Stability of Computational Methods for Constrained Dynamics Systems

by

TR 91-3

Uri Ascher

and

Linda Petzold

Technical Report 91-3
April, 1991

Department of Computer Science
University of British Columbia
Vancouver, B.C.
CANADA V6T 1W5

Stability of computational methods for constrained dynamics systems

Uri M. Ascher*
Department of Computer Science
University of British Columbia
Vancouver, British Columbia
Canada V6T 1W5

Linda R. Petzold†
Computing & Mathematics Research Division
Lawrence Livermore National Laboratory, L-316
Livermore, California 94550

May 1, 1991

Abstract

Many methods have been proposed for numerically integrating the differential-algebraic systems arising from the Euler-Lagrange equations for constrained motion. These are based on various problem formulations and discretizations. We offer a critical evaluation of these methods from the standpoint of stability.

Considering a linear model, we first give conditions under which the differential-algebraic problem is well-conditioned. This involves the concept of an essential underlying ODE. We review a variety of reformulations which have been proposed in the literature and show that most of them preserve the well-conditioning of the original problem. Then we consider stiff and nonstiff discretizations of such reformulated models. In some cases, the same implicit discretization may behave in a very different way when applied to different problem formulations, acting as a stiff integrator on some formulations and as a nonstiff integrator on others. We present the approach of projected invariants as a method for yielding problem reformulations which are desirable in this sense.

*The work of this author was partially supported under NSERC Canada Grant OGP0004306.

†The work of this author was partially supported by the Applied Mathematical Sciences subprogram of the Office of Energy Research, U.S. Department of Energy, by Lawrence Livermore National Laboratory under contract W-7405-Eng-48.

1 Introduction

Various techniques have been proposed in the literature for the numerical solution of the Euler-Lagrange equations which govern the motion of mechanical systems with constraints[15]. Several of these techniques are used in commercial codes. The equations to be solved form a system of second-order ordinary differential equations (ODEs) for the (generalized) multibody coordinates. They also involve Lagrange multiplier functions, and are subject to constraints, e.g. on configuration and/or motion. Mathematically, this may be considered as a system of differential-algebraic equations (DAEs) of index 3 in a semi-explicit, Hessenberg form [4]. It is well-known that a direct discretization of such a DAE yields numerical difficulties; this is what gives rise to a multitude of other, more specific solution techniques.

Typically, such a solution technique consists of a step of problem reformulation which involves reducing its index, followed by a discretization of the resulting formulation. In recent work [6, 7] it has been shown that for a certain model problem, some of these formulations can be equivalent. An important consideration in selecting an appropriate solution method (i.e. a combination of formulation and discretization) is the stability of the method and the subsequent stability restrictions that a chosen step size must satisfy. In this paper we investigate the stability of various solution techniques.

In order to be more specific, we write down the Euler-Lagrange equations for a constrained multibody system

$$\begin{aligned} M(\mathbf{p})\mathbf{p}'' &= \mathbf{f}(\mathbf{p},\mathbf{v}) - G^T(\mathbf{p})\lambda - \hat{G}^T(\mathbf{p})\hat{\lambda} && \text{(1.1a)} \\ \mathbf{0} &= \mathbf{g}(\mathbf{p}) && \text{(1.1b)} \\ \mathbf{0} &= \hat{G}(\mathbf{p})\mathbf{v} + \hat{\mathbf{g}}(\mathbf{p}) && \text{(1.1c)} \end{aligned}$$

Here the unknowns are $\mathbf{p}$ the generalized coordinates, $\mathbf{v} \equiv \frac{d\mathbf{p}}{dt} \equiv \mathbf{p}'$ the generalized velocities, and λ and $\hat{\lambda}$ the Lagrange multiplier functions. In (1.1a) M is the mass matrix ($M(\mathbf{p})(t) \in \mathcal{R}^{n_p \times n_p}$ is symmetric positive definite), $\mathbf{f}$ stands for the applied forces and $G(\mathbf{p})$ is the Jacobian matrix of the holonomic constraints

$$G(\mathbf{p}) = \mathbf{g_p}; \qquad G(\mathbf{p})(t) \in \mathcal{R}^{n_\lambda \times n_p} \tag{1.2}$$

In (1.1b) there are n_λ configuration (position) constraints and in (1.1c) there are $n_{\hat{\lambda}}$ motion or other constraints. For simplicity of presentation we shall often assume that either $n_{\hat{\lambda}} = 0$ (hence $\hat{\lambda}$ disappears from (1.1a) as well), i.e. that there are only holonomic constraints, or that $n_\lambda = 0$ (whence λ disappears from (1.1a)).

We assume that the constraints (1.1b) are independent in the sense that G has a full row rank $n_\lambda (\leq n_p)$. Then clearly two differentiations of the constraints (1.1b) allow elimination of λ from (1.1a). Thus the original DAE has index 3. On the other hand, if $n_\lambda = 0$ and $n_{\hat{\lambda}} > 0$ with $\hat{G}$ having a full row rank, then only one

differentiation of (1.1c) is needed to eliminate $\hat{\lambda}$ and obtain an ODE, so the index is 2. Both of these cases can be cast in the form (2.1) below with $m = 2$ and $m = 1$ (for the equivalent 1st order form of (1.1a)), respectively.

In order to give a methodical stability discussion we proceed in stages and consider the linearized form of the DAE (1.1). Since a nonlinear problem behaves like its linear variational form away from singularities (i.e. in a neighborhood of an isolated solution), our arguments will be valid in these general circumstances. We assume that the given linear DAE problem is well-conditioned, and in Section 2 specify precisely what this means using a constructed *essential underlying ODE*. The theory includes the linearizations of (1.1) as special cases.

In Section 3 we then consider a variety of problem reformulations and show that they are well-conditioned too under certain reasonable assumptions. We cover the Baumgarte stabilization technique, a variety of stabilized and unstabilized index reductions and transformations to state space form. This allows us to consider in Section 4 discretizations of the various formulations.

We consider stiff and nonstiff discretizations of such reformulated models. In some cases, the same BDF discretization (Backward Differentiation Formula, see e.g. [4]), or other stiff discretizations, may behave in a very different way when applied to different reformulations of the same problem, acting as a stiff integrator on some formulations and as a nonstiff integrator on others. The need to restrict the stepsize in BDF for numerical stability arises even in formulations which explicitly enforce the constraints. For (1.1), assuming say that there are only position constraints which vary on the scale of the solution, such a situation may arise for instance if the mass matrix M has both large and nonlarge eigenvalues, in which case $M^{-1}G^T$ may be a much less pleasant function than G. (A corresponding physical situation is a heterogeneous multibody system, i.e. a system which includes bodies with very different masses.[1]) We present the approach of *projected invariants* with a particular choice of the projection, as a method for yielding problem reformulations which are desirable in this sense. Section 5 concludes with a summary and recommendations based on our results.

Throughout this paper we use the following notation: Let $|\cdot|$ be the Euclidean vector norm. For a matrix A we denote the induced matrix norm by $\|A\|$. For a function $\mathbf{u}(t)$, $0 \le t \le 1$, we denote the corresponding max function norm by $\|\mathbf{u}\| := \max\{|\mathbf{u}(t)|, 0 \le t \le 1\}$.

