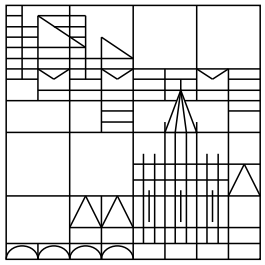

Universität Konstanz

One– and Multistep Discretizations of Index 2 Differential Algebraic Systems and their use in Optimization

Johannes Schropp

Konstanzer Schriften in Mathematik und Informatik
Nr. 148, Juni 2001

ISSN 1430–3558

Konstanzer Online-Publikations-System (KOPS)
URL: http://www.ub.uni-konstanz.de/kops/volltexte/2007/2174/
URN: http://nbn-resolving.de/urn:nbn:de:bsz:352-opus-21741

© Fachbereich Mathematik und Statistik
© Fachbereich Informatik und Informationswissenschaft
Universität Konstanz
Fach D 188, 78457 Konstanz, Germany
Email: preprints@informatik.uni–konstanz.de
WWW: http://www.fmi.uni–konstanz.de/Schriften

One- and Multistep Discretizations of Index 2 Differential Algebraic Systems and their use in Optimization

Johannes Schropp
Dept. of Mathematics and Computer Science
University of Konstanz
P.O. Box 5560
D-78434 Konstanz

Abstract

An approach to solve constrained minimization problems is to integrate a corresponding index 2 differential algebraic equation (DAE). Here corresponding means that the ω-limit sets of the DAE dynamics are local solutions of the minimization problem. In order to obtain an efficient optimization code we analyse the behavior of certain Runge-Kutta and linear multistep discretizations applied to these DAEs. It is shown that the discrete dynamics reproduces the geometric properties and the long time behavior of the continuous system correctly. Finally, we compare the DAE approach with a classical SQP-method.

1 Introduction

Differential algebraic problems of index 2 frequently arise when modelling phenomena from scientific computations. An important class for such problems is, e.g, multibody systems with constraints on the velocity level or in the Gupta, Gear, Leimkuhler [6] formulation. Due to Schropp [14] they also occur as auxiliary systems for minimization problems when searching for an evolution that approaches a local minimum of an objective function restricted by algebraic constraints.
To be more precise, Schropp [14] has been shown that the local minima of a smooth function $\hat{f} : D \to \mathbb{R}$, $D \subset \mathbb{R}^N$ open with respect to the constraints $\hat{g}(x) = 0$, $\hat{k}(x) \geq 0$, $\hat{g} : D \to \mathbb{R}^l$, $\hat{k} : D \to \mathbb{R}^q$ are computable in an indirect way as ω-limit sets of trajectories of a family of DAE's. We introduce the so called slack variables $y = (y_1, \ldots, y_q) \in \mathbb{R}^q$, define the functions

$$\bar{f}(x,y) \;:=\; \hat{f}(x), \quad \bar{g}(x,y) \;:=\; \begin{pmatrix} \hat{g}(x) \\ \hat{k}(x) - \operatorname{diag}(y_1, \ldots, y_q)y \end{pmatrix}$$

and minimize $\bar{f}(x,y)$ with respect to $\bar{g}(x,y)=0$. Moreover, regularity for the constrained conditions is assumed:

(R) There is $\tau > 0$ such that $v \in \mathbb{R}^{l+q}$, $\| v \|_2 \leq \tau$ is a regular value of $\bar{g}$.

Let $\bar{A} : D_\tau \to \mathbb{R}^{l+q,l+q}$, $D_\tau := \{(x,y) \in D \times \mathbb{R}^q \mid \| \bar{g}(x,y) \|_2 < \tau\}$ be a smooth family of symmetric, positive definite matrices such that

$$\bar{B}(x,y) \;:=\; D\bar{g}(x,y)D\bar{g}(x,y)^T\bar{A}(x,y) \tag{1.1}$$

fulfils

$$\inf\{\mu_2(-\bar{B}(x,y)) \mid (x,y) \in D \times \mathbb{R}^q,\ \| \bar{g}(x,y) \|_2 \leq \tau\} \;\leq\; -\eta,\ \eta > 0. \tag{1.2}$$

Here $\mu_2(C)$ denotes the logarithmic norm of the matrix C (see, e.g., Coppel [5], p.41). Then, with $u := (x,y)$ the family of differential algebraic equations appropriate to the underlying minimization problem suggested by Schropp [14] reads

$$\begin{aligned} \dot{u} &= -\nabla \bar{f}(u) + D\bar{g}(u)^T\lambda, \\ \dot{v} &= -\bar{B}(u)v, \\ 0 &= \bar{g}(u) - v. \end{aligned} \tag{1.3}$$

The regularity condition (R) ensures that the DAE (1.3) is of index 2. Consistent initial values for the DAE (1.3) satisfy $v_0 = \bar{g}(u_0)$, $\lambda_0 = \bar{r}(u_0, \bar{g}(u_0))$ with

$$\bar{r}(u,v) \;:=\; (D\bar{g}(u)D\bar{g}(u)^T)^{-1}(D\bar{g}(u)\nabla\bar{f}(u) - \bar{B}(u)v). \tag{1.4}$$

In Schropp [14] it is shown that the evolutions of (1.3) became stationary in the long-time run and x-components of stable equilibria are local solutions of the underlying minimization problem.
In order to obtain an optimization code one has to discretize the DAE (1.3) or its corresponding index 1 DAE or the index 0 ordinary differential equation (ODE) with a suitable numerical integration scheme and show a convergence result for the discrete dynamics. This approach in its different index realizations and the large number of numerical methods to integrate ODE's, index 1 and 2 DAE's open the route to a variety of nearly unexplored optimization methods. In addition, a characterization of numerical schemes particularly well suited for optimization problems is necessary. This has to be done on the basis of theoretical arguments and practical numerical tests.
The dynamical systems approach to solve optimization problems has been discussed for the index 0 ODE and its discretization with linear multistep and Runge-Kutta methods in Schropp ([15]). In particular, for the BDF-methods this leads to reliable and suitable optimization codes. Nevertheless, the computation of the right-hand side of the ODE is costly, because an $(l+q \times l+q)$-system has to be solved for every evaluation of the vectorfield. The underlying optimization problem is attacked more efficient with the same accuracy when

applying the BDF-discretization to the index 1 system. Numerical results can be found in Schropp [14] and a discrete convergence result is presented in Schropp [18].
In the present paper we will analyse the behavior of certain Runge-Kutta and linear multistep methods with constant step size when applied directly to the index 2 DAE (1.3). It will be shown that the discrete dynamics inherits all decisive properties from the continuous solution flow. We focus our interest to a subgroup of numerical methods which admit an efficient implementation for optimization problems. From that point of view the BDF-like multistep methods and the half-explicit Runge-Kutta methods are of particular interest. In a variable step size implementation after a phase of initialising, the BDF-methods admit huge step sizes. On the other hand, the half-explicit Runge-Kutta schemes are known to have low computational cost per iteration step.
In addition, let us remark that from the dynamical systems point of view there is also good reason to analyse a slightly more general class than (1.3). It will be shown that this class includes gradient dynamical systems on manifolds. In this sense our results generalize the convergence results of Humphries & Stuart [10] and Schropp [12], [13] for discrete gradient systems on open subsets of $\mathbb{R}^N$ to gradient dynamics on manifolds.

2 The main results

We are motivated to consider the DAE

$$\begin{aligned} \dot{u} &= f(u,\lambda), \\ \dot{v} &= -B(u)v,\ B(u)\in\mathbb{R}^{l,l},\ \ \mu_2(-B(u))\le -\eta,\ \eta>0, \\ 0 &= g(u)-v \end{aligned} \tag{2.1}$$

for $u\in D_\tau := \{u\in\mathbb{R}^N \mid \| g(u)\|_2 < \tau\}$ and $v,\ \lambda\in\mathbb{R}^l$. Obviously, (2.1) includes the DAE (1.3) of our minimization problem. For equation (2.1) we assume

(A1) f, g and B are sufficiently differentiable with globally bounded derivatives.

(A2) There is a function ψ satisfying $Dg(u)f(u,\psi(u,v))+B(u)v=0$ for $u\in D_\tau$, $\| v\|_2<\tau$.

(A3) $Dg(u)\frac{\partial f}{\partial\lambda}(u,\psi(u,v))$ is invertible for $u\in D_\tau$, $\| v\|_2<\tau$ and the inverse possesses bounded norm.

