{can}}(t) \subset S(t)$ *describes the homogeneous equation solution space, i.e. through each given* $t_0 \in J$, $x_0 \in \mathrm{im}\, \pi_{\text{can}}(t_0)$, *there passes exactly one solution.*

Theorem 2.1 as well as the next one are derived immediately by considering (2.9) – (2.11). While $S(t)$ is related to the first order constraint, $\mathrm{im}\, \pi_{\text{can}}(t)$ describes the second order one.

Theorem 2.2 *Let* (2.1) *be an index* 2 *DAE with continuously differentiable* Q_1.

(i) *Then the DAE* (2.1) *is solvable on* C_N^1 *for all* $q \in C_{(2)}^1 := \{q \in C : Q_1A_2^{-1}q \in C^1\}$.

(ii) *For the linear map* $L : C_N^1 \to C$, $Lx := A(Px)' + (B - AP')x$, $x \in C_N^1$, *the image space is* $C_{(2)}^1$, *i.e.* $\mathrm{im}\, L = C_{(2)}^1 \subset C$.

(iii) *The initial value problems*

$$Lx = q, \quad P(t_0)P_1(t_0)(x(t_0) - x^0) = 0 \tag{2.13}$$

are uniquely solvable on C_N^1 *for any given* $t_0 \in J$, $x^0 \in I\!R^m$, $q \in C_{(2)}^1$.

(iv) *Relating the max-norm to any compact interval* $J_0 \subseteq J$, $t_0 \in J_0$, *the inequality*

$$\|x\| := \|x\|_\infty + \|(Px)'\|_\infty \le K(J_0)(\|q\|_\infty + \|(Q_1A_2^{-1}q)'\|_\infty + |P(t_0)P_1(t_0)x^0|) \tag{2.14}$$

is satisfied by the IVP solution.

Corollary 2.3 *Let $M_{(2)} := \mathrm{im}\, P(t_0)P_1(t_0)$. Then, the IVP map $\mathcal{L} : C_N^1 \to C_{(2)}^1 \times M_{(2)} \subset C \times \mathbb{R}^m$,*

$$\mathcal{L}x := (Lx, P(t_0)P_1(t_0)x(t_0)), \quad x \in C_N^1,$$

acts bijectively from C_N^1 onto $C_{(2)}^1 \times M_{(2)}$.

Next, for a fixed compact interval $J_0 \subseteq J$, we equip the function spaces $C(J_0, \mathbb{R}^m)$ and $C_N^1(J_0, \mathbb{R}^m)$ with their natural norms $\|x\|_\infty := \max\{|x(t)| : t \in J_0\}$, $x \in C(J_0, \mathbb{R}^m)$, and $\|x\| := \|x\|_\infty + \|(Px)'\|_\infty$, $x \in C_N^1(J_0, \mathbb{R}^m)$, respectively. Clearly, by doing so, both $C(J_0, \mathbb{R}^m)$ and $C_N^1(J_0, \mathbb{R}^m)$ become Banach spaces and the linear map $\mathcal{L} : C_N^1(J_0, \mathbb{R}^m) \to C(J_0, \mathbb{R}^m) \times M_{(2)}$ is bounded. Furthermore, $C_{(2)}^1(J_0, \mathbb{R}^m) \subset C(J_0, \mathbb{R}^m)$ is a proper but nonclosed subset, which causes the inverse $\mathcal{L}^{-1}$ to become an unbounded map in the given natural topologies.
However, equipping $C_{(2)}^1(J_0, \mathbb{R}^m)$ with the stronger norm

$$\|w\|_{(2)} := \|w\|_\infty + \|(Q_1 A_2^{-1} w)'\|_\infty, \quad w \in C_{(2)}^1(J_0, \mathbb{R}^m),$$

we obtain again a Banach space, and may then turn to considering the map $\mathcal{L} : C_N^1(J_0, \mathbb{R}^m) \to C_{(2)}^1(J_0, \mathbb{R}^m) \times M_{(2)}$. Due to the inequality (2.14), $\mathcal{L}^{-1}$ is bounded in this new setting.
Additionally, the inequalities

$$\begin{aligned} &\|Lx\|_\infty \le K_1(\|x\|_\infty + \|(Px)'\|_\infty), \\ &\|Q_1 A_2^{-1} Lx\|_\infty = \|Q_1 x\|_\infty \le K_2 \|x\|_\infty, \quad x \in C_N^1(J_0, \mathbb{R}^m), \end{aligned}$$

are valid, hence

$$\|Lx\|_{(2)} \le K\|x\|, \quad x \in C_N^1(J_0, \mathbb{R}^m),$$

that is, L and further $\mathcal{L}$ are bounded also with respect to the new norm $\|\cdot\|_{(2)}$. So, we obtain the next assertion, which will prove its value in § 4.

Corollary 2.4 *Relate C_N^1, $\|\cdot\|$ and $C_{(2)}^1$, $\|\cdot\|_{(2)}$ to a compact interval $J_0 \subseteq J$. Then $\mathcal{L}$ is a homeomorphism of C_N^1 onto $C_{(2)}^1 \times M_{(2)}$ in these new topologies.*

Let us finish this section by discussing the so-called Hessenberg form index 2 equations in detail, i.e.

$$\left.\begin{aligned} x_1' + B_{11}x_1 + B_{12}x_2 &= q_1 \\ B_{21}x_1 \qquad\qquad &= q_2 \end{aligned}\right\}, \tag{2.15}$$

where $B_{21}(t)B_{12}(t)$ is supposed to be nonsingular on the given interval J. In our context this corresponds to

$$A = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & 0 \end{bmatrix}, \qquad Q = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix},$$

$$A_1 = G_1 = A + BQ = \begin{bmatrix} I & B_{12} \\ 0 & 0 \end{bmatrix},$$

$$S(t) = S_1(t) = \{(z_1^T, z_2^T)^T \in \mathbb{R}^m : B_{21}(t)z_1 = 0\}.$$

Obviously, $z \in S_1(t) \cap N_1(t)$ implies $z = 0$, hence (2.15) is index 2 tractable, indeed.

The block $H(t) := B_{12}(t)(B_{21}(t)B_{12}(t))^{-1}B_{21}(t)$ describes the projector onto im $B_{12}(t)$ along ker $B_{21}(t)$. Further, the projector function Q_1 now reads

$$Q_1 = \begin{bmatrix} H & 0 \\ -F & 0 \end{bmatrix},$$

where $F(t) := (B_{21}(t)B_{12}(t))^{-1}B_{21}(t)$. This leads to simple projector functions

$$PP_1 = \begin{bmatrix} I-H & 0 \\ 0 & 0 \end{bmatrix}, \qquad PQ_1 = \begin{bmatrix} H & 0 \\ 0 & 0 \end{bmatrix}.$$

However, the canonical projector π_{can} looks a bit more complex, namely

$$\pi_{\text{can}} = \begin{bmatrix} I-H & 0 \\ -FB_{11}(I-H)+F'(I-H) & 0 \end{bmatrix}.$$

Recall that $\pi_{\text{can}}(t)$ describes the subspace where the homogeneous equation solution proceeds. We have im $\pi_{\text{can}}(t) \subset S(t)$, $t \in J$. The larger subspace $S(t)$ relates to the first order constraint; im $\pi_{\text{can}}(t)$ relates to the second order constraint.

3 Nonlinear index 1 equations

In this section we return to the nonlinear equation

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

as it is described in the beginning of § 1. Besides the nullspace $N(t)$ we introduce the subspace

$$S(y,x,t) := \{z \in \mathbb{R}^m : f'_x(y,x,t)z \in \text{im } f'_{x'}(y,x,t)\}$$

and, moreover, the matrix

$$G_1(y,x,t) := f'_{x'}(y,x,t) + f'_x(y,x,t)Q(t).$$

Definition: The DAE (3.1) is said to be index 1 on the open set $\mathcal{G} \subseteq \mathbb{R}^m \times \mathcal{D} \times J$ if

$$N(t) \oplus S(y,x,t) = \mathbb{R}^m, \quad (y,x,t) \in \mathcal{G}. \tag{3.2}$$

is valid.