2 Problem conditioning

The DAE of order m

$$\mathbf{x}^{(m)} = \mathbf{f}(\mathbf{z}(\mathbf{x}), \mathbf{y}, t) \tag{2.1a}$$

$$\mathbf{0} = \mathbf{g}(\mathbf{x}, t) \tag{2.1b}$$

[1] We thank Dr. Dan Rosenthal of RASNA Corp. for illuminating us on this point.

where $\mathbf{x}^{(j)}(t) := \frac{d^j \mathbf{x}(t)}{dt^j}$ and

$$\mathbf{z}(\mathbf{x}) = (\mathbf{x}, \mathbf{x}', ..., \mathbf{x}^{(m-1)})^T, \tag{2.2}$$

has index $m+1$ if $\mathbf{g_x f_y}$ is nonsingular for all t, $0 \le t \le 1$. The Euler-Lagrange equations for dynamical systems with holonomic constraints are in this form with $m = 2$, $\mathbf{x}$ the generalized coordinates and $\mathbf{y}$ the Lagrange multipliers. Here we consider the linear (or linearized) form

$$\mathbf{x}^{(m)} = \sum_{j=1}^{m} A_j \mathbf{z}_j + B\mathbf{y} + \mathbf{q} \tag{2.3a}$$

$$\mathbf{0} = C\mathbf{x} + \mathbf{r} \tag{2.3b}$$

where A_j, B and C are smooth functions of t, $0 \le t \le 1$, $A_j(t) \in \mathcal{R}^{n_x \times n_x}$, $j = 1, .., m$, $B(t) \in \mathcal{R}^{n_x \times n_y}$, $C(t) \in \mathcal{R}^{n_y \times n_x}$, $n_y \le n_x$ and CB is nonsingular for each t (hence the DAE has index $m+1$). All matrices involved are assumed to be uniformly bounded in norm by a constant of moderate size. The inhomogeneities are $\mathbf{q}(t) \in \mathcal{R}^{n_x}$ and $\mathbf{r}(t) \in \mathcal{R}^{n_y}$.

We derive a stability result for this system. As in [1], there exists a smooth, bounded matrix function $R(t) \in \mathcal{R}^{(n_x - n_y) \times n_x}$ whose linearly independent rows form a basis for the nullspace of B^T (R can be taken to be orthonormal). Thus, for each t, $0 \le t \le 1$,

$$RB = 0. \tag{2.4}$$

We assume that there exists a constant K_1 of moderate size such that

$$\|(CB)^{-1}\| \le K_1 \tag{2.5}$$

uniformly in t, and obtain (Lemma 2.1 in [1]) that there is a constant K_2 of moderate size such that

$$\left\| \begin{pmatrix} R \\ C \end{pmatrix}^{-1} \right\| \le K_2. \tag{2.6}$$

The constant K_2 depends, in addition to K_1, also on $\|B\|$, $\|C\|$ and $\|R\|$. Let K_3 be a moderate bound on R and its derivatives:

$$\|R^{(j)}\| \le K_3 \qquad j = 0, 1, \ldots, m. \tag{2.7}$$

Define new variables

$$\mathbf{u} = R\mathbf{x}, \qquad 0 \le t \le 1. \tag{2.8}$$

Then, using (2.3b), the inverse transformation is given by

$$\mathbf{x} = \begin{pmatrix} R \\ C \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{u} \\ -\mathbf{r} \end{pmatrix} \equiv S\mathbf{u} - F\mathbf{r} \tag{2.9}$$

where $S(t) \in \mathcal{R}^{n_x \times (n_x - n_y)}$ satisfies

$$RS = I, \qquad CS = 0 \tag{2.10}$$

and

$$F := B(CB)^{-1}. \tag{2.11}$$

By our assumptions and (2.6) this mapping is well-conditioned. Both S and F are smooth and bounded. The first m derivatives of S and F are bounded by a constant involving K_2 and K_3. Taking m derivatives of (2.8) yields

$$\mathbf{u}^{(m)} = (R\mathbf{x})^{(m)} = \sum_{j=1}^{m} [RA_j + \binom{m}{j-1} R^{(m-j+1)}]\mathbf{z}_j + R\mathbf{q} \tag{2.12}$$

Further, using $m-1$ derivatives of (2.9) we obtain the *essential underlying ODE* (EUODE)

$$\mathbf{u}^{(m)} = \sum_{j=1}^{m} [RA_j + \binom{m}{j-1} R^{(m-j+1)}][(S\mathbf{u})^{(j-1)} - (F\mathbf{r})^{(j-1)}] + R\mathbf{q} \tag{2.13}$$

For a unique solution of (2.3) one needs to impose $m(n_x - n_y)$ independent boundary conditions

$$B_0\mathbf{z}(0) + B_1\mathbf{z}(1) = \beta. \tag{2.14}$$

These could be, for instance, initial conditions which, together with equation (2.3b) and its first $m-1$ derivatives all sampled at $t = 0$, form mn_x initial conditions which specify $\mathbf{z}(0)$. The boundary conditions can be written as $m(n_x - n_y)$ conditions on $\mathbf{u}$ and its first $m-1$ derivatives needed to specify a unique solution for the EUODE (2.13). If this ODE problem is stable, i.e. if Green's function $\mathcal{G}(t,s)$ and its first $m-1$ derivatives in t are bounded in norm by a constant of moderate size, say K_4, then a similar conclusion holds for the DAE. We obtain the following theorem:

Theorem 2.1 *Let the DAE (2.3) have smooth, bounded coefficients, and assume that (2.5) holds and that the underlying problem for (2.13) is stable. Then there is a constant K of moderate size such that*

$$\|\mathbf{z}\| \leq K(\|\mathbf{q}\| + \sum_{j=0}^{m-1} \|\mathbf{r}^{(j)}\| + |\beta|) \tag{2.15a}$$

$$\|\mathbf{y}\| \leq K(\|\mathbf{q}\| + \sum_{j=0}^{m} \|\mathbf{r}^{(j)}\| + |\beta|) \tag{2.15b}$$

Proof:

Our assumptions guarantee the well-conditioning of the transformation from $\mathbf{x}$ to $\mathbf{u}$ and back. The boundary data for $\mathbf{u}$ is therefore bounded by $\sum_{j=0}^{m-1} \|\mathbf{r}^{(j)}\| + |\beta|$

times a moderate constant. We may write $\mathbf{u}(t)$ in terms of Green's function $\mathcal{G}(t,s)$, differentiate $m-1$ times and take norms, obtaining

$$\|\mathbf{u}^{(l)}\| \le \tilde{K}(\|\mathbf{q}\| + \sum_{j=0}^{m-1} \|\mathbf{r}^{(j)}\| + |\beta|) \qquad 0 \le l \le m-1$$

with $\tilde{K}$ a moderate constant depending on K_2, K_3 and K_4. Conclusion (2.15a) is then obtained using (2.9).