In particular, the DAE (2.1) is of index 2. Moreover, (A3) implies that $Dg(u)$ has maximal rank for $u\in D_\tau$ and the solution set of $g(u)=0$ defines a submanifold of $\mathbb{R}^N$. After eliminating the v-variables the underlying index 0 ODE of (2.1) reads

$$\dot{u} = f(u,\psi(u,g(u))),\ u\in D_\tau, \tag{2.2}$$

that is, a classical ODE on an open subset of $\mathbb{R}^N$. Throughout the paper we denote the solution flow of (2.2) by $\bar{u}(t,u_0)$ and the flow generated by (2.1) with $(\bar{u}(t,u_0),\bar{v}(t,u_0),\bar{\lambda}(t,u_0))$,

$\bar{v}(t,u_0) = g(\bar{u}(t,u_0))$, $\bar{\lambda}(t,u_0) = \psi(\bar{u}(t,u_0),\bar{v}(t,u_0))$.
In the minimization case

$$f(u,\lambda) \;:=\; -\nabla \bar{f}(x) + D\bar{g}(x)^T\lambda,\; g = \bar{g},\; B = \bar{B} \tag{2.3}$$

(compare (1.3)) the function $\bar{r}$ from (1.4) plays the role of ψ in (A2) and equation (2.2) has the form

$$\dot{u} \;=\; (I - \bar{Q}(u))(-\nabla \bar{f}(u)) - D\bar{g}(u)^T(D\bar{g}(u)D\bar{g}(u)^T)^{-1}\bar{B}(u)\bar{g}(u). \tag{2.4}$$

Here, $\bar{Q}$ stands for the projector $\bar{Q}(u) = D\bar{g}(u)^T(D\bar{g}(u)D\bar{g}(u)^T)^{-1}D\bar{g}(u)$.

Lemma 2.1 *Let $\tau > 0$ such that $D_\tau := \{u \mid \| g(u) \|_2 < \tau\}$ is bounded and let (A1)-(A3) hold.*
Then every solution of (2.1) with initial value $u_0 \in D_\tau$, $v_0 = g(u_0)$ and $\lambda_0 = \psi(u_0,v_0)$ exists for all $t \geq 0$. Provided that every equilibrium of (2.1), (2.3) is hyperbolic, the trajectories of (2.1), (2.3) converge towards a steady state as $t \to \infty$.

A proof of Lemma 2.1 can be found in Schropp [14].
Here, $(\bar{u},\bar{v},\bar{\lambda})$ is a hyperbolic equilibrium of (2.1) if $\bar{u}$ is a hyperbolic equilibrium of (2.2).
Remark: Lemmata 2.1 and 2.2 in Schropp [14] ensure additionally that u-components of stationary points of (2.1), (2.3) are Kuhn-Tucker points for the underlying minimization problem and that stable equilibria are local minimas.
We are interested in the behavior of s-stage pth order half-explicit Runge-Kutta type methods with Butcher tableau

$$\begin{array}{c|c} c & A \\ \hline & b^T \end{array},\quad A = (A_{ij})_{1\leq i,j\leq s} \in \mathbb{R}^{s,s},\; A_{ij} = 0 \text{ for } i \leq j,\; b,c \in \mathbb{R}^s \tag{2.5}$$

as well as linear k-step BDF-type methods of order p with tableau

$$\begin{array}{ccc} \alpha_0 & \ldots & \alpha_k \\ \hline \beta_0 & \ldots & \beta_k \end{array}$$

and constant step size h when applied to (2.1) or (2.1), (2.3). We call a linear multistep method BDF-like, if $\beta_0 = \ldots \beta_{k-1} = 0$. Both classes of methods avoid the well known drift problems for DAE's of index greater than 1, since they retain the first order constraint $g(u) - v = 0$ exactly. For the half-explicit Runge-Kutta method introduced by Hairer, Lubich and Roche [8] we impose the conditions:

(B1) $A_{i+1,i} \neq 0$ for $i = 1,\ldots s-1$ and $b_s \neq 0$.

(B2) The method is of order $p \geq 1$.

The numerical scheme reads as follows. Solve

$$\begin{aligned} U - (\mathbb{I} \otimes u_n) &= h(A \otimes I)\bar{f}(U, \Lambda), \\ V - (\mathbb{I} \otimes v_n) &= -h(A \otimes I)\bar{B}(U)V, \\ 0 &= \bar{g}(U) - V \end{aligned} \tag{2.6}$$

in the case $A_{i,j} = 0$ for $j \geq i$ iteratively and obtain U^n, V^n and Λ_i^n, $i = 1, \ldots, s-1$. Finally, Λ_s^n and u_{n+1}, v_{n+1} are computed by

$$\begin{aligned} u_{n+1} &= u_n + h(b^T \otimes I)\bar{f}(U^n, \Lambda^n), \\ v_{n+1} &= v_n - h(b^T \otimes I)\bar{B}(U^n)V^n, \\ 0 &= g(u_{n+1}) - v_{n+1}. \end{aligned} \tag{2.7}$$

In the linear multistep situation we assume the method to be strictly stable, i.e., the polynomial $p(\mu) = \sum_{i=0}^{k} \alpha_i \mu^i$ possesses 1 as simple zero and all other roots $\bar{\mu}$ of $p(\mu) = 0$ satisfy $|\bar{\mu}| < 1$. We apply the numerical scheme to equation (2.1) and find

$$\begin{aligned} \sum_{i=0}^{k} \alpha_i u_{n+i} &= h \sum_{i=0}^{k} \beta_i f(u_{n+i}, \lambda_{n+i}), \\ \sum_{i=0}^{k} \alpha_i v_{n+i} &= -h \sum_{i=0}^{k} \beta_i B(u_{n+i}) v_{n+i}, \\ 0 &= g(u_{n+k}) - v_{n+k}. \end{aligned} \tag{2.8}$$

Our main result is the following discrete analogue of Lemma 2.1.

Theorem 2.2 *Let the conditions of Lemma 2.1 hold for equation (2.1). By (u_n, v_n, λ_n) we denote the sequences generated with a half-explicit Runge-Kutta method fulfilling (B1)-(B2) or with a strictly stable BDF-like linear multistep method when applied to (2.1) with initial values $u_0 \in D_{\hat{\tau}}$, $\hat{\tau} < \tau$ arbitrary, $v_0 = g(u_0)$ and $\lambda_0 = \psi(u_0, v_0)$.*
Then there is $h_0 > 0$ such that the half-explicit Runge-Kutta or BDF-like multistep iteration is well defined for $0 < h < h_0$, $n \in \mathbb{N}$. Moreover, let the maps f, g, B have the form (2.3) and let all equilibria of (2.1), (2.3) be hyperbolic. Then $h_1 \in]0, h_0]$ exists such that the half-explicit Runge-Kutta or BDF-like multistep sequences (u_n, v_n, λ_n) with constant step size $h \in]0, h_1]$ converge towards a steady state of (2.1), (2.3) as $n \to \infty$.

Finally, let us outline the practical consequences of Theorem 2.2. Theorem 2.2 directly opens the route to a class of minimization methods using index 2 DAE techniques. Since an efficient numerical integrator must use variable step sizes, we use Theorem 2.2 merely as a guideline how to proceed.
Let us compare the index 2 DAE approach with the classical minimization methods. The most powerful classical methods are the so called sequential quadratic programming methods (SQP). SQP methods can exploit structures of the functions $\hat{f}$, $\hat{g}$, $\hat{k}$ which define the

optimization problem. They work excellently when applied to linear problems and show only small weaknesses when attacking nonlinear problems. For a problem with N independent variables, l equality and q inequality constraints the computational effort in every iteration step is to solve a symmetric $(N+q+l)$-dimensional system of linear equations. Thus, SQP methods work efficient.

However, in the last 20 years there has been a huge development in the numerical treatment of DAEs. Due to formula (2.8) a nonlinear $(N+2q+2l)$-dimensional system has to be solved in every time step for the BDF-method. The costs per time step are higher as in the SQP-case but it is well known that after a few initialization steps a variable step size BDF-scheme can realize huge step sizes. When applying a half-explicit Runge-Kutta method to the DAE (1.3) we try to reduce the computational cost per time step. Thus, we focus our interest to the most simple method, the half-explicit Euler method. This means $s = 1$, $A = 0$ and $c = b = 1$ for the coefficients of the tableau (2.5). In this case, the system (2.6) has the solution $U = u_n$, $V = g(u_n)$ and we have to solve merely (2.7). Due to structure of equation (2.7) we insert the representation of u_{n+1} and v_{n+1} in the first two lines of (2.7) into the third one and obtain the resulting system

$$0 = g(u_n + hf(u_n, \lambda_n)) - g(u_n) + hB(u_n)g(u_n) \tag{2.9}$$

to determine λ_n. Then the first two relations of (2.7) determine u_{n+1}, v_{n+1}. Equation (2.9) shows that the computational cost per time step is to solve a nonlinear $(l+q)$-dimensional system. So the half-explicit Euler-method seems to be particularly well suited for minimization problems with $l + q \ll N + l + q$.