Condition (3.2) is well-known to be equivalent for the matrix $G_1(y,x,t)$ to be nonsingular, and further for the pencil $\{f'_{x'}(y,x,t), f'_x(y,x,t)\}$ to be regular with index 1.

What about linearizing an index 1 DAE (3.1) along a given C^1_N function x_*? It may be checked immediately that the resulting linear equation (1.6) has always index 1, too. In particular, we have precisely

$$S(t) = \{z \in \mathbb{R}^m : B(t)z \in \text{im } A(t)\} = S(y_*(t), x_*(t), t).$$

On the other hand, if we do not know what the index of (3.1) is, but if we are sure about (1.6) to have index 1, we can conclude that (3.1) has index 1 also in a neighbourhood of the graph of x_* in $\mathbb{R}^m \times \mathcal{D} \times J$, since the matrix $G_1(y,x,t)$ depends continuously on its argument and $G_1(y_*(t), x_*(t), t)$ is nonsingular.

Theorem 3.1 *Let $x_* \in C_N^1([t_0,T],\mathbb{R}^m)$ be a solution of the BVP* (1.9), (1.1), *and let the linearized equation* (1.6) *be index* 1. *Furthermore, let for the matrix*

$$S := r_1'(x_*(t_0),x_*(T))X(t_0) + r_2'(x_*(t_0),x_*(T))X(T)$$

the conditions

$$\begin{aligned}\ker S &= N(t_0),\\ \operatorname{im} S &= \operatorname{im}(r_1'(x_*(t_0),x_*(T)),r_2'(x_*(t_0),x_*(T)) =: M_{(1)}\end{aligned}$$

be valid, where the fundamental solution matrix X of (1.6) *is uniquely determined by $AX'+BX = 0$, $P(t_0)(X(t_0)-I)=0$.*
Then, the following assertions become true.

(i) *The perturbed BVPs*

$$\begin{aligned} f(x'(t),x(t),t) &= q(t), \quad t\in[t_0,T]\\ r(x(t_0),x(T)) &= d,\end{aligned}$$

$q\in C([t_0,T],\mathbb{R}^m)$, $d\in M_{(1)}$, $\|q\|_\infty$, $|d|$ sufficiently small, are uniquely solvable on $C_N^1([t_0,T],\mathbb{R}^m)$. For the BVP solution $x(\cdot;q,d)$ the inequality

$$\|x(\cdot;q,d)-x_*\| \le K(\|q\|_\infty + |d|)$$

with a constant $K>0$ is valid.

(ii) *For sufficiently small $\|x_0 - x_*\|$, the Newton-Kantorovich method* (1.10) – (1.12) *with the initial guess x_0 provides an uniquely determined sequence $\{x_j\}_{j\ge 0}$ which converges in $C_N^1([t_0,T],\mathbb{R}^m)$ to x_*. If, additionally, the partial Jacobians $f_{x'}'$, f_x' are locally Lipschitz, x_j tends to x_* quadratically.*

For the proof see März (1984).

In particular, if x_* is any solution of the DAE (1.1) on the interval $[t_0,T] =: J_*$, the initial value problems for (1.1) with the initial condition

$$P(t_0)x(t_0) = P(t_0)x^0, \quad x^0\in\mathbb{R}^m,\ \ |P(t_0)x^0| \text{ small},$$

are uniquely solvable at least on that interval J_*. To realize this, we simply choose $r(u,v) = P(t_0)u$ and apply Theorem 3.1(i).
It should be emphasized once more that there is no need for the initial guess function x_0 in (ii) to satisfy the constraint, that is, the derivative free part of (1.1). In particular, for the semi-explicit system (1.5), we may choose $x_0\in C_N^1$, which does not satisfy the second equation of (1.5).

4 Nonlinear index 2 equations

Continue to discuss equation

$$f(x'(t),x(t),t) = 0, \tag{4.1}$$

but now supposing $G_1(y,x,t)$ to be singular everywhere. At the same time also the matrix function

$$A_1(y,x,t) = G_1(y,x,t) - f'_{x'}(y,x,t)P'(t)Q(t)$$

becomes singular for all $(y,x,t) \in I\!R^m \times \mathcal{D} \times J$. Introduce the subspaces

$$\begin{aligned} N_1(y,x,t) &:= \ker A_1(y,x,t) \subset I\!R^m \\ S_1(y,x,t) &:= \{z \in I\!R^m : f'_x(y,x,t)P(t)z \in \operatorname{im} A_1(y,x,t)\}. \end{aligned}$$

Definition: The DAE (4.1) is said to be index-2 tractable on the open set $\mathcal{G} \subseteq I\!R^m \times \mathcal{D} \times J$ if the conditions

$$\left.\begin{aligned} &\dim N_1(y,x,t) = \text{const} > 0, \\ &N_1(y,x,t) \cap S_1(y,x,t) = \{0\},\ (y,x,t) \in \mathcal{G} \end{aligned}\right\} \tag{4.2}$$

are satisfied.

For index 2 DAEs, the matrix

$$G_2(y,x,t) := A_1(y,x,t) + f'_x(y,x,t)P(t)Q_1(y,x,t)$$

is everywhere nonsingular. Thereby, $Q_1(y,x,t)$ denotes the projector onto $N_1(y,x,t)$ along $S_1(y,x,t)$. Further, let $P_1(y,x,t) := I - Q_1(y,x,t)$.
The projector $P(y,x,t)P_1(y,x,t)$ is closely related to the state manifold of the DAE. Roughly speaking, its counterpart $\pi_{\text{can}}(y,x,t)$ (cf. § 2 for the linear case) describes the tangent space of that manifold.

Our notion of index-2 tractability is a straightforward generalization of the corresponding definition for the linear case, which, in its turn, represents a generalization of the Kronecker index. On the other hand, also nonlinear index-2 Hessenberg systems are known to be index-2 tractable (cf. also (4.5) below).

Let us turn to a linearization (1.6) taken along a fixed function $x_* \in C^1_N$ whose trajectory remains in $\mathcal{D}$. Now we find the relations

$$\begin{aligned} A_1(t) &= A_1(y_*(t), x_*(t), t), \qquad & G_2(t) &= G_2(y_*(t), x_*(t), t), \\ S_1(t) &= S_1(y_*(t), x_*(t), t), \qquad & N_1(t) &= N_1(y_*(t), x_*(t), t), \end{aligned}$$

which obviously imply the linearized DAE (1.6) to be index-2 tractable.
The opposite is not true in general. The index-2 tractability of a linearized at $x_* \in C^1_N$ DAE does not necessarily spread out onto a neighbourhood of $\{(y_*(t), x_*(t), t) : t \in J_*) \in I\!R^m \times \mathcal{D} \times J$ (cf. März and Tischendorf (1994) for an example). However, by means of certain structural restrictions of (4.1) we may guarantee that property. For this purpose, let $\mathcal{D}$ be constituted so that $x \in \mathcal{D}$ implies $\{P(t)x : t \in J\} \subset \mathcal{D}$.