Now, given $\mathbf{x}$ we obtain $\mathbf{y}$ through multiplying (2.3a) by C, yielding

$$\mathbf{y} = (CB)^{-1}C(\mathbf{x}^{(m)} - \sum_{j=1}^{m} A_j \mathbf{z}_j - \mathbf{q}). \tag{2.16}$$

Differentiating (2.3b) m times we substitute for $C\mathbf{x}^{(m)}$ in (2.16), and using (2.15a) obtain the bound (2.15b). □

Remark:

The EUODE (2.13) is nonunique. For any nonsingular, smooth, bounded transformation $T(t) \in \mathcal{R}^{(n_x-n_y)\times(n_x-n_y)}$, the transformed $R(t)$ given by

$$R \leftarrow TR \tag{2.17}$$

still satisfies (2.4), (2.6) and (2.7). Hence R is unique only up to such a transformation and, correspondingly, so is the EUODE. However, a transformation of the variables $\mathbf{u}$ in (2.8) corresponding to (2.17) does not alter the boundedness (or lack thereof) of the Green's function, and hence the stability properties are properly reflected in Theorem 2.1. For later theoretical purposes, we may wish to choose T such that the EUODE (2.13) is amenable to a direct discretization. In particular, for $m=1$ and a BDF discretization we can choose T so that the resulting matrix $(RA_1 + R')S$ is essentially diagonally dominant or block upper triangular (see [10], [11], [2]). □

We remark that a bound similar to (2.15) may also be obtained using Theorem 2.1 of [1] applied to the index-2 DAE

$$\begin{aligned}
\mathbf{z}_j' &= \mathbf{z}_{j+1} + B\mu_j, \qquad j = 1,\ldots,m-1 && \text{(2.18a)}\\
\mathbf{z}_m' &= \sum_{j=1}^{m} A_j\mathbf{z}_j + B\mathbf{y} + \mathbf{q} && \text{(2.18b)}\\
\mathbf{0} &= C\mathbf{z}_1 + \mathbf{r} && \text{(2.18c)}\\
\mathbf{0} &= C\mathbf{z}_2 + C'\mathbf{z}_1 + \mathbf{r}' && \text{(2.18d)}\\
&\vdots\\
\mathbf{0} &= \sum_{j=1}^{m} \binom{m-1}{j-1} C^{(m-j)}\mathbf{z}_j + \mathbf{r}^{(m-1)} && \text{(2.18e)}
\end{aligned}$$

subject to the original boundary conditions (2.14). Here we have applied a particular *stabilized index reduction* technique [8], adding multiplier functions $\mu_j(t) \in \mathcal{R}^{n_\nu}$ to compensate for insisting that the constraint (2.3b) and its first $m-1$ derivatives be all satisfied at all t. This DAE problem has the same exact solution as the original higher index problem (2.3), (2.14) because differentiation and substitution for each of the algebraic constraints in (2.18) yields $CB\mu_j = 0$, which implies $\mu_j = 0$, $j = 1, \ldots, m-1$. The EUODE for (2.18) is obtained using R of (2.4) m times, i.e. for the variables

$$\mathbf{w}_j := R\mathbf{z}_j = R\mathbf{x}^{(j-1)}, \qquad j = 1, \ldots, m, \tag{2.19}$$

we obtain

$$\mathbf{w}_j' = R\mathbf{z}_{j+1} + R'\mathbf{z}_j, \qquad j = 1, \ldots, m-1 \tag{2.20a}$$

$$\mathbf{w}_m' = \sum_{j=1}^{m} RA_j\mathbf{z}_j + R'\mathbf{z}_m + R\mathbf{q} \tag{2.20b}$$

where $\mathbf{z}_j$ are expressed in terms of $\mathbf{w}$ using the recursive relation

$$\mathbf{z}_j = S\mathbf{w}_j - F[\mathbf{r}^{(j-1)} + \sum_{l=1}^{j-1} \binom{j-1}{l-1} C^{(j-l)}\mathbf{z}_l] \qquad j = 1, \ldots, m. \tag{2.21}$$

It is easily shown that the stabilized index-2 form (2.18) with (2.14) is well-conditioned (or stable as an initial value problem) whenever the original high-index equation (2.3) is.

3 Other transformations

The EUODE (2.13) uses a minimal number of constraint differentiations. Therefore, we view the assumption that it is stable with the given boundary operator as essential. From this we now derive stability for a number of other problem reformulations which have appeared in the literature.

3.1 Baumgarte stabilization

The most straightforward transformation of the DAE (2.3) into an ODE involves replacing the constraint

$$\mathbf{g}(\mathbf{x}, t) \equiv C\mathbf{x} + \mathbf{r} = \mathbf{0}$$

with its m^{th} time derivative plus initial conditions:

$$\mathbf{g}^{(m)} = \frac{d^m \mathbf{g}(\mathbf{x}(t), t)}{dt^m} = \mathbf{0} \tag{3.1a}$$

$$\mathbf{g}(\mathbf{x}(0), 0) = \frac{d}{dt}\mathbf{g}(\mathbf{x}(0), 0) = \ldots = \frac{d^{m-1}}{dt^{m-1}}\mathbf{g}(\mathbf{x}(0), 0) = \mathbf{0} \tag{3.1b}$$

However, this causes well-known drift difficulties. A generalization of Baumgarte's method [3] replaces (3.1a) with the equation

$$\sum_{j=0}^{m} \alpha_j \frac{d^j}{dt^j} \mathbf{g}(\mathbf{x}(t), t) = \mathbf{0} \tag{3.2}$$

where α_j are chosen so that $\alpha_m = 1$ and the roots of the polynomial

$$\sigma(\tau) = \sum_{j=0}^{m} \alpha_j \tau^j \tag{3.3}$$

are all nonpositive. For instance, one may choose $\sigma(\tau) = (\tau + \gamma)^m$ for some $\gamma \geq 0$. We now investigate the stability of (2.3a), (3.2), (2.14) and (3.1b).

In (3.2) we have an expression for $C\mathbf{x}^{(m)}$ which we may substitute in (2.3a) multiplied by C and eliminate $\mathbf{y}$:

$$\mathbf{y} = -(CB)^{-1}\{\sum_{j=1}^{m}[CA_j + \binom{m}{j-1} C^{(m-j+1)}]\mathbf{z}_j + C\mathbf{q} + \mathbf{r}^{(m)} + \sum_{l=0}^{m-1} \alpha_l \mathbf{g}^{(l)}\} \tag{3.4}$$

Substituting back into (2.3a) we obtain an ODE for $\mathbf{x}$

$$\mathbf{x}^{(m)} = \sum_{j=1}^{m}[HA_j - \binom{m}{j-1} FC^{(m-j+1)}]\mathbf{z}_j + H\mathbf{q} - F\mathbf{r}^{(m)} - F\sum_{l=0}^{m-1} \alpha_l \mathbf{g}^{(l)} \tag{3.5a}$$

$$\mathbf{g}^{(l)} = \sum_{j=1}^{l+1} \binom{l}{j-1} C^{(l-j+1)}\mathbf{z}_j + \mathbf{r}^{(l)} \tag{3.5b}$$

with F given by (2.11) and H the projection

$$H = I - FC = SR. \tag{3.6}$$

We then ask the question regarding the stability of the ODE problem (3.5), (2.14), (3.1b).