We apply the half-explicit Euler method in a variable step size realization and the variable step size BDF-method (NAG routine D02NGF) to the index 2 DAE (1.3) with $B = I$. The NAG-routine D02NGF is driven with the option, that the nonlinear systems in every time step have to be solved with functional iterations instead of Newtons method to reduce computational costs.

Our test examples are the hydrostatic skeleton model of Beyn, Wadepuhl [3] and the optimization of an alkylation process in the chemotechnical industry (compare Bracken, Mc Cormick [4]). Numerical tests with these optimization problems arising in applications have been already made with an efficient realization of an SQP-method (NAG-routine E04UCF) and with a BDF-method (NAG-routine D02NGF) applied to the index 0 and index 1 version of equation (1.3) in Schropp [14], section 4.

Now, let us present the results for the index 2 DAE approach: Since $q = 0$, $l = 3$ and $N = 43$ hold in our skeleton example it is no surprise that the half-explicit Runge-Kutta method is the most efficient way to solve that problem. But similar to the SQP-method the half-explicit method shows some small weaknesses when getting started with an initial value $x_0 \in \mathbb{R}^N$ possessing the symmetry $(x_0)_i = (x_0)_{N+1-i}$, $i = 1, \ldots, N$ (see Schropp [14], section 4). Depending on the absolute and relative tolerances chosen for the accuracy of the numerical integration, the index 2 BDF-realization is as efficient as the SQP method. But the index 2 BDF-method can handle the problem for symmetric initial values too.

When dealing with the optimization of an alkylation process the situation is different. Here

we have $q = 28$, $l = 3$ and $N = 10$. Since $q = 28$ is the dominant value, the half-explicit Euler method has no advantage in the computational cost per time step. The SQP realization is the most efficient way to solve the problem. In particular, 20 of the 28 inequality constraints have the form $a_i \leq x_i \leq b_i$, $i = 1, \ldots, 10$ and they are treated very efficient by the SQP-method (compare Schropp [14], section 4). Since there are no hidden difficulties in that problem, it turns out that the SQP approach, the BDF index 2 approach and the half-explicit Euler method are able to solve that problem for suitable many initial values. But the SQP-method has small efficiency advantages compared to the index 2 BDF-method and major advantages in comparison with the half-explicit Euler method.
To summarize we can say that the index 2 BDF-realization is a very reliable and efficient optimization procedure. In comparison to the index 1 BDF-approach the cost per function evaluation is a bit lower in the index 2 case, but the software code is more complicated to initialise. Nevertheless, for problems with $l + q \gg N$ and a lot of linear constraints, an SQP approach is more efficient. But for problems with $l + q \ll N$ the half-explicit method is a particularly well suited reliable and efficient approach.

3 Existence of the discrete dynamics

In this section we will guarantee the existence and uniqueness of the discrete iterates generated by half-explicit Runge-Kutta and strictly stable BDF-like multistep methods for $n \in \mathbb{N}$. For $n \in \mathbb{N}$ with $0 < nh < T$, $T > 0$ fix the solvability of the discrete systems (2.6)-(2.8) is guaranteed by the standard theory, see e.g., Hairer and Wanner [9], Ch. VII.3 and VII.4. To establish the existence of the discrete iterates for $n \in \mathbb{N}$ it is useful to distinguish the dynamical variables u and v. This is possible with the concept of vector norms. A functional $|\cdot|: W \to \mathbb{R}^k$ on a vector space W is called a generalized norm, if

$$|v| \geq 0, \quad |v| = 0 \iff v = 0, \quad |v_1 + v_2| \leq |v_1| + |v_2|$$

holds with the natural ordering "$\leq$" on $\mathbb{R}^k$. Every norm $\|\cdot\|_*$ in $\mathbb{R}^k$ defines a norm $\|\cdot\|$ in W via $\|v\| = \| |v| \|_*$.

Lemma 3.1 *Let the assumptions of Theorem 2.2 hold and let $u_0 \in D_{\hat{\tau}}$, $\hat{\tau} < \tau$, $v_0 = g(u_0)$, $\lambda_0 = \psi(u_0, v_0)$ be a consistent initial value for the DAE (2.1). Then the BDF-iterates (u_n, v_n, λ_n) exist for $n \in \mathbb{N}$.*

Proof: We assume $\alpha_k = 1$ and replace equation (2.8) for $h > 0$ by the equivalent system

$$\begin{aligned} u_{n+k} &= \sum_{i=0}^{k-1} -\alpha_i u_{n+i} + h\beta_k f(U, \Lambda), \\ v_{n+k} &= \sum_{i=0}^{k-1} -\alpha_i v_{n+i} - h\beta_k B(U)V, \\ \lambda_{n+k} &= \Lambda. \end{aligned} \tag{3.1}$$

Here (U,V,Λ) denotes the solution of

$$\begin{aligned} U + \sum_{i=0}^{k-1} \alpha_i u_{n+i} - h\beta_k f(U,\Lambda) &= 0, \\ V + \sum_{i=0}^{k-1} \alpha_i v_{n+i} + h\beta_k B(U)V &= 0, \qquad (3.2) \\ \int_0^1 Dg(\sum_{i=0}^{k-1} -\alpha_i u_{n+i} + \tau(U - (\sum_{i=0}^{k-1} -\alpha_i u_{n+i}))) \, d\tau \; \beta_k f(U,\Lambda) & \\ +\beta_k B(U)V + \frac{1}{h}(g(-\sum_{i=0}^{k-1} \alpha_i u_{n+i}) + \sum_{i=0}^{k-1} \alpha_i g(u_{n+i})) &= 0 \end{aligned}$$

(compare Hairer, Wanner [9], p.483).
We prove Lemma 3.1 by applying Lemma 4.2 of Beyn, Schropp [2] onto

$$T(h,w,U,V,\Lambda) = \begin{pmatrix} U - \sigma(w) - h\beta_k f(U,\Lambda) \\ V - \eta(w) + h\beta_k B(U)V \\ \int_0^1 Dg(\sigma(w) + \tau(U - \sigma(w))) \, d\tau \; \beta_k f(U,\Lambda) \\ +\beta_k B(U)V + h^{-1}(g(\sigma(w)) - \eta(w)) \end{pmatrix} = 0 \qquad (3.3)$$

with $w := (w_1, \ldots, w_k)$, $\sigma(w) := -\sum_{i=0}^{k-1} \alpha_i w_{i+1}$, $\eta(w) := -\sum_{i=0}^{k-1} \alpha_i g(w_{i+1})$.
We introduce the generalized norm $\mid (U,V,\Lambda) \mid = (\parallel U \parallel, \parallel V \parallel, \parallel \Lambda \parallel) \in \mathbb{R}^3$ and the central point $v_0(w) := (\sigma(w), g(\sigma(w)), \psi(\sigma(w), g(\sigma(w))))$. Using (A2) we obtain

$$T(h,w,v_0(w)) = (O(h), g(\sigma(w)) - \eta(w) + O(h), h^{-1}(g(\sigma(w) - \eta(w)))) \qquad (3.4)$$

for $w \in \mathbb{R}^{Nk}$, $\sigma(w) \in D_{\hat{\tau}}$. Moreover, we can can calculate

$$\frac{\partial}{\partial(U,V,\Lambda)} T(h,w,v_0(w)) = \begin{pmatrix} I + O(h) & 0 & O(h) \\ O(h) & I + O(h) & 0 \\ O(1) & O(1) & \beta_k Dg(\sigma(w)) \frac{\partial f}{\partial \lambda}(\tau(w)) \end{pmatrix} \qquad (3.5)$$

with $\tau(w) = (\sigma(w), \psi(\sigma(w), g(\sigma(w))))$. Obviously, $\frac{\partial}{\partial(U,V,\Lambda)} T(h,w,v_0(w))$ is invertible for small h. Formulae (3.4), (3.5) ensure that we can apply Lemma 4.2 of Beyn, Schropp [2] for parameters $w \in \mathbb{R}^{Nk}$, $\sigma(w) \in D_{\hat{\tau}}$, $\hat{\tau} < \tau$, provided $h > 0$ and $r(h,w) := h^{-1}(g(\sigma(w) - \eta(w)))$ are sufficiently small.
This is not true in general but in applications we have $w = (w_1, \ldots, w_k) = (u_n, \ldots, u_{n+k-1})$ for some $n \in \mathbb{N}$. These iterations satisfy

$$\parallel u_{n+k} - u_{n+k-1} \parallel \; \leq \; Ch, \; n \in \mathbb{N}. \qquad (3.6)$$