Lemma 4.1 *Given $x_* \in C^1_N(J_*, I\!R^m)$, $J_* \subseteq J$, $x_*(t) \in \mathcal{D}$ for $t \in J_*$. Let the linearized DAE* (1.6) *be index-2 tractable on J_*, and let the structural condition*

$$Q_1(t)G_2(t)^{-1}\{f(y,x,t) - f(0,P(t)x,t)\} = 0, \quad (y,x,t) \in \mathcal{N} \tag{4.3}$$

be given on a neighbourhood $\mathcal{N}$ of the graph

$$\mathcal{T} := \{(y_*(t), x_*(t), t) : t \in J_*\} \subset I\!R^m \times \mathcal{D} \times J.$$

Then, the DAE (4.1) *is index-2 tractable on a neighbourhood $\mathcal{N}_1 \subseteq \mathcal{N}$ of $\mathcal{T}$.*

Proof: Condition (4.3) implies $Q_1(t)G_2(t)^{-1}f'_{x'}(y,x,t)=0$, further $Q_1(t)G_2(t)^{-1}f'_x(y,x,t)=Q_1(t)G_2(t)^{-1}f'_x(0,P(t)x,t)P(t)$, and consequently $Q_1(t)G_2(t)A_1(y,x,t)=0$ for all $(y,x,t)\in\mathcal{N}$. Since the nullspace of $Q_1(t)$ is precisely $S_1(t)$, it follows that

$$\text{im}\, G_2(t)^{-1}A_1(y,x,t)\subseteq S_1(t),$$

thus rank $G_2(t)^{-1}A_1(y,x,t)\le\mu:=\dim S_1(t)=\text{rank}\, P_1(t)$. On the other hand, due to

$$G_2(t)^{-1}A_1(y_*(t),x_*(t),t)=G_2(t)^{-1}A_1(t)=P_1(t)$$

the matrix $A_1(y,x,t)$ is of constant rank μ in a neighbourhood $\mathcal{N}_0\subseteq\mathcal{N}$ of $\mathcal{T}$.
Now, the orthoprojector $Q_1^{\perp}(y,x,t)$ onto $N_1(y,x,t)$ depends continuously on its arguments there, since $A_1(y,x,t)$ does so. It follows further that the matrix

$$G_2^{\perp}(y,x,t):=A_1(y,x,t)+f'_x(y,x,t)P(t)Q_1^{\perp}(y,x,t)$$

is also continuous with respect to (y,x,t). Due to Lemma A.1 in März and Tischendorf $G_2^{\perp}(y,x,t)$ is nonsingular for $(y,x,t)\in\mathcal{T}$, but for reasons of continuity for $(y,x,t)\in\mathcal{N}_1\subseteq\mathcal{N}_0$, $\mathcal{N}_1\supset\mathcal{T}$, too. Applying that Lemma A.1 once more, we conclude that
$N_1(y,x,t)\cap S_1(y,x,t)=\{0\}$ holds for all $(y,x,t)\in\mathcal{N}_1$. □

Condition (4.3) seems to be more transparent and easier to be checked in its equivalent form

$$S(t)\{f(y,x,t)-f(0,P(t)x,t)\}\in\text{im}\, S(t)B(t)Q(t), \tag{4.4}$$

where $S(t):=I-A(t)A(t)^+$. The equivalence is given by means of the relation

$$\ker Q_1(t)G_2(t)^{-1}=\{z\in\mathbb{R}^m : S(t)z\in\text{im}\, S(t)B(t)Q(t)\}$$

(e.g. März and Tischendorf (1994), Lemma 3.3). In particular, the relation (4.4) is satisfied trivially for all linear DAEs, where $f(y,x,t)=A(t)y+B(t)x-q(t)$, thus

$$S(t)\{f(y,x,t)-f(0,P(t)x,t)\}=S(t)B(t)Q(t)x.$$

Most authors discussing index-2 DAEs restrict their interest to so-called Hesssenberg form systems from the very beginning, that is, to systems

$$\left.\begin{array}{rcl} x_1'+\varphi(x_1,x_2,t) &=& 0\\ \psi(x_1,t) &=& 0\end{array}\right\}, \tag{4.5}$$

where $\varphi'_{x_2}(x_1,x_2,t)\psi'_{x_1}(x_1,t)$ is supposed to be nonsingular. For this kind of special DAEs we derive

$$A=P=\begin{bmatrix} I & 0\\ 0 & 0\end{bmatrix},\qquad A_1=\begin{bmatrix} I & \varphi'_{x_2}\\ 0 & 0\end{bmatrix},$$

$$N_1(y,x,t)=\{z\in\mathbb{R}^m : z_1+\varphi'_{x_2}(x_1,x_2,t)z_2=0\}$$

and

$$S_1(y,x,t)=\{z\in\mathbb{R}^m : \psi'_{x_1}(x_1,x_2,t)z_1=0\}.$$

Now, index-2 tractability becomes obvious to be equivalent with the above nonsingularity condition. Furthermore, the structural condition (4.4) is satisfied because the nullspace component x_2 does not appear at all in the second equation of (4.5). Namely, in this case we simply have

$$S(t)\{f(y,x_1,x_2,t)-f(0,x_1,0,t)\}=\begin{bmatrix} 0 & 0\\ 0 & I\end{bmatrix}\begin{pmatrix} y_1+\varphi(x_1,x_2,t)-\varphi(x_1,0,t)\\ 0\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}.$$

We have seen that condition (4.4) covers both the linear equations and the Hessenberg form ones. However, also more general equations may be considered. In particular, when transforming quasilinear index-1 DAEs whose leading nullspace varies with x and t into its enlarged form we typically obtain a system

$$\left.\begin{array}{rcl} x_1' + \varphi(x_1, x_2, t) & = & 0 \\ \psi(x_1, t) + \chi(x_1, t)x_2 & = & 0 \end{array}\right\} \tag{4.6}$$

which has index 2 and satisfies the structural condition (4.4), supposed the range of $\chi(x_1, t)$ does not vary with x_1. We will discuss this case in § 5 in more detail.

Lemma 4.2 *Let the structural condition* (4.3) *be satisfied for* (1.1). *Then it is also valid for the enlarged systems*

(i) $$\left.\begin{array}{l} f(x'(t), x(t), t) = 0 \\ g(z(t), x(t), P(t)x'(t), t) = 0 \end{array}\right\},$$

where g_z' is supposed to be nonsingular, and

(ii) $$\left.\begin{array}{l} f(x'(t), x(t), t) = 0 \\ h(w'(t), w(t), P(t)x'(t), x(t), t) = 0 \end{array}\right\},$$

where $h_{w'}'$ is supposed to be nonsingular.

Proof: (i) We put the enlarged system back into the form $\hat{f}(\hat{x}'(t), \hat{x}(t), t) = 0$ with $\hat{x} = \binom{x}{z}$. Then we compute

$$\begin{aligned} \hat{f}_{\hat{x}'}' &= \begin{bmatrix} f_{x'}' & 0 \\ g_{x'}' & 0 \end{bmatrix}, \quad \hat{f}_{\hat{x}}' = \begin{bmatrix} f_x' & 0 \\ g_x' & g_z' \end{bmatrix}, \quad \hat{Q} = \begin{bmatrix} Q & 0 \\ 0 & I \end{bmatrix}, \\ \hat{A}_1 &= \begin{bmatrix} A_1 & 0 \\ & g_z' \end{bmatrix}, \qquad := g_{x'}' + (g_x' + g_{x'}' P')Q, \\ \hat{S}_1 &= \left\{ \begin{pmatrix} u \\ v \end{pmatrix} : f_x' P u \in \operatorname{im} A_1 \right\} = \left\{ \begin{pmatrix} u \\ v \end{pmatrix} : u \in S_1 \right\}, \\ \hat{Q}_1 &= \begin{bmatrix} Q_1 & 0 \\ \quad Q_1 & 0 \end{bmatrix}, \text{ where } \qquad := -{g_z'}^{-1} \quad Q_1, \end{aligned}$$

further

$$\hat{G}_2 = \begin{bmatrix} G_2 & 0 \\ \quad + g_x' P Q_1 & g_z' \end{bmatrix}, \quad \hat{Q}_1 \hat{G}_2^{-1} = \begin{bmatrix} Q_1 G_2^{-1} & 0 \\ \quad Q_1 G_2^{-1} & 0 \end{bmatrix}.$$