To resolve this question, define

$$\mathbf{u} = R\mathbf{x}, \qquad \mathbf{v} = C\mathbf{x} \tag{3.7}$$

Then

$$\mathbf{x} = S\mathbf{u} + F\mathbf{v} \tag{3.8}$$

(cf. (2.9)). So, by (2.6),

$$\|\mathbf{x}\| \leq K_2(\|\mathbf{u}\| + \|\mathbf{v}\|).$$

To see what $\mathbf{u}$ and $\mathbf{v}$ satisfy, multiply (3.5a) by R and by C. This gives

$$\mathbf{u}^{(m)} = \sum_{j=1}^{m}[RA_j + \binom{m}{j-1} R^{(m-j+1)}][(S\mathbf{u})^{(j-1)} + (F\mathbf{v})^{(j-1)}] + R\mathbf{q} \tag{3.9a}$$

$$\mathbf{v}^{(m)} = -\sum_{j=0}^{m-1} \alpha_j \mathbf{v}^{(j)} + \sum_{j=0}^{m} \alpha_j \mathbf{r}^{(j)} \tag{3.9b}$$

$$\mathbf{v}^{(j)}(0) = -\mathbf{r}^{(j)}(0), \quad j = 0, \ldots, m-1 \tag{3.9c}$$

Now, in (3.9b), (3.9c) we have a uniformly stable initial value problem for $\mathbf{v}$ so long as at most one root of $\sigma(\tau)$ is 0 and the rest have negative real parts. For instance, with

$$\sigma(\tau) = (\tau + \gamma)^m \tag{3.10}$$

any choice of $\gamma > 0$ yields a uniformly, *asymptotically* stable problem for $\mathbf{v}$, while the choice $\gamma = 0$, which corresponds to using (3.1a) in place of (3.2), allows for a mild instability, viz. a polynomial error growth (of degree $m-1$), to occur. Moreover, if the EUODE is asymptotically stable and $\gamma > 0$ then the ODE (3.5) is also asymptotically stable.

Once $\mathbf{v}$ is integrated it may be substituted into (3.9a) to obtain the EUODE for $\mathbf{u}$. The problem for (3.9) is therefore stable, and by (3.8) so is the problem which the Baumgarte stabilization technique yields.

This analysis agrees with practical observations. First, the unstabilized index reduction (3.1) has only a mild instability for a well-conditioned original problem. This instability gets worse as m increases, i.e. it is worse for the DAE with position constraints (1.1a), (1.1b) than for the DAE with motion constraints (1.1a), (1.1c). Second, any $\gamma > 0$ in (3.10) yields a stable problem in (3.5). The difference in performance for different values of $\gamma > 0$, when there is any, is due to discretization effects applied to stable ODE problems. This is taken up in the Section 4, but we may already expect here that if K in (2.15) is indeed of moderate size and the discretization mesh is very fine, then the results will not be sensitive to the choice of γ. In practice, the choice of γ in sensitive situations is far from clear.

3.2 Reduction to index 2

In (3.1) and (3.2) we have differentiated $\mathbf{g}(\mathbf{x}, t)$ m times, reducing the index to 1. A subsequent elimination of $\mathbf{y}$ gives an ODE. If instead we differentiate the constraints only $m-1$ times, we obtain a DAE of index 2 consisting of (2.3a) and

$$\sum_{j=0}^{m-1} \hat{\alpha}_j \frac{d^j}{dt^j} \mathbf{g}(\mathbf{x}(t), t) = \mathbf{0} \tag{3.11}$$

with $\hat{\alpha}_{m-1} = 1$. This is subject to (2.14) and

$$\mathbf{g}(\mathbf{x}(0), 0) = \ldots = \frac{d^{m-2}}{dt^{m-2}} \mathbf{g}(\mathbf{x}(0), 0) = \mathbf{0} \tag{3.12}$$

The stability analysis for this problem formulation proceeds precisely as before: using the transformation (3.7), (3.8) we obtain (3.9a) and

$$\sum_{j=0}^{m-1} \hat{\alpha}_j \mathbf{v}^{(j)} = -\sum_{j=0}^{m-1} \hat{\alpha}_j \mathbf{r}^{(j)} \tag{3.13a}$$

$$\mathbf{v}^{(j)}(0) = -\mathbf{r}^{(j)}(0), \quad j = 0, \ldots, m-2 \tag{3.13b}$$

The ODE (3.13a) is asymptotically stable if the roots of $\hat{\sigma}(\tau) = \sum_{j=0}^{m-1} \hat{\alpha}_j \tau^j$ all have negative real parts. Considering in particular

$$\hat{\sigma}(\tau) = (\tau + \gamma)^{m-1} \tag{3.14}$$

there is asymptotic stability if $\gamma > 0$ and a polynomial growth of order $m-2$ if $\gamma = 0$. The latter corresponds to unstabilized index reduction. In particular, for mechanical systems with $m = 2$ one differentiation of the constraints without stabilization ($\gamma = \hat{\alpha}_0 = 0$) yields a stable, although not asymptotically stable, problem (3.13) for $\mathbf{v}$.

The stability of the index-2 problem (2.3a), (3.11), (2.14), (3.12) follows, as before, from that of (3.13) and the analysis of Section 2.

The justification for considering this type of index reduction is that certain implicit discretization schemes like BDF may already be successfully applied to the resulting formulation (cf. [4], [12], [1]). This is considered in Section 4.

Another problem reformulation which reduces the index to 2 is, of course, the stabilized index reduction of (2.18). In Section 3.5 below we consider an entire family of additional stabilized index reductions.

3.3 State space form

The problem formulations considered hitherto in this section all end up in an ODE of size mn_x, requiring supplementary boundary conditions. In contrast, the EUODE (2.13) only has size $m(n_x - n_y)$, and no supplementary conditions are required for the problem reformulation. Moreover, incorporation of the constraint (2.3b) and its first $m-1$ derivatives into the transformation has insured no drift in a subsequent discretization.

This can be done more generally: Let $\tilde{R}(t)$ be a smooth, bounded function, $\tilde{R}(t) \in \mathcal{R}^{(n_x-n_y)\times n_x}$ such that

$$\left\| \begin{pmatrix} \tilde{R} \\ C \end{pmatrix}^{-1} \right\| \le \tilde{K}, \qquad \|\tilde{R}^{(j)}\| \le \tilde{K} \qquad j = 0, 1, \ldots, m \tag{3.15}$$

for a constant $\tilde{K}$ of moderate size. (We do not require $\tilde{R}B = 0$.) Define

$$\tilde{\mathbf{u}} = \tilde{R}\mathbf{x}, \qquad \mathbf{x} = \begin{pmatrix} \tilde{R} \\ C \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{u}} \\ -\mathbf{r} \end{pmatrix} \equiv \tilde{S}\tilde{\mathbf{u}} - \tilde{F}\mathbf{r} \tag{3.16}$$

Taking $\alpha_0 = \ldots = \alpha_{m-1} = 0$ in (3.5a) (i.e. $\gamma = 0$ in (3.10)) we multiply it by $\tilde{R}$ to obtain

$$\begin{aligned} \tilde{\mathbf{u}}^{(m)} &= \sum_{j=1}^{m} [\tilde{R}HA_j + \binom{m}{j-1} (\tilde{R}^{(m-j+1)} - \tilde{R}FC^{(m-j+1)})]\mathbf{z}_j \\ &\quad + \tilde{R}H\mathbf{q} - \tilde{R}F\mathbf{r}^{(m)} \qquad (3.17a) \\ \mathbf{z}_j &= (\tilde{S}\tilde{\mathbf{u}})^{(j-1)} - (\tilde{F}\mathbf{r})^{(j-1)} \qquad 1 \le j \le m \qquad (3.17b) \end{aligned}$$

This state-space ODE is subject to the boundary conditions (2.14), suitably transformed. The EUODE is obtained as a special case with $\tilde{R} = R$.