Assuming (3.6), using $\sum_{i=0}^{k-1} \alpha_i = -1$ and Taylor expansion we can compute

$$\begin{aligned} r(h, u_n, \ldots, u_{n+k-1}) &\leq h^{-1}[g(-\sum_{i=0}^{k-1} \alpha_i u_n - \sum_{i=0}^{k-1} \alpha_i (u_{n+i} - u_n)) \\ &\quad + \sum_{i=0}^{k-1} \alpha_i g(u_n + u_{n+i} - u_n)] \\ &= h^{-1}[g(u_n) + Dg(u_n)(\sum_{i=0}^{k-1} -\alpha_i (u_{n+i} - u_n)) \\ &\quad + \sum_{i=0}^{k-1} \alpha_i (g(u_n) + Dg(u_n)(u_{n+i} - u_n) + O(h^2))] \\ &\leq \hat{C}h, \ n \in \mathbb{N}. \end{aligned} \tag{3.7}$$

Now, we outline a proof of formula (3.6).
By construction, we have $u_{n+k} = -\sum_{i=0}^{k-1} \alpha_i u_{n+i} + h\beta_k f(U, \Lambda)$ and

$$u_{n+k} - u_{n+k-1} = \sum_{i=0}^{k-2} \alpha_i (u_{n+k-1} - u_{n+i}) + O(h) \tag{3.8}$$

follows. Moreover, (3.8) implies for $j = 1, \ldots, k-2$ the relation

$$u_{n+k} - u_{n+j} = \sum_{i=0}^{k-2} \alpha_i (u_{n+k-1} - u_{n+i}) + u_{n+k-1} - u_{n+j} + O(h). \tag{3.9}$$

With

$$\gamma_{n+k-1} = \begin{pmatrix} u_{n+k-1} - u_n \\ u_{n+k-1} - u_{n+1} \\ \vdots \\ u_{n+k-1} - u_{n+k-2} \end{pmatrix} = \begin{pmatrix} \gamma_{n+k-1,1} \\ \gamma_{n+k-1,2} \\ \vdots \\ \gamma_{n+k-1,k-1} \end{pmatrix} \in \mathbb{R}^{N(k-1)} \tag{3.10}$$

we can rewrite (3.8) and (3.9) in the form

$$\gamma_{n+k} = (\hat{A} \otimes I)\gamma_{n+k-1} + O(h).$$

Here $\hat{A}$ stands for

$$\hat{A} = \begin{pmatrix} \Theta & I_{k-2} \\ 0 & \Theta^T \end{pmatrix} + \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \cdot (\alpha_0, \ldots, \alpha_{k-2})^T.$$

Since we have $\mid \lambda \mid < 1$ for all eigenvalues of $\hat{A} \otimes I$ (see Schropp [16], Appendix), there is a norm $\| \cdot \|_*$ on $\mathbb{R}^{N(k-1)}$ such that $\| (\hat{A} \otimes I) \|_* \leq \rho < 1$ holds. This gives us

$$\begin{aligned} \| \gamma_{n+k} \|_* &\leq \rho \| \gamma_{n+k-1} \|_* + \tilde{C}h \leq \sum_{r=0}^{n} \rho^r \tilde{C}h + \rho^{n+1} \| \gamma_{k-1} \|_* \\ &\leq \frac{1}{1-\rho} \tilde{C}h + \rho^{n+1} \hat{C}h \leq Ch, \ n \in \mathbb{N} \end{aligned} \tag{3.11}$$

and (3.6) is shown.
With formula (3.7) and $v_0(u_n, \ldots, u_{n+k-1}) = (\sigma_n, g(\sigma_n), \psi(\sigma_n, g(\sigma_n)))$, $\sigma_n = -\sum_{i=0}^{k-1} \alpha_i u_{n+i}$ we can deduce

$$\mid T(h, u_n, \ldots, u_{n+k-1}, v_0(u_n, \ldots, u_{n+k-1})) \mid \ = \ O(h)(1,1,1).$$

In addition, (3.5) shows that $\frac{\partial}{\partial(U,V,\Lambda)} T(h, u_n, \ldots, u_{n+k-1}, v_0(u_n, \ldots, u_{n+k-1}))$, $0 < h \leq h_0$ is invertible for $n \in \mathbb{N}$. Thus, the new iterate $(u_{n+1}, v_{n+1}, \lambda_{n+1})$ exists by Lemma 4.2 of Beyn, Schropp [2]. Additionally, an application of the implicit function theorem onto equation (3.3) guarantees that (U, V, Λ) depend smoothly on (h, w).
It remains to show that $\sigma_{n+1} \in D_{\hat{\tau}}$ holds. After extracting V from (3.1) the v-component reads

$$v_{n+k} \ = \ (I + h\beta_k B(U(h, u_n, \ldots, u_{n+k-1})))^{-1} (-\sum_{i=0}^{k-1} \alpha_i v_{n+i}). \tag{3.12}$$

We rewrite (3.12) as a one-step method in $\mathbb{R}^{ls}$. With the matrix
$\Gamma_h^n = (I + h\beta_k B(U(h, u_n, \ldots, u_{n+k-1})))^{-1}$ we find the scheme

$$\begin{pmatrix} v_{n+1} \\ \vdots \\ v_{n+k-1} \\ v_{n+k} \end{pmatrix} = \begin{pmatrix} 0 & I & & \\ & \ddots & \ddots & \\ & & 0 & I \\ -\alpha_0\Gamma_h^n & -\alpha_1\Gamma_h^n & \ldots & -\alpha_{k-1}\Gamma_h^n \end{pmatrix} \cdot \begin{pmatrix} v_n \\ \vdots \\ v_{n+k-2} \\ v_{n+k-1} \end{pmatrix}. \tag{3.13}$$

Now, we want to ensure a norm on $\mathbb{R}^{ls}$ such that the iteration matrix in (3.13) has a corresponding operator norm less than 1 for $h > 0$ sufficiently small. In the case
$B(U(h, u_n, \ldots, u_{n+k-1})) = C$, $C \in \mathbb{R}^{l,l}$ fix this is shown in Beyn (1987), Lemma 4.2. Since we have

$$\mu_2(-B(U(h, u_n, \ldots u_{n+k-1}))) \ \leq \ -\eta \quad \text{uniformly for } n \in \mathbb{N}$$

we can adapt Lemma 4.2, Beyn (1987) to our situation. This ensures a norm on $\mathbb{R}^{ls}$ and $s > 0$ such that the corresponding operator norm of the iteration matrix in (3.13) is less equal $1 - sh$. Hence, $\| (v_{n+1}, \ldots, v_{n+k}) \| \leq \| (v_n, \ldots, v_{n+k-1}) \|$ follows and with $v_n = g(u_n)$ the relation $u_n \in D_{\hat{\tau}}$ is justified. This finishes our proof.

Our next issue is to enlarge the domain of the functions (U, V, Λ). This will be needed in the process of proving Theorem 2.2. We introduce the cut-off function $\chi \in C^\infty(\mathbb{R}^{Nk}, [0,1])$ defined by

$$\chi(w) \;=\; 1 \text{ for } w \in S_C, \qquad \chi(w) \;=\; 0 \text{ for } w \in \mathbb{R}^{Nk} \backslash S_{2C}$$

with $S_C = \{w = (w_1, \ldots, w_k) \in \mathbb{R}^{Nk} \mid \| w - (\mathbb{I} \otimes w_1) \| \leq kC\}$ and C from (3.6) for a k-step method. Then, with $S(w) := g(\sigma(w)) - \eta(w)$ we replace T from (3.3) by the modified operator

$$\begin{aligned} \hat{T}(h, w, U, V, \Lambda) \;&=\; \begin{pmatrix} U - \sigma(w) - h\beta_k f(U, \Lambda) \\ V - \eta(w) + h\beta_k B(U)V + (\chi(h^{-1}w) - 1)S(w) \\ \int_0^1 Dg(\sigma(w) + \tau(U - \sigma(w)))\, d\tau \;\beta_k f(U, \Lambda) \\ + \beta_k B(U)V + \chi(h^{-1}(w))h^{-1}S(w) \end{pmatrix} \;=\; 0 \\ &\qquad \text{for } \; 0 < h < h_0, \; w \in \mathbb{R}^{Nk}. \end{aligned} \tag{3.14}$$