Now it becomes obvious that

$$\hat{Q}_1 \hat{G}_2^{-1} \{ \hat{f}(\hat{y}, \hat{x}, t) - \hat{f}(0, \hat{P}(t)\hat{x}, t) \} = 0$$

is valid if and only if it holds that

$$Q_1 G_2^{-1} \{ f(y, x, t) - f(0, P(t)x, t) \} = 0.$$

(ii) Letting $\hat{x} = \binom{x}{w}$ we proceed analogously as in the first part. Now we have

$$
\begin{aligned}
\hat{f}'_{\hat{x}'} &= \begin{bmatrix} f'_{x'} & 0 \\ h'_{x'} & h'_{w'} \end{bmatrix}, \quad \hat{f}'_{\hat{x}} = \begin{bmatrix} f'_{x} & 0 \\ h'_{x} & h'_{w} \end{bmatrix}, \quad \hat{Q} = \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix}, \\
\hat{A}_1 &= \begin{bmatrix} A_1 & 0 \\ & h'_{w'} \end{bmatrix}, \qquad := h'_{x'} + (h'_x + h'_{x'}P')Q, \\
\hat{S}_1 &= \left\{ \begin{pmatrix} u \\ v \end{pmatrix} : f'_x P u \in \operatorname{im} A_1 \right\} = \left\{ \begin{pmatrix} u \\ v \end{pmatrix} : u \in S_1 \right\}, \\
\hat{Q}_1 &= \begin{bmatrix} Q_1 & 0 \\ -h'_{w'}{}^{-1} \; Q_1 & 0 \end{bmatrix}, \\
\hat{G}_2 &= \begin{bmatrix} G_2 & 0 \\ & h'_{w'} \end{bmatrix}, \quad \hat{Q}_1 \hat{G}_2^{-1} = \begin{bmatrix} Q_1 G_2^{-1} & 0 \\ -h'_{w'}{}^{-1} \; Q_1 G_2^{-1} & 0 \end{bmatrix}.
\end{aligned}
$$

Again we obtain

$$\hat{Q}_1 \hat{G}_2^{-1} \{ \hat{f}(\hat{y}, \hat{x}, t) - \hat{f}(0, \hat{P}(t)\hat{x}, t) \} = 0$$

if and if $Q_1 G_2^{-1} \{ f(y, x, t) - f(0, P(t)x, t) \} = 0$. $\square$

Due to Lemma 4.2 we may add certain further index 1 and index 0 equations to the original index 2 DAE satisfying the structural condition (4.3). The resulting enlarged DAE has index 2, too, and fulfils condition (4.3) again.

Next we try to reformulate our DAE into an appropriate operator equation to make use of well-known standard arguments for those equations like the Implicit Function Theorem and the Newton-Kantorovich method.
Let $x_* \in C^1_N(J_*, I\!R^m)$ be fixed as before, $J_* \subseteq J$. Introduce the map

$$F : \quad (x_*, \varrho) \subset C^1_N \longrightarrow C$$

by means of

$$(Fx)(t) := f((Px)'(t) - P'(t)x(t), x(t), t), \quad t \in J_*, \quad x \in \quad (x_*, \varrho).$$

Thereby, (x_*, ϱ) denotes an open ball in C^1_N, and $\varrho > 0$ is assumed to be small enough to keep the trajectories of all functions $x \in (x_*, \varrho)$ within the region $\mathcal{D}$ we started with.
It is well-known (März 1986, cf. also § 2) that F is not Fredholm in the given natural topologies. However, Corollary 2.4 causes discussions whether we should turn from C, $\|\cdot\|_\infty$ to $C^1_{(2)}$, $\|\cdot\|_{(2)}$ also in the nonlinear case. If F is continuously differentiable and $F'(x_*)$ maps C^1_N surjectively onto $C^1_{(2)}$ in that new image space topology, the resulting linearized boundary value problem map $\mathcal{F}'(x_*)$is a homeomorphism. Then standard argument apply if F itself maps into $C^1_{(2)}$, too. Unfortunately, now the space $C^1_{(2)}$ may depend on x_*.
However the structural condition (4.3) proves its value once more in this context. Namely, supposed $\mathcal{N}$ is large enough (resp. $\varrho > 0$ is small enough) so that the graphs corresponding to $x \in (x_*, \varrho)$ proceed in $\mathcal{N}$, we have

$$
\begin{aligned}
&Q_1(t)G_2(t)^{-1} f((Px)'(t) - P'(t)x(t), x(t), t) = \\
&\quad = Q_1(t)G_2(t)^{-1} f(0, P(t)x(t), t) \quad \text{for } t \in J_*, \; x \in \quad (x_*, \varrho).
\end{aligned} \tag{4.7}
$$

Due to that property, we may easily realize that $x \in (x_*, \varrho)$ implies $Fx \in C^1_{(2)}$.

Lemma 4.3 *Let the linearized at* $x_* \in C^1_N(J_*, \mathbb{R}^m)$ *DAE be index-2 tractable,* $J_* \subset J$ *be compact, and* (4.3) *be valid.*
Let $\tilde{g}(x,t) := Q_1(t)G_2(t)^{-1}f(0, P(t)x, t)$, $x \in \mathcal{D}$, $t \in J_*$, *as well as its partial derivatives* $\tilde{g}'_t(x,t)$, $\tilde{g}'_x(x,t)$, $\tilde{g}''_{xx}(x,t)$, $\tilde{g}''_{xt}(x,t)$ *depend continuously on* (x,t).
Then, F *maps* $(x_*, \varrho) \subset C^1_N(J_*, \mathbb{R}^m)$ *into* $C^1_{(2)}(J_*, \mathbb{R}^m)$, *and it is continuously (Fréchet) differentiable.*

Proof: First of all it is worth mentioning that

$$Q_1(t) = Q_1(t)G_2(t)^{-1}B(t) = \tilde{g}'_x(P(t)x_*(t), t)$$

depends continuously differentiably on t.
For given $x \in (x_*, \varrho)$, we obtain $Fx \in C^1_{(2)}$ immediately by means of relation (4.7). It remains to check the differentiability. As usually, we calculate for $x \in (x_*, \varrho)$, $z \in C^1_N$

$$\lim_{\tau \to 0} \frac{1}{\tau}(F(x + \tau z) - F(x)) =: F'(x)z$$

and show $F'(x) : C^1_N \to C^1_{(2)}$ to be a linear, bounded map. After that we prove $F'(x)$ to depend continuously on x.
First of all we compute

$$F'(x)z = A_x(Pz)' + (B_x - A_xP')z,$$

where $A_x(t) := f'_{x'}((Px)'(t) - P'(t)x(t), x(t), t)$, $B_x(t) := f'_x((Px)'(t) - P'(t)x(t), x(t), t)$, $t \in J_*$.
The inequality

$$\|F'(x)z\|_\infty \le K_1(x)\|z\|$$

results immediately with $K_1(x) \in \mathbb{R}^+$.
Further, (4.3) yields for all $(\bar{y}, \bar{x}, t) \in \mathcal{N}$

$$\begin{aligned} Q_1(t)G_2(t)^{-1}f'_{x'}(\bar{y}, \bar{x}, t) &= 0 \\ Q_1(t)G_2(t)^{-1}f'_x(\bar{y}, \bar{x}, t) &= Q_1(t)G_2(t)^{-1}f'_x(0, P(t)\bar{x}, t)P(t). \end{aligned}$$

Therefore,

$$Q_1(t)G_2(t)^{-1}(F'(x)z)(t) = \tilde{g}'_x(P(t)x, t)P(t)z(t), \quad t \in J_*,$$

represents a C^1 function with respect to t. Consequently, $F'(x)z \in C^1_{(2)}$ is actually true.
Now, the inequality

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}Q_1(t)G_2(t)^{-1}(F'(x)z)(t)\right| \le K_2(x)\|z\|, \quad t \in J_*,$$

follows, thus $\|F'(x)z\|_{(2)} \le K_3(x)\|z\|$, i.e. $F'(x) : C^1_N \to C^1_{(2)}$ is a bounded linear map as expected.
It remains to show that $F'(x)$ is continuous with x.
For $x, \tilde{x} \in (x_*, \varrho)$, $z \in C^1_N$, we derive on the one hand

$$\|(F'(x) - F'(\tilde{x}))z\|_\infty \le \max_{t \in J_*}\{|A_x(t) - A_{\tilde{x}}(t)|\,|I + P'(t)| + |B_x(t) - B_{\tilde{x}}(t)|\}\|z\|.$$

On the other hand, we also compute

$$\Big|\frac{\mathrm{d}}{\mathrm{d}t}Q_1(t)G_2(t)^{-1}((F'(x)-F'(\tilde{x}))z)(t)\Big| \leq$$
$$\leq \max_{t\in J_*}\{|\tilde{g}'_x(P(t)x(t),t)-\tilde{g}'_x(P(t)\tilde{x}(t),t)|+|\tilde{g}''_{xt}(P(t)x(t),t)-\tilde{g}''_{xt}(P(t)\tilde{x}(t),t)|+$$
$$+|\tilde{g}''_{xx}(P(t)x(t),t)(Px)'(t)-\tilde{g}''_{xx}(P(t)\tilde{x}(t),t)(P\tilde{x})'(t)|\}\|z\|$$

This shows that $\|\tilde{x}-x\|\to 0$ implies $\|F'(\tilde{x})-F'(x)\|_{(2)}\to 0$, in fact. □

Now we are ready to state solvability of perturbed nonlinear index 2 IVPs locally around a given solution x_*. Intending e.g. to approximate x_* by numerical integration, we should be aware even of those solvability results. Roughly speaking, the next theorem says how to catch, locally around x_*, the implicitly given but practically unknown ("hidden") second order constraint or state manifold, where the neighbouring solutions proceed. It also describes in which sense the solutions depend on the perturbations. Notice also that Theorem 4.4 represents the nonlinear version of Theorem 2.2(iii).