The stability of the problem formulation (3.17) follows immediately upon relating $\tilde{\mathbf{u}}$ and $\mathbf{u}$ through $\mathbf{x}$, i.e. using (3.16), (2.8), (2.9) and their derivatives. The obtained stability bound depends on $\tilde{K}$ (and of course on K of (2.15)).

A favorite practical choice for $\tilde{R}$ is as a piecewise constant function [16], [14]. Thus, choosing $\tilde{R}$ at a certain reference time t_c so that (3.15) is satisfied, one proceeds to integrate in t holding this $\tilde{R}$ constant so long as (3.15) holds with a reasonable $\tilde{K}$. When (3.15) is deemed violated, a new constant matrix $\tilde{R}$ is chosen based on a new reference point, giving a different ODE (3.17). The segments are connected in such switching points through continuity of $\mathbf{z}$. The lack of nonzero derivatives of $\tilde{R}$ over the integrated segment gives (3.17) an attractive form. A robust detection scheme for the necessity to change $\tilde{R}$ may prove to be the more difficult aspect of such a procedure, as discussed in Section 4.

3.4 Overdetermined DAE

Consider deriving the EUODE from the 1^{st} order form

$$\mathbf{z}_j' = \mathbf{z}_{j+1} \qquad j = 1, \ldots, m-1 \tag{3.18a}$$

$$\mathbf{z}_m' = \sum_{j=1}^{m} A_j \mathbf{z}_j + B\mathbf{y} + \mathbf{q} \tag{3.18b}$$

We proceed to define $\mathbf{w}_j = R\mathbf{z}_j$ and obtain the back-transformation using

$$\mathbf{0} = \sum_{l=1}^{j} \binom{j-1}{l-1} C^{(j-l)} \mathbf{z}_l + \mathbf{r}^{(j-1)} \qquad j = 1, \ldots, m \tag{3.19}$$

The transformation matrix for each j is $\begin{pmatrix} R \\ C \end{pmatrix}$. If we now write down the equations to be satisfied (3.18), (3.19), they form an overdetermined DAE (ODAE). This overdetermination is subsequently resolved when multiplying (3.18a) and (3.18b) by R, obtaining the EUODE (2.20) in terms of

$$\mathbf{w} = (\mathbf{w}_1, \ldots, \mathbf{w}_m)^T. \tag{3.20}$$

The fact that (3.18), (3.19) is indeed an ODAE is reflected by the fact that we could replace (3.18a) by the "stabilized form"

$$\mathbf{z}_j' = \mathbf{z}_{j+1} + B\mu_j \tag{3.21}$$

and obtain precisely the same EUODE, as in (2.18). Note that the DAE (3.21), (3.18b), (3.19) is not overdetermined any more, and that we have not used the fact that $\mu_j = \mathbf{0}$.

The ODAE (3.18), (3.19) subject to (2.14) has a unique solution, but when we replace it by a discretized form using one of the conventional difference schemes, we cannot expect an exact solution to exist. Still, one can multiply the discretized (3.18) by R, and (3.21) can replace (3.18a) provided that R and B are sampled at the same point t. The discretized DAE (3.21), (3.18b), (3.19) is therefore equivalent to a particular projection for solving the problem of minimizing the residual of the discretized ODAE subject to satisfying the discretized (3.19) (cf. [6, 7]).

Similar arguments apply when replacing R by a more general smooth, bounded $\tilde{R}$ satisfying (3.15). Before applying $\tilde{R}$ we must differentiate the constraints once more and substitute (3.4) into (3.18b) to eliminate $\mathbf{y}$, obtaining

$$\mathbf{z}'_m = \sum_{j=1}^{m} [HA_j - \binom{m}{j-1} FC^{(m-j+1)}]\mathbf{z}_j + H\mathbf{q} - F\mathbf{r}^{(m)} \tag{3.22}$$

A particular projected solution for the discretized ODAE (3.18a), (3.19), (3.22) with (3.19) treated as constraints is then obtained from the same discretization applied to (3.17) written in 1st order form, and this in turn is equivalent to the discretized form of the DAE (3.22), (3.19) and

$$\mathbf{z}'_j = \mathbf{z}_{j+1} + \tilde{B}\mu_j \qquad j = 1, \ldots, m-1 \tag{3.23}$$

with $\tilde{B}(t) \in \mathcal{R}^{n_x \times n_y}$ having full rank and satisfying

$$\tilde{R}\tilde{B} = 0. \tag{3.24}$$

For other possibilities, see [6, 7]. Here we note that the stability treatment of the ODAEs we have described is covered by our previous stability analysis.

3.5 Projected invariants

Consider the following general procedure: Differentiating the constraints (2.3b) m times and eliminating $\mathbf{y}$ from (3.18b), we obtain (3.22) which, together with (3.18a), form an explicit ODE system for $\mathbf{z}$. However, this allows for unacceptable drifts in the constraints after discretization, so we reimpose the first k constraints of (3.19), for some integer $k \leq m$

$$\mathbf{0} = \sum_{l=1}^{j} \binom{j-1}{l-1} C^{(j-l)}\mathbf{z}_l + \mathbf{r}^{(j-1)} \qquad j = 1, \ldots, k. \tag{3.25}$$

The constraints (3.25) form an invariant of the ODE (3.18a),(3.24) (cf. Gear [9]). To satisfy these constraints even after discretization, we project the ODE as follows.

For k given smooth, full-rank bounded matrix functions $R_j(t) \in \mathcal{R}^{(n_x - n_y)\times n_x}$ satisfying (3.15) (for $\tilde{R} = R_j$), require that

$$R_j\mathbf{z}'_j \;=\; R_j\mathbf{z}_{j+1} \qquad j = 1, \ldots, \min(k, m-1) \tag{3.26a}$$

$$R_m \mathbf{z}'_m = R_m \{ \sum_{j=1}^{m} [HA_j - \binom{m}{j-1} FC^{(m-j+1)}] \mathbf{z}_j + H\mathbf{q} - F\mathbf{r}^{(m)} \} \quad (\text{if } k = m) \tag{3.26b}$$

This is equivalent to writing

$$\mathbf{z}'_j = \mathbf{z}_{j+1} + B_j \mu_j \qquad j = 1, \ldots, m-1 \tag{3.27a}$$

$$\mathbf{z}'_m = \sum_{j=1}^{m} [HA_j - \binom{m}{j-1} FC^{(m-j+1)}] \mathbf{z}_j + H\mathbf{q} - F\mathbf{r}^{(m)} + B_m \mu_m \tag{3.27b}$$

where $B_j = 0$ if $j > k$, but for $1 \le j \le k$, $B_j(t) \in \mathcal{R}^{n_x \times (n_x - n_y)}$ have full rank and satisfy for each t

$$B_j = N_j C^T, \qquad R_j B_j = 0 \tag{3.28}$$

with N_j smooth nonsingular matrices. The additional unknowns $\mu_j(t) \in \mathcal{R}^{n_y}$ are multipliers.