By construction $\hat{T}(h, u_n, \ldots, u_{n+k-1}, U, V, \Lambda) = T(h, u_n, \ldots, u_{n+k-1}, U, V, \Lambda)$ holds for $n \in \mathbb{N}$. Moreover, for S we can deduce

$$\begin{aligned} S(w) \;&=\; S(\mathbb{I} \otimes w_1) + DS(\mathbb{I} \otimes w_1)(w - (\mathbb{I} \otimes w_1)) + O(\| w - (\mathbb{I} \otimes w_1) \|^2) \\ &=\; O(\| w - (\mathbb{I} \otimes w_1) \|^2). \end{aligned} \tag{3.15}$$

Thus, with formula (3.15) and the definition of χ

$$\begin{aligned} \sup\{\| \hat{r}(h, w) \| \mid \; w \in \mathbb{R}^{Nk}\} \;&=\; \sup\{\| \hat{r}(h, w) \| \mid \| h^{-1}(w - (\mathbb{I} \otimes w_1)) \| \leq 2kC\} \\ &=\; h^{-1} \sup\{\| S(w) \| \mid \| (w - (\mathbb{I} \otimes w_1)) \| \leq 2kCh\} \\ &=\; O(h) \end{aligned} \tag{3.16}$$

follows and $\hat{T}(h, w, v_0(w)) = O(h)$ is valid for $0 < h \leq h_0$, $(w_1, \ldots, w_k) \in D_{\hat{r}}^k$, $\sigma(w) \in D_{\hat{r}}$. The corresponding discrete iteration scheme reads

$$\begin{aligned} u_{n+k} \;&=\; \sum_{i=0}^{k-1} -\alpha_i u_{n+i} + h\beta_k f(U(h, u_n, \ldots, u_{n+k-1}), \Lambda(h, u_n, \ldots, u_{n+k-1})), \\ v_{n+k} \;&=\; \sum_{i=0}^{k-1} -\alpha_i v_{n+i} - h\beta_k B(U(h, u_n, \ldots, u_{n+k-1}))V(h, u_n, \ldots, u_{n+k-1}), \\ \lambda_{n+k} \;&=\; \Lambda(h, u_n, \ldots, u_{n+k-1}). \end{aligned} \tag{3.17}$$

Obviously, the u-component of equation (3.17) can be regarded as nonlinear multistep method of order k applied to the index 0 equation (2.2). This follows from the local error relation

$$\begin{aligned} \bar{u}(t + kh, u_0) \;&=\; U(h, \bar{u}(t, u_0), \ldots, \bar{u}(t + (k-1)h, u_0)) + O(h^{k+1}), \\ \bar{\lambda}(t + kh, u_0) \;&=\; \Lambda(h, \bar{u}(t, u_0), \ldots, \bar{u}(t + (k-1)h, u_0)) + O(h^k) \end{aligned} \tag{3.18}$$

(see Hairer, Wanner [9], p.485) for a strictly stable k-step BDF-method.

Finally, let us complete the section with the existence results for the half-explicit Runge-Kutta methods. The existence of the iterates (u_n, v_n, λ_n), $n \in \mathbb{N}$ is shown in Schropp ([17]), section 3 and 4. For the u-component we obtain the iteration scheme

$$u_{n+1} \;=\; u_n + h(b^T \otimes I)\bar{f}((u_n, U_2(h,u_n), \ldots, U_s(h,u_n)), \Lambda(h,u_n)). \tag{3.19}$$

Here, $U_i(h,u_n)$, $i = 2,\ldots,s$, $\Lambda(h,u_n)$ denote the appropriate parts of the solution of (2.6), (2.7). Obviously, formula (3.19) can be regarded as a one-step method applied to the corresponding index 0 initial value problem

$$\dot{u} \;=\; f(u, \psi(u, g(u))),\; u(0) = u_0. \tag{3.20}$$

The local error properties of (3.19), (3.20) are discussed in Hairer, Brasey [7]). It turns out that the coefficients A_{ij}, b_j of the Runge-Kutta scheme have to satisfy additional order conditions. Finally if desired, we compute λ_{n+1}. This can be done either very efficient by $\lambda_{n+1} = \Lambda_s(h,u_n)$ or more accurate by the use of the index 2 condition (A3), that is, $\lambda_{n+1} = \psi(u_{n+1}, g(u_{n+1}))$.

4 The gradient case

The first point in this section to show is that for the proposed numerical schemes the fixed point set coincides for small step sizes h with the set of equilibria of the corresponding continuous equation.

Lemma 4.1 *Consider the differential equation (2.2) together with its discretization (3.19) arising from a half-explicit Runge-Kutta scheme (2.6), (2.7). Then the fixed point set of the discrete dynamics and the set of stationary points of (2.2) coincide for $h > 0$ sufficiently small provided the function f in (2.2) is globally lipschitzian.*

Proof: Let $\bar{u}$ be an equilibrium of the ODE $\dot{u} = f(u, \psi(u, g(u)))$. By definition of ψ we have

$$g(\bar{u}) \;=\; -B(\bar{u})^{-1} Dg(\bar{u}) f(\bar{u}, \psi(\bar{u}, g(\bar{u}))) \;=\; 0.$$

Hence, we can conclude that $U(h,\bar{u}) = \mathbb{I} \otimes \bar{u}$, $V(h,\bar{u}) = \mathbb{I} \otimes g(\bar{u})$, $\Lambda_i(h,\bar{u}) = \psi(\bar{u}, g(\bar{u}))$, $i = 1,\ldots,s-1$ in (2.6). Moreover, we obtain $u_{n+1} = \bar{u}$, $v_{n+1} = g(\bar{u})$, $\Lambda_s(h,\bar{u}) = \psi(\bar{u}, g(\bar{u}))$ from (2.7).
On the other hand, let $\bar{u}$ be a fixed point of the discrete dynamics (3.19). Using the stability inequality (3.21) in Schropp [17] for half-explicit Runge-Kutta schemes, we can compute

$$\begin{aligned}
\| f(\bar{u}, \psi(\bar{u}, g(\bar{u}))) \| &= \| (b^T \otimes I)\bar{f}(\mathbb{I} \otimes \bar{u}, \mathbb{I} \otimes \psi(\bar{u}, g(\bar{u}))) \| \\
&\leq \| (b^T \otimes I)[\bar{f}(\mathbb{I} \otimes \bar{u}, \mathbb{I} \otimes \psi(\bar{u}, g(\bar{u}))) \\
&\quad -\bar{f}(U(h,\bar{u}), \Lambda(h,\bar{u})))] \| \\
&\leq C(\| U(h,\bar{u}) - \mathbb{I} \otimes \bar{u} \| + \| \Lambda(h,\bar{u}) - \mathbb{I} \otimes \psi(\bar{u}, g(\bar{u})) \|) \\
&= O(h)[\| f(\bar{u}, \psi(\bar{u}, g(\bar{u}))) \| + \| g(\bar{u}) \|].
\end{aligned} \tag{4.1}$$

Our next step is to establish an estimation for $g(\bar{u})$. With the use of the second order constraint we obtain

$$\begin{aligned} \| g(\bar{u}) \| &= \| B(\bar{u})^{-1} Dg(\bar{u}) f(\bar{u}, \psi(\bar{u}, g(\bar{u}))) \| \\ &\leq \| B(\bar{u})^{-1} Dg(\bar{u}) \| \cdot \| f(\bar{u}, \psi(\bar{u}, g(\bar{u}))) \| \\ &\leq O(h)(\| g(\bar{u}) \| + \| f(\bar{u}, \psi(\bar{u}, g(\bar{u}))) \|). \end{aligned} \tag{4.2}$$

The inequalities (4.1), (4.2) can be rewritten as a system. This gives

$$\begin{pmatrix} 1 - O(h) & -O(h) \\ -O(h) & 1 - O(h) \end{pmatrix} \cdot \begin{pmatrix} \| f(\bar{u}, \psi(\bar{u}, g(\bar{u}))) \| \\ \| g(\bar{u}) \| \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and $\| f(\bar{u}, \psi(\bar{u}, g(\bar{u}))) \| = 0$, $\| g(\bar{u}) \| = 0$ follows. This finishes our proof.

Remark: When applying a BDF-like multistep method onto (2.1), (2.3) the u-component of the resulting discrete dynamics has the shape (3.17), that is, a nonlinear multistep discretization of (2.2). For linear multistep methods applied to ODE's it is well known that for $h > 0$ sufficiently small the fixed point set of the discrete dynamics and the set of equilibria of the ODE coincide (see, e.g., Stuart, Humphries [19], Theorem 5.3.8). This proof can be adapted to handle the nonlinear multistep situation too, provided the function f in (2.2) is globally lipschitzian.