Theorem 4.4 *Given a solution $x_*\in C^1_N(J_*,\mathbb{R}^m)$ of the DAE (4.1), $J_*\subseteq J$ compact. Let the linearized at x_* DAE (1.6) be index-2 tractable. Let (4.3) be valid. Moreover, let $\tilde{g}$ be continuous together with its partial derivatives $\tilde{g}'_x$, $\tilde{g}'_t$, $\tilde{g}''_{xx}$, $\tilde{g}''_{xt}$.*
Then, for $t_0\in J_*$ and sufficiently small $\sigma>0$, $\tau>0$, the perturbed initial value problem

$$\begin{aligned}
&f(x'(t),x(t),t)=q(t), \quad t\in J_*,\\
&P(t_0)P_1(t_0)(x(t_0)-x^0)=0, \quad x^0\in\mathbb{R}^m,\\
&|P(t_0)P_1(t_0)(x^0-x(t_0))|\leq\tau,\\
&q\in C^1_{(2)}(J_*,\mathbb{R}^m), \quad \|q\|_\infty+\|(Q_1G_2^{-1}q)'\|_\infty\leq\sigma,
\end{aligned}$$

is uniquely solvable on $C^1_N(J_*,\mathbb{R}^m)$. Its solution $x(\cdot,x^0,q)$ depends continuously on $(x^0,q)\in\mathbb{R}^m\times C^1_{(2)}$, where $C^1_{(2)}$ is equipped with $\|\cdot\|_{(2)}$.

Proof: Define $\mathcal{F}x:=(Fx,P(t_0)P_1(t_0)(x(t_0)-x_*(t_0)))$, $x\in\mathcal{B}(x_*,\varrho)$. $\mathcal{F}$ maps (x_*,ϱ) into $C^1_{(2)}\times M_{(2)}$, where $M_{(2)}:=\operatorname{im}P(t_0)P_1(t_0)$.
Obviously, $\mathcal{F}$ is as smooth as F. By construction, we have $\mathcal{F}x_*=0$. Due to Corollary 2.4, $\mathcal{F}'(x_*)$ is a homeomorphism from C^1_N onto $C^1_{(2)}\times M_{(2)}$. Hence, our assertion is a direct consequence of the Implicit Function Theorem (e.g. Krasnosel'ski et al. (1969), p. 23) □

Remarks:

1. In particular, the inequality

$$\|x(\cdot, x^0, q) - x_*\|_{C_N^1} \leq K\{|P(t_0)P_1(t_0)(x^0 - x_*(t_0))| + \|q\|_\infty + \|(Q_1 G_2^{-1} q)'\|_\infty\}$$

results by Theorem 4.4, what shows the so-called perturbation index also to be 2 (cf. Hairer and Wanner (1991)).

2. The solution $x(t, x^0)$ of the initial value problem $f(x'(t), x(t), t) = 0$, $P(t_0)P_1(t_0)(x(t_0) - x^0) = 0$, $|P(t_0)P_1(t_0)(x_*(t_0) - x^0)| \leq \tau$, depends continuously differentiably on x^0, but the partial derivative $X(t) := \dfrac{\partial}{\partial x^0} x(t, x^0)$ satisfies the first variation equation

$$A_x(PX)' + (B_x - A_x P')X = 0, \qquad P(t_0)P_1(t_0)(X(t_0) - I) = 0.$$

3. A similar result may be obtained for parameter dependent DAEs $f(x'(t), x(t), t, p) = 0$, where x_* solves this equation for a given parameter value p_*, and the structural condition (4.3) is satisfied uniformly for all parameter values to be considered.

4. Recall once more that the relations $x(t_0, x^0) = x^0$ or $P(t_0)x(t_0, x^0) = P(t_0)x^0$ cannot be expected at all. Even if x^0 was close to $x_*(t_0)$ but $x^0 \notin \operatorname{im} P(t_0)P_1(t_0)$, the initial condition $x(t_0) = x^0$ would yield a non-solvable initial value problem. If x^0 is a consistent initial value, the solution exists a priori. However, there is no easy way for describing the manifold of consistent initial values at all.

5. For the special case of quasilinear DAEs the assertion given by Theorem 4.4 is also proved in März and Tischendorf (1994), where the nonlinear equation itself is decoupled via its linear part. Moreover, the backward differentiation formula is shown to work well in this context so that we are able to integrate those IVPs numerically.

Next we turn to BVPs. What about the Newton-Kantorovich method applied to index 2 DAEs? Recall that, as usually, neither the initial guess $x_0 \in C_N^1$ nor the approximations provided are expected to satisfy the original DAE (4.1) itself.
In the following, we will show that quasilinearization should work well also in the index 2 case, supposed the structural condition (4.3) is valid. In this context, denote again by

$$S := r_1'(x_*(t_0), x_*(T))X(t_0) + r_2'(x_*(t_0), x_*(T))X(T), \tag{4.8}$$

the solvability matrix of the boundary value problem (1.1), (1.9), where now the fundamental solution matrix X is uniquely determined by

$$A(PX)' + (B - AP')X = 0, \qquad P(t_0)P_1(t_0)(X(t_0) - I) = 0.$$

Theorem 4.5 *Let all the assumptions of Theorem 4.4 be satisfied, and let x_* solve the BVP (4.1), (1.9), $J_* = [t_0, T]$. Let the boundary condition (1.9) be stated properly, i.e.*

$$\ker S = \ker P(t_0)P_1(t_0), \tag{4.9}$$

$$\operatorname{im} S = \operatorname{im}(r_1'(x_*(t_0), x_*(T)), r_2'(x_*(t_0), x_*(T))). \tag{4.10}$$

Then, for sufficiently accurate initial guess $x_0 \in C_N^1$, $\|x_0 - x_*\|$ small enough, the Newton-Kantorovich method (1.10) – (1.12) provides a uniquely determined sequence $\{x_j\}_{j\geq 0} \subset \quad (x_*, \varrho)$, and

$$\|x_j - x_*\| \longrightarrow 0 \qquad (j \to \infty).$$

Proof: Form the boundary value problem map
$\tilde{\mathcal{F}}: \quad (x_*, \varrho) \to C^1_{(2)} \times \tilde{M}_{(2)}$ by $\tilde{\mathcal{F}}x := (Fx, r(x(t_0), x(T)))$, $x \in \quad (x_0, \varrho)$, $\tilde{M_{(2)}} := \mathrm{im}(D_1, D_2)$, $D_i := r'_i(x_*(t_0), x_*(T))$, $i = 1, 2$. Again, $\tilde{\mathcal{F}}$ is as smooth as F.
Further, it holds that $\tilde{\mathcal{F}}x_* = 0$, and $\tilde{\mathcal{F}}'(x_*)z = (q, d)$ represents the linearized at x_* BVP

$$A(Pz)' + (B - AP')z = q, \qquad D_1 z(t_0) + D_2 z(T) = d, \tag{4.11}$$

which is uniquely solvable for each $q \in C^1_{(2)}$, $d \in \mathrm{im}(D_1, D_2)$ (cf. März (1992)).
Again, Corollary 2.3 implies $\tilde{\mathcal{F}}'(x_*)$ to be a homeomorphism from C^1_N onto $C^1_{(2)} \times \tilde{M}_{(2)}$. Now, applying standard arguments on the Newton-Kantorovich method (Krasnosel'skij et al. (1969), § 11) we are done. □

Remark: Supposed all involved partial derivatives of f and $\tilde{g}$ are Lipschitzian, the resulting approximations x_j converge quadratically to x_*.