Using a discretization on a mesh, the discretized equations (3.26) or (3.27) are required to hold together with (3.25) at all mesh points.

Clearly, the obtained system (3.27), (3.25) is a DAE of index 2 in Hessenberg form which, together with (2.14), is well-conditioned if the original problem is. Also, the projected invariant approach can be viewed as an ODAE approach, although we feel that it gives more insight. The advantage here compared with the stabilized index reduction (2.18) is that there the stabilizer B is dictated by the problem while here we may choose B_j (i.e. R_j, so long as (3.15) is satisfied). This proves useful in cases where C is a much better behaved function than B, because here we may in fact choose $B_j = C^T$ – see Example 3 in Section 4. □

Summarizing the results of this section, we have seen that the stability of the original problem is preserved by problem reformulations such as stabilized index reduction, introduction of (properly chosen) Baumgarte parameters, and reduction to state-space form. Unstabilized index-reduction leads to a mild instability[2] which becomes progressively worse for higher-index problems. Finally, overdetermined DAEs can be regarded as a special case of one of the above forms, depending on the projection which is used in the numerical solution procedure. In the next section, we will consider the stability of discretization methods applied to these various formulations.

[2]We note that this result appears to be in conflict with statements in [7] which imply that unstabilized index-reduction can lead to an instability which is worse than mild. This is because we are considering the stability of the time-dependent system as determined by the boundedness of the Green's function, whereas in [7], only the local eigenvalues are considered. We also note that in some engineering applications, any significant drift from the original constraint manifold may be considered to be unacceptable; thus even the mild instability may pose a problem.

4 Discretization

4.1 Backward Euler for an index-2 DAE

To better understand the stability behaviour of numerical methods applied to the above formulations, consider the Hessenberg index-2 system

$$\mathbf{x}' = A\mathbf{x} + B\mathbf{y} + \mathbf{q} \tag{4.1a}$$
$$\mathbf{0} = C\mathbf{x} + \mathbf{r} \tag{4.1b}$$

which is a special case of (2.3) for $m = 1$ and may arise from stabilized or unstabilized reduction of a higher-index system to index-2. For simplicity of presentation, we will consider discretizing (4.1) by the backward Euler method, which gives

$$\mathbf{x}_n = \mathbf{x}_{n-1} + hA_n\mathbf{x}_n + hB_n\mathbf{y}_n + h\mathbf{q}_n \tag{4.2a}$$
$$\mathbf{0} = C_n\mathbf{x}_n + \mathbf{r}_n \tag{4.2b}$$

Note that, if we derive first an explicit ODE in $\mathbf{x}$ by differentiating (4.1b) and using this to eliminate $\mathbf{y}$ and then discretize using backward Euler, we get

$$\mathbf{x}_n = (I - hHA + hFC')^{-1}[\mathbf{x}_{n-1} + hH\mathbf{q} - hF\mathbf{r}']$$

(all quantities are sampled at t_n, unless otherwise noted). So the amplification matrix is

$$(I - hHA + hFC')^{-1} \tag{4.3}$$

But for (4.2) we obtain, upon multiplying (4.2a) by C and substituting (4.2b) to eliminate $\mathbf{y}_n$,

$$\mathbf{x}_n = H(\mathbf{x}_{n-1} + hA\mathbf{x}_n + h\mathbf{q}) - F\mathbf{r}$$

Then, using

$$H_n\mathbf{x}_{n-1} = \mathbf{x}_{n-1} - F(C_{n-1} + hC'_{n-1} + O(h^2))\mathbf{x}_{n-1}$$

yields

$$\mathbf{x}_n = (I - hHA)^{-1}[(I - hFC' + O(h^2))\mathbf{x}_{n-1} + hH\mathbf{q} - F(\mathbf{r}_n - \mathbf{r}_{n-1})]$$

so the amplification matrix is approximately

$$(I - hHA)^{-1}(I - hFC'). \tag{4.4}$$

Taking for simplicity $A = 0$ we see that, while in (4.3) we have the backward Euler matrix for the ODE

$$\mathbf{x}' = -FC'\mathbf{x}, \tag{4.5}$$

in (4.4) we have the forward Euler matrix for the same ODE. If (4.5) is stiff then the backward Euler scheme for (4.1) behaves like a nonstiff method!

The same phenomenon can also be seen as follows. Let $\mathbf{u}_n = R_n\mathbf{x}_n$, where $R_n = R(t_n)$, cf. (2.4). Then $\mathbf{x}_n = S_n\mathbf{u}_n - F_n\mathbf{r}_n$. Multiplying (4.2a) by R_n and changing variables to $\mathbf{u}$, we find that

$$\mathbf{u}_n = \mathbf{u}_{n-1} + h((R'S)_{n-1} + O(h))\mathbf{u}_{n-1} + hRAS\mathbf{u}_n - h(R' + O(h))F\mathbf{r} + hR\mathbf{q} \quad (4.6)$$

We note that (4.6) is a consistent discretization of the EUODE, but it is not the same as backward Euler applied directly to the EUODE because in (4.6), the term involving $R'S$ is discretized *explicitly*. Thus for problems where $R'S$ is large but the solution is smooth, we would expect that the stepsize for (4.2) must be restricted to maintain numerical stability. [3] A similar problem of numerical instability arises when the higher-index problem is discretized directly by such a method.

The analysis using $\mathbf{u}$ has the advantage that the amplification matrix has a smaller size (because there are fewer unknowns). But it depends on the choice of R as per (2.17), whereas (4.3)-(4.5) are independent of the choice of R. When considering (4.6) and $R'S$ one must avoid premature conclusions about instability of the backward Euler scheme, although a positive stability conclusion (upon finding that $R'S$ is not large for a given problem) is immediate.

It is natural to ask under what conditions and for which formulations can FC' (or $R'S$ for the best scaling) become large. The question is more immediately answered for $FC' = B(CB)^{-1}C'$. If we assume that the solution $\mathbf{x}$ varies at a rate similar to that of C, so that the step size taken for accuracy reasons satisfies $h||C'|| << 1$, then FC' can be large only if $||B||$ or $||(CB)^{-1}||$ are large (and forming the product $B(CB)^{-1}$ does not cancel this effect). Assume also that $||C||, ||(CC^T)^{-1}|| = O(1)$. In such a case the projected invariant approach (3.27) with $k = m = 1$ and $B_1 = C^T$ is rather useful: the obtained index-2 DAE which is subsequently discretized is

$$\mathbf{x}' = (HA - FC')\mathbf{x} + H\mathbf{q} - F\mathbf{r}' + C^T\mu \quad (4.7a)$$

$$\mathbf{0} = C\mathbf{x} + \mathbf{r} \quad (4.7b)$$

so C^T plays the role which B plays in (4.1) and a BDF discretization is expected to behave like a stiff solver because $C^T(CC^T)^{-1}C'$ is not large in norm. This is demonstrated below in Example 2. The price paid to obtain (4.7) does include an additional differentiation of the constraints. In a forthcoming paper, we will show for

[3] This property of inherently explicit treatment of $R'S$ when the index-2 problem is discretized directly is shared also by higher-order BDF and by most implicit Runge-Kutta schemes. An investigation of a projected midpoint method for which the stability is the same as for the discretization of the EUODE by midpoint will be reported in the near future. We note also that because of the strong relationship between semi-explicit index-two problems and fully-implicit index-one problems [8], this problem of numerical instability can also be expected to occur for certain fully-implicit index-one DAEs.

mechanical systems how these equations can be formulated so that they are generated and solved very efficiently.