In the sequel we consider the DAE (2.1) with gradient structure, that is,

$$f(u, \lambda) := -\nabla \bar{f}(u) + D\bar{g}(u)^T \lambda, \quad g = \bar{g} \quad \text{and} \quad B = \bar{B}$$

which originates from our minimization problem (1.3). In this situation, the underlying ODE reads

$$\begin{aligned} \dot{u} &= f(u, \psi(u, g(u))) = k(u) \\ &=: (I - \bar{Q}(u))(-\nabla \bar{f}(u)) - D\bar{g}(u)^T (D\bar{g}(u) D\bar{g}(u)^T)^{-1} \bar{B}(u) \bar{g}(u) \end{aligned} \tag{4.3}$$

with $\bar{Q}(u) = D\bar{g}(u)^T (D\bar{g}(u) D\bar{g}(u)^T)^{-1} D\bar{g}(u)$.
We apply a half-explicit Runge-Kutta method to (2.1), (2.3). It is shown in the previous sections that the half-explicit Runge-Kutta iterates are well defined and the u-component of the discrete iteration can be regarded as smooth one-step method applied to the ODE (4.3). Next, the reader should check that the discrete dynamics fulfils the assumptions of Theorem 2.4 in Schropp [15]. Application of Theorem 2.4, Schropp [15] then shows

$$\lim_{n \to \infty} u_n = \bar{u} \text{ for some } \bar{u} \text{ satisfying } f(\bar{u}, \psi(\bar{u}, g(\bar{u}))) = 0$$

and the conservation of the first order constraint yields $\lim_{n \to \infty} v_n = g(\bar{u}) = 0$. Additionally, for λ- component of the half-explicit Runge-Kutta method we obtain

$$\lim_{n \to \infty} \lambda_n = \psi(\bar{u}, 0)$$

provided $\lambda_{n+1} = \Lambda_s(h, u_n)$ or $\lambda_{n+1} = \psi(u_{n+1}, g(u_{n+1}))$ hold.

Now we analyse the behavior of the BDF-like multistep method when applied onto (2.1), (2.3). By (3.17), the u-component of the resulting discrete dynamics can be regarded as nonlinear multistep method applied onto the underlying ODE (3.20).
Slightly more general than the u-iteration in (3.17) we consider the ODE $\dot{u} = F(u)$ in $\Omega \subset \mathbb{R}^N$ open and apply an arbitrary multistep method of the form

$$\sum_{i=0}^{k} \alpha_i' u_{n+i} \quad = \quad h f_h(u_n, \ldots, u_{n+k-1}, u_{n+k}) \tag{4.4}$$

with $f_h : \Pi \to \mathbb{R}^N$, $0 < h < h_0$, $\Pi := \{(w_1, \ldots, w_{k+1}) \in \Omega^{k+1} \mid \sum_{i=0}^{k-1} -\alpha_i' w_{i+1}, \sum_{i=0}^{k-1} -\alpha_i' w_{i+2} \in \Omega\} \subset \Omega^{k+1}$, $\alpha_k' = 1$ completed by a starting procedure

$$u_{n+1} \quad = \quad u_n + h f_{h,s}(u_n), \; n = 0, \ldots, k-2. \tag{4.5}$$

We assume that (4.4) and the starting procedure possess order p and that (4.4) is strictly stable, that is, $p(\mu) = \sum_{i=0}^{k} \alpha_i' \mu^i$ fulfils the strong root condition. Moreover, let the method function f_h be sufficiently smooth with globally bounded derivatives. Similar to Beyn [1], section 4, we assume

(V1) There exist β_i', $i = 0, \ldots, k$ such that $f_h(u, \ldots, u) = \sum_{i=0}^{k} \beta_i' F(u) + O(h)$, $\sum_{i=0}^{k} \alpha_i' = 0$, $\sum_{i=0}^{k} i\alpha_i' = \sum_{i=0}^{k} \beta_i'$ and $\frac{\partial}{\partial u_{n+i}} f_h(u, \ldots, u) = \beta_i' DF(u) + O(h)$, $i = 0, \ldots, k$ hold.

(V2) The set of fixed points of the multistep dynamics and the set of equilibria of $\dot{u} = F(u)$ coincide for $h > 0$ sufficiently small.

(V3) The method is of order $p \geq 1$.

(V1) ensures that (4.4) is not too far from a classical linear multistep method. Then the following extension of Theorem 2.4, Schropp [15] is valid.

Lemma 4.2 *Suppose the assumptions of Lemma 2.1 hold. Let u_n denote the strictly stable pth order multistep approximation (4.4), (4.5) of the solution $\phi(nh, u_0)$ of equation (3.20) with $u_0 \in D_{\hat{\tau}}$, $\hat{\tau} < \tau$ and let the multistep method fulfil (V1)-(V3). Then there is $h_0 > 0$ such that the multistep approximation u_n with step size $h \in]0, h_0]$ exists for $n \in \mathbb{N}$. Additionally,*

$$\lim_{n \to \infty} u_n \quad = \quad \bar{u}$$

holds for some $\bar{u}$ satisfying $f(\bar{u}, \psi(\bar{u}, g(\bar{u}))) = 0$, $0 < h < h_0$.

A proof of Lemma 4.2 will be given in next section.

Our objective here to show is that in the case $F(u) = f(u, \psi(u, g(u)))$ (compare (3.20))

the u component of the special scheme (3.17) satisfies (V1). (V2) is already shown, (V3) follows from the local error relation (3.18) and with the modified operator $\hat{T}$ from (3.14) the iteration is defined on the set $\Pi \subset D_{\hat{\tau}}^{k+1}$.
To verify (V1), with $\sigma(w) = -\sum_{i=0}^{k-1} \alpha_i w_{i+1}$, (A2), (3.16) and $\hat{T}$ from (3.14) we compute

$$\hat{T}(h, w, \sigma(w), g(\sigma(w)), \psi(\sigma(w), g(\sigma(w)))) \;=\; (O(h), O(h), O(h)).$$

In section 3 it has been shown that Lemma 4.2 of Beyn, Schropp [2] applies to equation (3.14). The stability inequality of that Lemma then yields

$$\begin{aligned} U(h,w) &= \sigma(w) + O(h), \\ V(h,w) &= g(\sigma(w)) + O(h), \\ \Lambda(h,w) &= \psi(\sigma(w), g(\sigma(w))) + O(h) \end{aligned} \tag{4.6}$$

and with $w = \mathbb{I} \otimes u$ the relation

$$f_h(\mathbb{I} \otimes u, u) \;=\; \beta_k f(u, \psi(u, g(u))) + O(h) \tag{4.7}$$

follows. In addition, a careful differentiation argument shows

$$\begin{aligned} \frac{\partial U}{\partial w_i}(h,w) &= -\alpha_{i-1} I + O(h), \\ \frac{\partial V}{\partial w_i}(h,w) &= -\alpha_{i-1} Dg(\sigma(w)) + O(h), \\ \frac{\partial \Lambda}{\partial w_i}(h,w) &= (-\alpha_{i-1})\Big(\frac{\partial \psi}{\partial u}(\sigma(w), g(\sigma(w))) \\ &\qquad + \frac{\partial \psi}{\partial v}(\sigma(w), g(\sigma(w))) Dg(\sigma(w)) \Big) + O(h). \end{aligned} \tag{4.8}$$

We insert $w = \mathbb{I} \otimes u$ and obtain

$$\frac{\partial f_h}{\partial w_i}(\mathbb{I} \otimes u, u) \;=\; -\beta_k \alpha_{i-1} \frac{d}{du} f(u, \psi(u, g(u))) + O(h). \tag{4.9}$$

Formulae (4.7), (4.9) say that (V1) is satisfied with $\alpha_i' = \alpha_i$, $\beta_i' = -\beta_k \alpha_i$, $i = 0, \ldots, k-1$ and $\beta_k' = 0$, $\alpha_k' = \alpha_k = 1$.
Since the conditions (V1)-(V3) are fulfilled for the u-component of the discrete scheme (3.17), Lemma 4.2 applies and ensures

$$\lim_{n\to\infty} u_n \;=\; \bar{u} \text{ for some } \bar{u} \text{ satisfying } f(u, \psi(\bar{u}, g(\bar{u}))) = 0.$$

In addition, $\lim_{n\to\infty} v_n = \lim_{n\to\infty} g(u_n) = 0$ follows with (2.8). For the λ-component we can compute

$$\lim_{n\to\infty} \lambda_{n+k} \;=\; \lim_{n\to\infty} \Lambda(h, u_n, u_{n+1}, \ldots, u_{n+k-1}) \;=\; \Lambda(h, \bar{u}, \ldots, \bar{u}) \;=\; \psi(\bar{u}, 0)$$

and $(\bar{u}, 0, \psi(\bar{u}, 0))$ is an equilibrium of equation (2.1), (2.3).