Stress again that the initial guess x_0 may be chosen not to satisfy the first and second order constraint. In particular, in case of the Hessenberg system (4.5) there is no need for x_0 to satisfy the second equation $\psi(x_1, t) = 0$, but also the hidden constraint $\psi'_{x_1}(x_1, t)\varphi(x_1, x_2, t) - \psi'_t(x_1, t) = 0$.

5 Quasilinear index 1 DAEs whose leading nullspace varies with x and t

In this section we specify results of § 4 for enlarged systems resulting from index 1 DAEs whose leading nullspace depends also on x. Note that the index 1 theory described in § 3 (cf. Griepentrog and März (1986)) does not apply to those equations since the nullspace condition (1.2) is no more valid. Consider the following equation

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

where the leading Jacobian $A(x, t)$ has constant rank $r < m$. Form the basic subspaces to be

$$\begin{aligned} N(x,t) &:= \{z \in \mathbb{R}^m : A(x,t)z = 0\}, \\ S(y,x,t) &:= \{z \in \mathbb{R}^m : B(y,x,t)z \in \mathrm{im}\, A(x,t)\}, \\ B(y,x,t) &:= g'_x(x,t) + A'_x(x,t)y, \quad (y,x,t) \in \mathbb{R}^m \times \mathcal{D} \times J. \end{aligned}$$

Definition: Equation (5.1) is said to be an index 1 DAE on the open set $\mathcal{G} \subseteq \mathbb{R}^m \times \mathcal{D} \times J$ if the condition

$$N(x,t) \cap S(y,x,t) = \{0\}, \quad (y,x,t) \in \mathcal{G},$$

is fulfilled.

Since the previous index 1 results apply only for those DAEs having $N(x,t)$ invariant of x, we turn to the enlarged system

$$\left.\begin{aligned} &x'(t) - y(t) = 0 \\ &A(x(t),t)y(t) + g(x(t),t) = 0 \end{aligned}\right\}, \tag{5.2}$$

which has a constant leading nullspace. But the price for the nicer form of (5.2) is a higher index, as the next lemma will show.

Lemma 5.1 *The enlarged system* (5.2) *is index-2 tractable if and only if* (5.1) *itself has index* 1.

Proof: Put the enlarged system (5.2) into the form

$$\hat{f}(\hat{x}'(t), \hat{x}(t), t) = 0, \quad \hat{x} := \begin{pmatrix} x \\ y \end{pmatrix}.$$

Compute the partial Jacobians

$$\hat{f}_{\hat{x}'}(\hat{y}, \hat{x}, t) = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \quad \hat{f}_{\hat{x}}(\hat{y}, \hat{x}, t) = \begin{bmatrix} 0 & -I \\ B(y, x, t) & A(x, t) \end{bmatrix},$$

further

$$\hat{Q} = \begin{bmatrix} 0 & \\ & I \end{bmatrix}, \quad \hat{A}_1(\hat{y}, \hat{x}, t) = \begin{bmatrix} I & -I \\ 0 & A(x, t) \end{bmatrix}.$$

Obviously, $\hat{A}_1(\hat{y}, \hat{x}, t)$ is singular since $A(x, t)$ is so. Moreover, we have

$$\begin{aligned} \hat{N}_1(\hat{y}, \hat{x}, t) &= \{z \in I\!R^m \times I\!R^m : z_1 = z_2,\ A(x, t)z_2 = 0\}, \\ \hat{S}_1(\hat{y}, \hat{x}, t) &= \{z \in I\!R^m \times I\!R^m : B(y, x, t)z_1 \in \operatorname{im} A(x, t)\}. \end{aligned}$$

Evidently, $z \in \hat{N}_1(\hat{y}, \hat{x}, t) \cap \hat{S}_1(\hat{y}, \hat{x}, t)$ is equivalent with $z_1 = z_2 \in N(x, t) \cap S(y, x, t)$. □

It was Lubich (1989) who discovered that a differentiation index 1 DAE the leading nullspace of which rotates with varying x behaves rather than an index 2 DAE. This was one of the reasons for introducing the perturbation index. Recall Lubich's example of a differentiation index 1 but perturbation index 2 problem in more detail.

Example: $m = 3$,

$$\left.\begin{aligned} x_1' - x_3 x_2' + x_2 x_3' - g_1 &= 0 \\ x_2 - g_2 &= 0 \\ x_3 - g_3 &= 0 \end{aligned}\right\}. \tag{5.3}$$

We are interested in solving the IVP for (5.3) with the initial condition $x_1(0) = 0$ on $[0, 2\pi]$, letting

$$g_1(t) = 0, \quad g_2(t) = \frac{1}{n}\sin n^k t, \quad g_3(t) = \frac{1}{n}\cos n^k t, \quad n, k \in I\!N.$$

The solution is

$$x_1(t) = n^{k-2}t, \quad x_2(t) = g_2(t), \quad x_3(t) = g_3(t).$$

For fixed $k \geq 3$, we have $\|g\|_\infty \to 0$ $(n \to \infty)$, but $\|x\|_\infty \to \infty$ $(n \to \infty)$, although the solution of (5.3) vanishes identically if g does so. Obviously, this behaviour confirms once again the understanding to consider this problem rather as a higher index one.
Linearizing (5.3) at a fixed $x_* \in C^1$ gives the coefficients for (1.6)

$$A(t) = \begin{pmatrix} 1 & -x_{*3}(t) & x_{*2}(t) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad B(t) = \begin{pmatrix} 0 & x_{*3}'(t) & -x_{*2}'(t) \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which clearly form an index-1 tractable equation (cf. § 2). A possible nullspace projector is

$$Q(t) = \begin{pmatrix} 0 & x_{*3}(t) & -x_{*2}(t) \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad \square$$

Lemma 5.2 *Let $Q(y,x,t)$ denote the projector onto $N(x,t)$ along $S(y,x,t)$. Then*

$$\hat{Q}_1(\hat{x},t) := \begin{bmatrix} Q(y,x,t) & 0 \\ Q(y,x,t) & 0 \end{bmatrix}$$

represents the projector onto $\hat{N}_1(\hat{y},\hat{x},t)$ along $\hat{S}_1(\hat{y},\hat{x},t)$, which is invariant of $\hat{y}$.

Proof: For $z \in \hat{N}_1(\hat{y},\hat{x},t)$, it holds that $z_1 = z_2$, $z_2 = Q(y,x,t)z_2$, further

$$\hat{Q}_1(\hat{x},t)z = \begin{bmatrix} Q(y,x,t)z_1 \\ Q(y,x,t)z_1 \end{bmatrix} = z.$$

On the other hand, $z \in \hat{S}_1(\hat{y},\hat{x},t)$ implies $z_1 \in S(y,x,t)$, that is $Q(y,x,t)z_1 = 0$, thus $\hat{Q}_1(\hat{x},t)z = 0$. $\square$

Next, for given $x_* \in C^1(J_*, I\!R^m)$, $y_* := x_*'$ we may consider both the linearization of (5.1) at x_* and that of the enlarged system (5.2) at $(x_*, y_*) \in C^1_{\hat{N}} := \{(x,y) : x \in C^1, y \in C\}$. The first linearization leads to

$$A(t)z'(t) + B(t)z(t) = q(t) \tag{5.4}$$

with

$$A(t) := A(x_*(t),t), \quad B(t) := B(y_*(t), x_*(t), t), \quad t \in J_*,$$

but linearizing (5.2) yields

$$\begin{aligned} &\hat{A}(t)\hat{z}'(t) + \hat{B}(t)\hat{z}(t) = \hat{q}(t), \\ &\hat{A}(t) := \begin{bmatrix} I & \\ & 0 \end{bmatrix}, \quad \hat{B}(t) := \begin{bmatrix} 0 & -I \\ B(t) & A(t) \end{bmatrix}. \end{aligned} \tag{5.5}$$

Clearly, (5.5) represents the enlarged system of the linear equation (5.4) simultaneously. In this sense, enlarging the system and linearizing commute.
Due to Lemma 5.1, the DAE (5.4) has index 1 if and only if (5.5) is an index 2 DAE.