We note further that for systems such as (4.7) where $B = C^T$, $R'S$ is of moderate size, *even if* $||(CC^T)^{-1}||$ *is large.* To see this, assume that C is analytic. Then there is an analytic SVD

$$C^T = U \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V^T$$

and we may choose $R = (\,0 \quad I\,)\,U^T$. Then (for each t)

$$(\,F \quad S\,) = \begin{pmatrix} C \\ R \end{pmatrix}^{-1} = U \begin{pmatrix} (V\Sigma)^{-1} & 0 \\ 0 & I \end{pmatrix},$$

so,

$$S = U \begin{pmatrix} 0 \\ I \end{pmatrix}.$$

It follows that

$$R'S = (\,0 \quad I\,)\,U'^T U \begin{pmatrix} 0 \\ I \end{pmatrix}.$$

Hence for most reasonable C, the performance of numerical methods based on discretization of the projected invariants formulations should not degrade due to stability when approaching a singularity in C (i.e. no smaller step sizes are enforced due to stability).

4.2 Discretization and stiffness

A number of numerical methods currently implemented in CAD codes consist of more or less standard stiff or nonstiff discretizations applied to one of the formulations in Section 3. By a "nonstiff discretization" we mean a difference scheme (e.g. explicit Runge-Kutta) which works efficiently for a nonstiff initial value ODE, but becomes inefficient for a stiff ODE because absolute stability restrictions force a step size selection h which is much smaller than what accuracy requirements alone would dictate. A "stiff discretization", e.g. a BDF scheme, does not usually suffer from such absolute stability restrictions and is inherently an implicit difference scheme.

We now consider such methods:

1. Baumgarte stabilization, followed by (i) a nonstiff discretization or by (ii) a stiff discretization.

2. Stable reduction to index 2 (as in Sections 3.5 or 3.2 or in (2.18)), followed by a stiff discretization.

3. Reduction to state space form, followed by (i) a nonstiff discretization or by (ii) a stiff discretization.

Given that we consider essentially the same discretization schemes applied to problem reformulations which we have just proved equivalent under mild conditions, one might expect all of these methods to perform equally well. As it turns out, however, it is surprisingly easy to give examples (as we shall do below) where each of the three methods significantly outperforms the other two. Indeed, it is often not very clear in the literature what is meant by the term "stiffness" when it is applied to a higher index DAE. To understand this, we distinguish among four cases for the EUODE (2.13) (or (2.20)).

1. The EUODE is nonstiff; B and C vary slowly.
2. The EUODE is nonstiff; B or C do not vary slowly.
3. The "frozen coefficients part" of the EUODE, viz. $\hat{\mathbf{u}}^{(m)} = \sum_{j=1}^{m} RA_j S\hat{\mathbf{u}}^{(j-1)}$ (or the homogeneous (2.20) with constant (frozen) coefficients) is stiff; B and C vary slowly. In this case, the stiffness is caused by the 'ODE part' of the system.
4. The "variable coefficients part" of the EUODE, i.e. what remains after subtracting out the frozen coefficient part, is stiff. In this case, the stiffness is caused by time- or solution- dependent coupling of the constraints with the differential equations.

CASE 1

Many mechanical systems yield ODEs which are not stiff. If no part of the mechanical system moves rapidly in time and the system is not heterogeneous, we may expect a nonstiff ODE to result in all problem formulations of Section 3. In this case, a Baumgarte stabilization (3.5) with say $\gamma = 1$ in (3.10), can be efficiently solved using a nonstiff discretization. For such examples, see [3]. Note that γ should not be taken large in this case, because this may introduce artificial stiffness (cf. (3.9b)).

While the Baumgarte technique yields a nonstiff ODE, so for instance an explicit difference scheme may be applied to (3.5), the other two reformulations require satisfaction of constraints and therefore have an implicit part, even if the reduced ODE is discretized using an explicit scheme [14] (similarly [12]). In simple situations (where the m^{th} constraint differentiation is not a bother either), such a Baumgarte technique is therefore more efficient.

CASE 2

Generally, a robust discretization would have to use a step size commensurate with the variation of B and C. With such a step size, a Baumgarte technique or an index-2 reduction method should perform well, as above, except that the additional constraint differentiation or a poorly-scaled choice of the parameter in the Baumgarte technique might increase errors.

With C varying significantly, however (e.g. corresponding to a rapidly rotating shaft), the robustness of a state space form reduction using a constant $\tilde{R}$ may be called into question.

Example 1

For

$$C(t) = (\sin(\nu t), \cos(\nu t)), \qquad 0 \leq t$$

with $\nu \geq 1$ a parameter, an appropriate choice for $\tilde{R}$ satisfying (3.15) at $t_c = 0$ is

$$\tilde{R} = (1, 0), \qquad 0 \leq t \leq 1$$

But then, $\begin{pmatrix} \tilde{R} \\ C \end{pmatrix}$ is singular at $t = (j + 1/2)\pi/\nu$ for all j integer. It is clear that one must restart $\tilde{R}$ (i.e. switch coordinates) at steps $O(1/\nu)$ apart. While the discretization step for any of the other methods must be $O(1/\nu)$ as well, a simple discretization step involves much less effort than a restart.

What is potentially worse, the detection of restart points is not an easy matter in practice. (This is somewhat similar to using a Riccati method for stiff boundary value ODEs, see [5].) To see what happens when a singularity point is missed, we continue the example as follows:

$$\begin{aligned} \mathbf{x}' &= -\mathbf{x} + By + \mathbf{q} \\ 0 &= C\mathbf{x} + r \end{aligned}$$

with $B = C^T$, $x_1(0) = 1$, and $\mathbf{q}$ and r are chosen to be

$$\begin{aligned} \mathbf{q} &= \begin{pmatrix} 2e^t + \frac{\sin(\nu t)e^t}{(2-t)} \\ 2e^t + \frac{\cos(\nu t)e^t}{(2-t)} \end{pmatrix} \\ r &= -(\sin(\nu t) + \cos(\nu t))e^t \end{aligned}$$

such that the solution is $\mathbf{x}^T = e^t(1, 1)$, $y = -e^t/(2 - t)$. With

$$R(t) = (\cos(\nu t), -\sin(\nu t)),$$

we have $S^T = R$, $F^T = C$ and the homogeneous part of the EUODE is

$$u' = -u$$

with $u(0)$ given. This problem is stable, with $K = O(1)$ in (2.15) independently of ν.

The state space form with $\tilde{R} = (1, 0)$ gives, on the other hand, an ODE whose homogeneous part is

$$\tilde{u}' = -(1 + \nu \tan(\nu t))\tilde{u}$$

So, if one ignores, or misses, a singularity point, then one may end up integrating an unstable ODE.