5 Some properties of nonlinear multistep methods

In this section we give a proof of Lemma 4.2. The first step is to generalize Theorem 4.1 in Schropp [13] to the class of general multistep methods (4.4) satisfying (V1)-(V3).
We consider an autonomous ordinary differential equation $\dot{u} = F(u)$ of class C^r, $r \geq 2$ in a neighbourhood of a hyperbolic equilibrium $\bar{u}$. Let the sequence $(u_n)_{n\in\mathbb{N}}$ be generated via a strictly stable multistep method (4.4) with starting values $z_0 = (u_0, u_1, \ldots, u_{k-1})$ applied to $\dot{u} = F(u)$. In smooth dynamical systems with solution flow $\phi(t, u_0)$ the local stable manifold at $\bar{u}$ with respect to a neighbourhood V of $\bar{u}$ in $\mathbb{R}^N$ is defined by

$$W_s^V(\bar{u}) \;:=\; \{u \in V \mid \; \phi(t,u) \in V,\; t \geq 0 \text{ and } \phi(t,u) \to \bar{u} \text{ as } t \to \infty\}.$$

Additionally, if we define

$$B_s^V(\bar{u}) \;:=\; \{u \in V \mid \; \phi(t,u) \in V \text{ for } t \geq 0\},$$

the Hartman-Grobman Lemma shows the relation $W_s^V(\bar{u}) = B_s^V(\bar{u})$ for V sufficiently small. In the discrete multistep case (4.4) the analogous definitions read

$$\begin{aligned} W_{s,h}^V(\bar{u}) &:= \{z_0 \in V^k \mid \; u_n \in V,\; n \in \mathbb{N},\; u_n \to \bar{u} \text{ as } n \to \infty\}, \\ B_{s,h}^V(\bar{u}) &:= \{z_0 \in V^k \mid \; u_n \in V \text{ for } n \in \mathbb{N}\}. \end{aligned}$$

If $K_r(w)$ stands for the ball with center w and radius r in the appropriate normed space, the generalization of Theorem 4.1, Schropp [13] reads as follows.

Lemma 5.1 *Let $\bar{u}$ be a hyperbolic equilibrium of $\dot{u} = F(u)$ and let the iterates u_n be generated by applying a strictly stable multistep method (4.4), (4.5) with constant step size $h \in]0, h_0]$ onto $\dot{u} = F(u)$. In addition, we assume that (V1)-(V3) hold for the discrete dynamics (4.4). Then $\delta > 0$, $h_1 \in]0, h_0]$ exist such that*

$$W_{s,h}^{K_\delta(\bar{u})}(\bar{u}) \;=\; B_{s,h}^{K_\delta(\bar{u})}(\bar{u}),\;\; 0 < h \leq h_1$$

is satisfied.

Proof: We rewrite the multistep method (4.4) as a one-step method in $\mathbb{R}^{Nk}$. Defining $z^n = (u_n, \ldots, u_{n+k-1})$ the iteration scheme reads

$$z^{n+1} \;=\; (A \otimes I) z^n + h(e_k \otimes I) f_h(z^n, U(h, z^n)) \;=:\; G_h(z^n) \tag{5.1}$$

with

$$A \;=\; \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ -\alpha_0' & -\alpha_1' & \ldots & -\alpha_{k-2}' & -\alpha_{k-1}' \end{pmatrix} \in \mathbb{R}^{k,k}. \tag{5.2}$$

Here, $U(h, z^n)$ denotes the solution of the equation

$$U \;=\; (-\alpha_0' I, \ldots, -\alpha_{k-1}' I) z^n + h f_h(z^n, U). \tag{5.3}$$

The reader may recall that equation (5.3) possesses a unique solution for $0 < h < h_0$ since f_h is globally lipschitzian with respect to the last variable.
Now, let $\bar{z} = \mathbb{I} \otimes \bar{u}$, $F(\bar{u}) = 0$ be a fixed point of the multistep iteration. We will show that

$$\begin{aligned} DG_h(\bar{z}) \;&=\; \begin{pmatrix} 0 & I & & & \\ & 0 & I & & \\ & & \ddots & \ddots & \\ & & & 0 & I \\ -D_{hk}^{-1} D_{h0} & -D_{hk}^{-1} D_{h1} & \ldots & -D_{hk}^{-1} D_{hk-2} & -D_{hk}^{-1} D_{hk-1} \end{pmatrix} + O(h^2) \\ &=:\; Q_h + O(h^2) \end{aligned} \tag{5.4}$$

holds with $D_{hj} = \alpha_j' I - h\beta_j' DF(\bar{u})$, $j = 0, \ldots, k$.
Differentiation of (5.1) directly shows the first $(k-1)$-block rows of the jacobian of $G_h(\bar{z})$. In order to compute the remaining k-block row, we combine the formulae (5.1), (5.3) and see

$$U(h, u_n, \ldots, u_{n+k-1}) \;=\; (G_h((u_n, \ldots, u_{n+k-1})))_k.$$

We differentiate the defining equation (5.3) for U with respect to u_{n+i}. This yields for $i = 0, \ldots, k-1$

$$\begin{aligned} \frac{\partial}{\partial u_{n+i}} U(h, z) + \alpha_i' I \;&=\; h\Big(\frac{\partial}{\partial u_{n+i}} f_h(z, U(h, z)) \\ &\quad + \frac{\partial}{\partial u_{n+k}} f_h(z, U(h, z)) \frac{\partial}{\partial u_{n+i}} U(h, z) \Big). \end{aligned} \tag{5.5}$$

Then, for $z = \mathbb{I} \otimes \bar{u}$ we obtain with $U(h, \mathbb{I} \otimes \bar{u}) = \bar{u}$ and

$$\frac{\partial}{\partial u_{n+i}} f_h(\mathbb{I} \otimes \bar{u}, \bar{u}) \;=\; \beta_i' DF(\bar{u}) + O(h),\ i = 0, \ldots, k$$

(compare (V1)) the relation

$$(I - h\beta_k' DF(\bar{u}) + O(h^2)) \frac{\partial U(h, \mathbb{I} \otimes \bar{u})}{\partial u_{n+i}} \;=\; -\alpha_i' I + h\beta_i' DF(\bar{u}) + O(h^2). \tag{5.6}$$

Using the definition of D_{hj} as well as $(D_{hk} + O(h^2))^{-1} = D_{hk}^{-1} + O(h^2)$ we can calculate

$$\frac{\partial}{\partial u_{n+j}} U(h, \mathbb{I} \otimes \bar{u}) \;=\; -D_{hk}^{-1} D_{hj} + O(h^2),\ j = 0, \ldots, k-1$$

and formula (5.4) is shown.
In what follows we mimic the proof of Theorem 4.1 in Schropp [13]. To simplify the notation we assume without loss of generality $\bar{z} = 0$. Let Z_s^h, Z_u^h denote the stable and unstable subspace of $DG_h(\bar{z})$, and let S_0, S_b denote the Banach spaces of zero convergent, respectively, bounded $\mathbb{R}^{Nk}$-valued sequences $(z_n)_{n\in\mathbb{N}}$ with the norm

$$\|\,(z_n)_{n\in\mathbb{N}}\,\|_\infty := \sup\{\|\,z_n\,\| \mid n \in \mathbb{N}\}.$$

Consider the scaled and cut-off vector field

$$\hat{G}_h(z) \;:=\; Q_h + R_h(z),\; R_h(z) \;:=\; \chi(z)\frac{1}{\delta}(G_h(\delta z) - Q_h\delta z)$$

for $0 < h \le h_0$ and scaling factor $\delta > 0$. Here $\chi \in C_b^\infty(\mathbb{R}^{Nk}, [0,1])$ is a cut-off function satisfying

$$\chi(z) = 1,\ \text{if}\ \|\,z\,\| \le 1, \qquad \chi(z) = 0,\ \text{if}\ \|\,z\,\| \ge 2.$$

For the transformed map $\hat{G}_h$, we define the operators

$$\begin{aligned} \Gamma_i^h : S_i &\to Z_s^h \times S_i \\ (z_n)_{n\in\mathbb{N}} &\to ([z_0]_s, (z_{n+1} - \hat{G}_h(z_n))_{n\in\mathbb{N}}) \end{aligned}$$

for $i \in \{0, b\}$. The aim is to apply the Lipschitz inverse mapping Theorem (see the Appendix of Irwin [11]) to the equations

$$\Gamma_i^h((z_n)_{n\in\mathbb{N}}) \;=\; (v, (0)_{n\in\mathbb{N}}),\ v \in Z_s^h$$

simultaneously and with the same data for $i = 0$ and $i = b$. Since $S_b \subset S_0$ our conclusion then follows from the uniqueness of the solutions. The main ingredient for the application of the Lipschitz inverse mapping theorem is the estimate

$$\|\,DR_h(z)\,\| \;\le\; \frac{\mu h}{4} \text{ for } 0 < h < h_0,\ z \in \mathbb{R}^{Nk} \tag{5.7}$$