Lemma 5.3 *Let A and g belong to the class C^1 and g have continuous partial derivatives g''_{xx}, g''_{xt}. Let* im $A(x,t)$ *be invariant of x, i.e.*

$$\operatorname{im} A(x,t) = R(t), \quad x \in \mathcal{D},\ t \in J. \tag{5.6}$$

Moreover, let the linearized at $x_* \in C^1$ equation (5.4) be index 1.
Then, the DAE (5.5) is index 2 tractable with a C^1 projector function $\hat{Q}_1$. Further, the structural condition (4.3) is valid for the enlarged system (5.2).

Proof: $A(t) = A(x_*(t),t)$ depends continuously differentiably on t and has constant rank r. Consequently, the orthoprojector $Q^\perp(t)$ onto $N(t) := \ker A(t)$ is also continuously differentiable. Moreover, condition (5.6) implies $\operatorname{im}(A'_x(x,t)y) \subset R(t)$ which simplifies, in its turn, the subspace $S(t) := S(y_*(t), x_*(t), t)$ to $S(t) = \{z \in I\!R^m : g'_x(x_*(t),t)z \in \operatorname{im} A(t)\}$. Next,

$$Q^\perp(t)(A(t) + g'_x(x_*(t),t)Q^\perp(t))^{-1} g'_x(x_*(t),t) =: Q(t) \tag{5.7}$$

represents the canonical projector onto $N(t)$ along $S(t)$. Due to our smoothness assumptions, Q becomes a C^1 projector function. Applying Lemma 5.2 to the linear DAEs (5.4) and (5.5), we arrive at the C^1 projector function

$$\hat{Q}_1(t) = \left[\begin{array}{cc} Q(t) & 0 \\ Q(t) & 0 \end{array} \right]$$

for (5.5).
It remains to check the condition (4.3) to be fulfilled for (5.2). For that, compute

$$\hat{G}_2(t) = \left[\begin{array}{cc} I & -I \\ B(t)Q(t) & A(t) \end{array} \right], \quad \hat{G}_2(t)^{-1} = \left[\begin{array}{cc} P(t) & G_1(t)^{-1} \\ -Q(t) & G_1(t)^{-1} \end{array} \right],$$

$$\hat{Q}_1(t)\hat{G}_2(t)^{-1} = \left[\begin{array}{cc} 0 & Q(t)G_1(t)^{-1} \\ 0 & Q(t)G_1(t)^{-1} \end{array} \right],$$

and then, for $\hat{y} \in I\!R^m \times I\!R^m$, $\hat{x} \in \mathcal{D} \times I\!R^m$, $t \in J_*$,

$$\hat{Q}_1(t)\hat{G}_2(t)^{-1}\{\hat{f}(\hat{y}, \hat{x}, t) - \hat{f}(0, \hat{P}\hat{x}, t)\} = \left(\begin{array}{c} Q(t)G_1(t)^{-1} A(x,t)y) \\ Q(t)G_1(t)^{-1} A(x,t)y) \end{array} \right) = 0.$$

Thereby, $G_1 := A+BQ$ remains nonsingular, and the last relation is a consequence of assumption (5.6) and the property $QG_1^{-1}A = 0$. □

Now we are well-prepared to specify Theorem 4.4 for (5.2).

Theorem 5.4 *Given a solution $x_* \in C^1(J_*, I\!R^m)$ of (5.1), $J_* \subseteq J$ compact, and let all assumptions of Lemma 5.3 be satisfied. Let $Q(t)$ denote the projector onto $N(t) := \ker A(x_*(t),t)$ along the subspace*

$$\begin{aligned} S(t) &:= \{z \in I\!R^m : g_x'(x_*(t),t)z \in R(t)\}, \\ \tilde{G}_1(t) &:= A(t) + g_x'(x_*(t),t)Q(t). \end{aligned}$$

(i) *Then, for $t_0 \in J_*$ and sufficiently small $\sigma > 0$, $\tau > 0$, the IVP*

$$\begin{aligned} & A(x(t),t)x'(t) + g(x(t),t) = q(t), \quad t \in J_*, \\ & P(t_0)(x(t_0) - x^0) = 0, \quad x^0 \in I\!R^m, \\ & |P(t_0)(x^0 - x_*(t_0))| \le \tau, \\ & q \in C, \;\; Q\tilde{G}_1^{-1}q \in C^1, \;\; \|q\|_\infty + \|(Q\tilde{G}_1^{-1}q)'\|_\infty \le \sigma \end{aligned}$$

is uniquely solvable on $C^1(J_*, I\!R^m)$.

(ii) *The IVP solution depends continuously differentiably on x^0.*

(iii) *The IVP solution satisfies the inequality*

$$\|x(\cdot, x^0, q) - x_*\|_{C^1} \le K\{|P(t_0)(x^0 - x_*(t_0))| + \|q\|_\infty + \|(Q\tilde{G}_1^{-1}q)'\|_\infty\}.$$

Proof: First of all, with the denotations used when proving Lemma 5.3, we have $QG_1^{-1} = Q\tilde{G}_1^{-1}$. Next we turn to the enlarged form of the DAE to be considered in (i), that is,

$$\left. \begin{array}{l} x'(t) - y(t) = 0 \\ A(x(t),t)y(t) + g(x(t),t) = q(t) \end{array} \right\}. \tag{5.8}$$

Due to Lemma 5.3, Theorem 4.4 applies immediately to that system. It holds that

$$\hat{Q}_1\hat{G}_2^{-1}\begin{pmatrix}0\\q\end{pmatrix} = \begin{bmatrix} 0 & QG_1^{-1} \\ 0 & QG_1^{-1} \end{bmatrix}\begin{pmatrix}0\\q\end{pmatrix} = \begin{pmatrix}Q\tilde{G}_1^{-1}q\\Q\tilde{G}_1^{-1}q\end{pmatrix}$$

and

$$\hat{P}\hat{P}_1(t_0) = \begin{bmatrix} P(t_0) & 0 \\ 0 & 0 \end{bmatrix},$$

thus

$$\hat{P}\hat{P}_1(t_0)(\hat{x}(t_0) - \hat{x}^0) = \begin{pmatrix} P(t_0)(x(t_0) - x^0) \\ 0 \end{pmatrix}. \qquad \square$$

Remarks:

1. It should be emphasized once more that now $P(t_0)$ may depend on $x_*(t_0)$.
2. Theorem 5.4(iii) says that the perturbation index of (5.1) is not greater than two. As Lubich (1989) has shown, a solution-depending leading nullspace may force the perturbation index to become two in fact. On the other hand, comparing with standard results, which apply in case the nullspace condition (1.2) holds true, we know the perturbation index to be one then.
3. Supposed $\ker A(x,t)$ does not vary with x, i.e. condition (1.2) is valid, we can do with lower smoothness to obtain solvability on the function space C_N^1. In particular, we can do without demanding $QG_1^{-1}q \in C^1$, but $q \in C$ will suffice. However, Theorem 5.4 provides C^1 solutions. For that, the additional smoothness, e.g. $QG_1^{-1}q \in C^1$, is necessary.
4. The partial derivative $X(t) := \frac{\partial}{\partial x^0}x(t,x^0)$ satisfies the first variation equation
$$\begin{aligned} A(x(t,x^0),t)X'(t) + B(x'(t,x^0),x(t,x^0),t)X(t) = 0 \\ P(t_0)(X(t_0) - I) = 0. \end{aligned}$$
5. Condition (5.5) is not even restrictive. It may be achieved by corresponding scalings.