(see Schropp [13], p.94). Using $\chi(z) = 0$ for $\|\,z\,\| \ge 2$ and (5.4), we obtain

$$\begin{aligned} \|\,DR_h(z)\,\| \;\le\; & \sup\{\|\,D\chi(z)\,\| \mid \|\,z\,\| \le 2\} \\ & * \sup\{\|\,\frac{1}{\delta}(G_h(\delta z) - G_h(0) - DG_h(0)\delta z)\,\| \mid \|\,z\,\| \le 2\} \\ & + \sup\{\|\,DG_h(\delta z) - DG_h(0)\,\| \mid \|\,z\,\| \le 2\} + O(h^2) \\ \le\; & \frac{\mu h}{4}, \quad z \in \mathbb{R}^{Nk} \end{aligned}$$

and (5.7) is shown.
From now on we follow the proof of Theorem 4.1 in Schropp [13]. This finishes the proof of Lemma 5.1.
Our next step is the generalization of Lemma A.1, Schropp [16] from linear to general multistep methods satisfying (V1)-(V3).

Lemma 5.2 *Let the iterates u_n be generated by applying a pth order strictly stable multistep method of the form (4.4), (4.5) onto a smooth initial value problem $\dot{u} = F(u)$, $u(0) = u_0$ with solution flow $\phi(t, u_0)$. We assume that the multistep method satisfies (V1)-(V3) and that F is globally lipschitzian. Then*

$$\| u_{n+k} - \phi(h, u_{n+k-1}) \| \leq Ch^{p+1}, \quad n \in \mathbb{N}$$

holds.

Proof: First we rewrite the multistep method (4.4), (4.5) as a one-step method in $\mathbb{R}^{Nk}$, that is, in the equivalent form (5.1)-(5.3). A careful inspection of the proof of Lemma A.1 in Schropp [16] shows that it is sufficient to establish

$$\begin{aligned} &U(h, u_n, \ldots, u_{n+k-1}) - U(h, \phi((1-k)h, u_{n+k-1}), \ldots, \phi(-h, u_{n+k-1}), u_{n+k-1}) \\ &= \sum_{i=0}^{k-2} (\alpha_i' I + O(h))(u_{n+k-1} - \phi((k-1-i)h, u_{n+i})) \end{aligned} \tag{5.8}$$

(compare formula (A.7) in Schropp [16]). Provided (5.8) is shown a proof of Lemma 5.2 can be obtained along the lines of the proof of Lemma A.1 in Schropp [16].
Let $\hat{U} = g(h, \phi(-(k-1)h, u_{n+k-1}), \ldots, \phi(-h, u_{n+k-1}), u_{n+k-1})$. Using formula (5.3) we can compute

$$\begin{aligned} U - \hat{U} &= \sum_{i=0}^{k-2} -\alpha_i'(u_{n+i} - \phi((-(k-1)+i)h, u_{n+k-1})) \\ &\quad + h\Big(\sum_{i=0}^{k-2} -\Delta_i \alpha_i'(u_{n+i} - \phi((-(k-1)+i)h, u_{n+k-1})) + \Delta_k (U - \hat{U}) \Big). \end{aligned} \tag{5.9}$$

Here, the matrices Δ_i, $i = 0, \ldots, k$ arise from an application of the mean value theorem. Extracting $U - \hat{U}$ in (5.9) yields

$$(I + O(h))(U - \hat{U}) = \sum_{i=0}^{k-2} (-\alpha_i' I + O(h))(u_{n+i} - \phi((-(k-1)+i)h, u_{n+k-1}))$$

Then, with $(I + O(h))^{-1} = I + O(h)$ the representation

$$U - \hat{U} = \sum_{i=0}^{k-2} (-\alpha_i' I + O(h))(u_{n+i} - \phi((-(k-1)+i)h, u_{n+k-1})) \tag{5.10}$$

follows. Next the reader should recall the flow property $\phi(t, y) - \phi(t, x) = (I + O(t))(y - x)$. Using this property, we can calculate for $i = 0, \ldots, k-1$

$$\begin{aligned} (u_{n+i} - \phi((-(k-1)+i)h, u_{n+k-1})) &= (-I + O(h)) \\ &\quad \cdot (u_{n+k-1} - \phi(((-k+1)+i)h, u_{n+i})). \end{aligned} \tag{5.11}$$

Finally, inserting (5.11) into (5.10) gives the desired result (5.8).

Proof of Lemma 4.2: Lemmata 5.1, 5.2 as well as

$$\| u_{n+k} - u_{n+k-1} \| \leq Ch \quad \forall n \in \mathbb{N} \tag{5.12}$$

(see (3.6)) enable us now to prove Lemma 4.2 by adapting the proof of Theorem 2.4, Schropp [15] in the multistep case. Indeed, a careful inspection of this proof shows that the special structure of the linear k-step method is merely used to establish the statements of the Lemmata 5.1, 5.2 and (5.12). Thus, following the lines of the proof of Theorem 2.4, Schropp [15] shows Lemma 4.2.

References

[1] Beyn, W.-J., *On the Numerical Approximation of Phase Portraits near Stationary Points*, SIAM J. Numer. Anal., **24** (1987), pp. 1095–1113.

[2] Beyn, W.-J., Schropp, J., *Runge-Kutta discretizations of singularly perturbed gradient equations*, BIT Numer. Mathematics, **40** (2000), pp. 415–433.

[3] Beyn, W.-J., Wadepuhl, M., *Computer Simulation of the Hydrostatic Sceleton, The Physical Equivalent, Mathematics and Applications to Wormlike Forms*, J. Theor. Biol., **136** (1989), pp. 379–402.

[4] Bracken, J., Mc Cormick, G.P., *Selected Applications of Nonlinear Programming*, (1968), Wiley, New York.

[5] Coppel, W.A., *Stability and asymptotic behaviour of differential equations*, (1965), D.C. Heath and Company, Boston Englewood Chicago.

[6] Gear, C.W., Gupta, G.K, Leimkuhler, B., *Automatic integration of Euler-Lagrange equations with constraints*, J. Comp. Appl. Math., Vol. **12** & **13** (1985), pp. 77–90.

[7] Hairer, E., Brasey, V., *Half-explicit Runge-Kutta methods for Differential-Algebraic Systems of index 2*, SIAM J. of Numer. Anal., **30** (1993), pp. 538–552.

[8] Hairer, E., Lubich, Ch., Roche, M., *Error of Runge-Kutta methods for stiff problems studied via differential algebraic equations*, BIT Numer. Math., **28** (1989), pp. 678–700.

[9] Hairer, E., Wanner, G., *Solving Ordinary Differential Equations II*, second edition (1996), Springer.

[10] Humphries, A.R., Stuart, A.M., *Runge-Kutta methods for dissipative and gradient systems*, SIAM J. Numer. Anal., **31** (1994), pp. 1452–1485.

[11] Irwin, M.C., *Smooth dynamical systems*, (1980), Academic Press, London and New York.

[12] Schropp, J., *Using dynamical systems methods to solve minimization problems*, Appl. Num. Math., **18** (1995), pp. 321–335.

[13] Schropp, J., *A note on minimzation problems and multistep methods*, Numer. Math., **78** (1997), pp. 87–101.

[14] Schropp, J., *A dynamical systems approach to constrained minimization*, Numer. Funct. Anal. and Optimiz., **21** (2000), pp. 537–551.

[15] Schropp, J., *One- and Multistep procedures for constrained minimization problems*, IMA J. of Numer. Anal., **20** (2000), pp. 135–152.

[16] Schropp, J., *Conserving first integrals under discretization with variable step size integration procedures*, J. of Comp. and Appl. Math., **115** (2000), pp. 503–517.

[17] Schropp, J., *Geometric Properties of Runge-Kutta Discretizations for Index 2 Differential-Algebraic Systems*, Konstanzer Schriften in Mathematik und Informatik, Nr. **128** (2000).

[18] Schropp, J. , *Ordinary Differential Equations, Differential Algebraic Equations and their use in Optimization*, Proceedings of the International Conference on Differential Equations, EQUADIFF-99, Vol. **2** (2000), pp. 928–933.

[19] Stuart, A.M., Humphries, A.R., *Dynamical Systems and Numerical Analysis*, (1996), Cambridge University Press.