Now, let us specify the Newton-Kantorovich method for the boundary value problem

$$\left.\begin{aligned} &A(x(t),t)x'(t) + g(x(t),t) = 0 \\ &r(x(t_0),x(T)) = 0 \end{aligned}\right\}. \tag{5.9}$$

Given an initial guess $x_0 \in C^1([t_0,T],I\!R^m)$ we put $y_0 := x_0'$ and apply method (1.10) – (1.12) to the enlarged system (5.2). This yields $y_j = x_j'$ for all $j \geq 1$ so that we are able to describe the whole iteration process in terms of the original system as follows.
For $j \geq 0$, we solve the linear BVP

$$\left.\begin{aligned} A(x_j(t),t)z_{j+1}'(t) + B(x_j'(t),x_j(t)t)z_{j+1}(t) + \\ +A(x_j(t),t)x_j'(t) + g(x_j(t),t) = 0 \\ r_1'(x_j(t_0),x_j(T))z_{j+1}(t_0) + r_2'(x_j(t_0),x_j(T))z_{j+1}(T) = \\ = -r(x_j(t_0),x_j(T)) \end{aligned}\right\} \tag{5.10}$$

and put $x_{j+1} = x_j + z_{j+1}$ then.
Obviously, (5.10) looks like the iteration directly applied to (5.9). What is only left to do is checking the unique solvability of the linear BVPs to be solved.
Given a BVP solution $x_* \in C^1([t_0,T], \mathbb{R}^m)$ we define the matrix

$$S := r_1'(x_*(t_0), x_*(T))X(t_0) + r_2'(x_*(t_0), x_*(T))X(T)$$

as usually, where the fundamental solution matrix X is given by

$$AX' + BX = 0, \quad P(t_0)(X(t_0) - I) = 0.$$

Comparing with (4.9), (4.10) applied to the enlarged system and taking into account the representation $\hat{P}\hat{P}_1(t_0) = \mathrm{diag}(P(t_0), 0)$, we derive the conditions

$$\left.\begin{array}{rcl} \ker S &=& \ker P(t_0) \\ \operatorname{im} S &=& \operatorname{im}(r_1'(x_*(t_0), x_*(T)), r_2'(x_*(t_0), x_*(T))) \end{array}\right\} \tag{5.11}$$

to be the relevant ones for the proper statement of the boundary conditions. Emphasize that, formally, (5.11) are the same conditions as those used in Theorem 3.1. But now $\ker P(t_0) = N(x_*(t_0), t_0)$ may depend on the solution.
On this background, Theorem 4.5 applied to the BVP (5.9) simplifies as given below.

Theorem 5.5 *Given a solution x_* of the BVP* (5.9)*, and let the boundary conditions be stated properly, i.e.* (5.11) *be fulfilled. Let the assumptions of Lemma* 5.3 *be satisfied.*
Then, for any sufficiently good initial guess $x_0 \in C^1([t_0,T], \mathbb{R}^m)$, $\|x_0 - x_*\|_{C^1}$ small enough, the Newton-Kantorovich process (5.10) *provides uniquely determined x_j, $j \geq 0$, and $\|x_j - x_*\|_{C^1} \to 0$ $(j \to \infty)$.*

Note again, all linear BVPs to be solved for z_j are uniquely solvable. But now, all of them have perturbation index one. This fact can be considered as a further advantage of the linearization.

The statements given in the present section for equation (5.1) may be immediately generalized for fully implicit DAEs (1.1), supposed the partial Jacobian $f'_{x'}(y,x,t)$ has constant rank and its range is invariant of (y,x), i.e. $\operatorname{im} f'_{x'}(y,x,t) = R(t)$, $(y,x,t) \in \mathbb{R}^m \times \mathcal{D} \times J$. But additionally, in this case we have either to assume the resulting projector function $\hat{Q}_1$ (cf. Lemma 5.3) to belong to the class C^1 or to require the respectively higher smoothness of the two functions f and x_* for obtaining that property via the lines of Lemma 5.3.

Finally, return briefly to the example of Lubich mentioned above. Complete the DAE (5.3) by the boundary resp. initial condition with $r(z_1, z_2) := z_{1,1}$, $z_1, z_2 \in \mathbb{R}^m$, such that $r_1'(z_1,z_2) \equiv \mathrm{diag}(1,0,0)$, $r_2'(z_1,z_2) \equiv 0$. Denote by $x_* \in C^1$ the solution of the resulting IVP, that is

$$\begin{array}{rcl} x_{*i} &=& g_i, \quad i = 2,3, \\ x_{*1}' &=& g_1 + g_3 g_2' - g_2 g_3', \quad x_{*1}(0) = 0. \end{array}$$

The solution matrix

$$S = \begin{bmatrix} 1 & & \\ & 0 & \\ & & 0 \end{bmatrix} \begin{bmatrix} 1 & -x_{*3}(0) & x_{*2}(0) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = P(0)$$

satisfies the conditions (5.11) trivially. Also Lemma 5.3 applies immediately. Compute further

$$\tilde{G}_1 = \begin{pmatrix} 1 & -x_{*3} & x_{*2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad Q\tilde{G}_1^{-1} = \begin{pmatrix} 0 & x_{*3} & -x_{*2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Moreover, the Newton-Kantorovich method (5.10) yields exact second and third components $x_{j+1,i} = g_i$, $i = 2, 3$, after the first iteration, i.e. for $j \geq 0$, independently of the choice of the initial guess x_0. After the second iteration step, i.e. for $j \geq 1$, also the first component becomes exact, i.e. $x'_{j+1,1} = g_1 + g_3 g'_2 - g_2 g'_3$, $x_{j+1,1}(0) = 0$ is satisfied. Hence, due to the very simple structure of this special example the exact IVP solution is obtained after the second iteration step, independently of the chosen initial guess we started with.

References

R. März: On correctness and numerical treatment of boundary value problems in differential-algebraic equations. Z. vych. matem. i matem. fiziki 26 (1986) 1, 50-64.

R. März and C. Tischendorf: Solving more general index-2 differential algebraic equations. Computers and Mathematics with Applications, 28 (1994) 10-12, 77-105.

E. Hairer and G. Wanner: Solving ordinary differential equations II. Springer 1991.

E. Griepentrog and R. März: Differential-algebraic equations and their numerical treatment. Teubner Texte zur Mathematik 88, Leipzig 1986.

E. Griepentrog and R. März: Basic properties of some differential-algebraic equations. Z. für Anal. u. ihre Anwendungen 8 (1989), 25-40.

Ch. Lubich: Linearly implicit extrapolation methods for differential-algebraic systems. Numer. Math. 55 (1989), 197-211.

M.A. Krasnosel'skij et al.: Priblizhennoe reshenie operatornykh uravnenij. Nauka, Moskva 1969.

R. März: Boundary value problems in differential-algebraic equations and their numerical treatment. Humboldt-Univ. Berlin, Sektion Mathematik, Seminarbericht 55 (1984), 115-150.

L.V. Kantorovich and G.P. Akilov: Funktsional'nyj analiz. Nauka, Moskva 1977.

F.L. Lewis: A survey of linear singular systems. Circuits Systems Signal Process 5 (1986) 1, 3-36.

R. März: Canonical projectors for linear differential algebraic equations. Humboldt-Univ. Berlin, Fachbereich Mathem. Preprint 93-17, 1993.

C. Tischendorf: On stability of solutions of autonomous index-1 tractable and quasilinear index-2 tractable dae s. Circuits Systems Signal Process 13 (1994), 139-154.

R. März: On quasilinear index 2 differential algebraic equations. Humboldt-Univ. Berlin, Fachbereich Mathematik, Seminarbericht 92-1 (1992), 39-60.

S.M. Roberts and J.S. Shipman: Two point boundary value problems: Shooting methods. American Elsevier Publ., New York 1972.

Z. Aktas and H.J. Stetter: A classification and survey of numerical methods for boundary value problems in ordinary differential equations. Intern. J. for Num. Methods in Engineering 11 (1977), 771-796.

A. Miele and R.R. Iyer: Modified quasilinearization method for solving nonlinear two-point boundary value problems. J. Mathem. Anal. and Appl. 36 (1971), 674-692.

R. März: Index-2 differential-algebraic equations. Results in Mathematics 15 (1989), 149-171.