In Table 4.1 we list some results obtained using a backward Euler discretization with step size $h = .01$ for $\nu = 1000$. The problem is solved over the interval $t \in [0, 1]$. We denote by γ the Baumgarte parameter (i.e. we have replaced the constraint

$g(\mathbf{x},t) = 0$ with $g' + \gamma g = 0$ except for the case $\gamma = \infty$ which corresponds to a direct discretization of the given problem). The discretization of (4.7) is referred to as Projected Invariant. The reported errors are the max-norm of errors in $\mathbf{x}$, and the reported drift is the magnitude of the residual of the original constraint, at the endpoint of the time interval.

Method	γ	Error	Drift
Baumgarte	0.	.26e+79	.33e+79
Baumgarte	1.	.10e+79	.13e+79
Baumgarte	10.	.37e+75	.48e+75
Baumgarte	100.	.63e+53	.85e+53
Baumgarte	1000.	.21e+08	.13e+09
Baumgarte	10000.	.23e-3	.23e-4
Baumgarte	∞	.20e-3	.14e-15
Projected Invariant	NA	.20e-3	.25e-16
'State-space form' ($\tilde{R} = (1,0)$)	NA	.42e+1	0

Table 4.1: Behavior of Methods for Example 1

This example shows the index-2 reduction method in a particularly favourable light: since $|g'| >> |g|$, a rather large γ is needed for the Baumgarte technique to work well. Insisting on satisfying (4.1b) or (4.7b) in the context of an index-2 Hessenberg DAE is advantageous.

□

CASE 3

We apply the same BDF scheme to discretize the three formulations. It is well-known that BDF schemes usually perform well for stiff initial value ODEs. It is less well-known that the theory justifying this performance is at present incomplete, and applies mainly to scalar equations. Consider a stiff initial value ODE

$$\mathbf{x}' = A(t)\mathbf{x} \tag{4.8}$$

and its backward Euler discretization

$$h_n^{-1}(\mathbf{x}_n - \mathbf{x}_{n-1}) = A_n\mathbf{x}_n \tag{4.9}$$

where $0 = t_0 < t_1 < \ldots < t_N = 1$, $h_n = t_n - t_{n-1}$, $A_n := A(t_n)$. Given a nonsingular transformation $T(t)$, let

$$\mathbf{w} = T^{-1}\mathbf{x} \tag{4.10}$$

Then $\mathbf{w}$ satisfies the ODE

$$\mathbf{w}' = (T^{-1}AT - T^{-1}T')\mathbf{w} \equiv U\mathbf{w} \tag{4.11}$$

If U is upper triangular with off-diagonal elements which are not too large, or if U is essentially diagonally dominant (see [10], [11], or Ch.10 of [2]), then a backward Euler

scheme applied to (4.11), i.e. the discretization is applied *after* the transformation, performs well as a stiff discretization scheme. (To see this we may consider the diagonal part of U first, obtaining stability results for a scalar equation for each of the equations in (4.11), and follow this by a contraction argument for the full U.) But if we apply the transformation (4.10) after the discretization (4.9), we obtain ($\mathbf{w}_n := T_n^{-1}\mathbf{x}_n$)

$$h_n^{-1}(\mathbf{w}_n - \mathbf{w}_{n-1}) = T_n^{-1}A_nT_n\mathbf{w}_n - (T_{n-1}^{-1}T_{n-1}' + O(h_n))\mathbf{w}_{n-1} \qquad (4.12)$$

so, the variable transformation term $T^{-1}T'\mathbf{w}$ is discretized explicitly at $n-1$ instead of at n. Usually the term $T^{-1}AT$ dominates, accounting for the practical success of the backward Euler and higher order BDF schemes.

Our case 3 corresponds to the domination of $T^{-1}AT$ in (4.11), (4.12): It is easy to see that the frozen coefficient part of the EUODE (2.20) is preserved in various transformations even after discretization (i.e. it is not significant whether the reformulation precedes discretization or vice versa). Therefore, a BDF discretization of any of the three formulations in this case results in a (stiffly) stable method and performs well. The method of reduction to index 2 without the extra differentiation is most straightforward under these circumstances.

CASE 4

In contrast to case 3 above, the variable coefficient part of the various transformations, i.e. those terms involving derivatives of R, S, C etc. in (2.13), (2.20), (3.5) and (3.17), does not generally get reproduced under discretization, as we saw in Section 4.1. The phenomenon is much like in (4.12), but it may be practically worse because unlike in the ODE case R and S (and $\tilde{R}$) do not depend on A_j of (2.3) at all, so it is easy to envision situations where the variable coefficients part of the EUODE dominates. In such circumstances a backward Euler discretization may behave like a nonstiff discretization, causing a possible slowdown in an automatic integrator. Application of a state space form method may be advantageous then, if the restart difficulty is not present.

Example 2

Consider for $0 \le t \le 1$

$$\begin{aligned} x_1' &= (2-t)\nu y + q_1(t) \\ x_2' &= (\nu - 1)y + q_2(t) \\ 0 &= (t+2)x_1 + (t^2-4)x_2 + r(t) \end{aligned}$$

with $x_1(0) = 1$. Here $\nu \ge 1$ is a parameter. The inhomogeneities $\mathbf{q}$ and r are chosen to be

$$\begin{aligned} \mathbf{q} &= \begin{pmatrix} (1+\nu)e^t \\ (1+\frac{\nu-1}{2-t})e^t \end{pmatrix} \\ r &= -(t^2+t-2)e^t \end{aligned}$$

such that the exact solution is $x_1 = x_2 = e^t$, $y = -\frac{e^t}{2-t}$.

This is essentially the same example as Example 1 in [1], but with $A_1 = 0$, so there is no frozen coefficient part in the EUODE. With

$$R(t) = \nu^{-1}(1-\nu, (2-t)\nu)$$

we have $\begin{pmatrix} R \\ C \end{pmatrix}^{-1} = (4-t^2)^{-1} \begin{pmatrix} (4-t^2)\nu & (2-t)\nu \\ (t+2)\nu & \nu - 1 \end{pmatrix}$ so (2.5), (2.6), (2.7) are satisfied with $K_1 = O(1)$, $K_2 = O(\nu)$, $K_3 = O(1)$. The EUODE for the homogeneous problem is

$$u' = R'Su = -\frac{\nu}{2-t}u$$

subject to an initial condition. Hence this is a stable problem with $K_4 = O(1)$, $K = O(\nu)$ in (2.15). Note also that $||F|| = O(\nu)$ and $||C'|| = O(1)$.

We certainly expect any of the numerical methods mentioned in this section to work well when the discretization step size $h = \max_n h_n$ satisfies $h\nu << 1$. It is more interesting to find out what happens, say, when $h\nu = 10$, which for $\nu = 1000$ yields a rather small step relative to the smoothness of the solution.

At first consider a state space reduction using $\tilde{R} = (1, 0)$. Thus $\tilde{S}^T = (1, 1/(2-t))$ and the homogeneous part of the ODE (3.17) is

$$\tilde{u}' = -\tilde{R}FC'\tilde{S}\tilde{u} = -\frac{\nu}{2-t}\tilde{u}$$

If the inhomogeneous version of this ODE is discretized by a BDF scheme, or any other L-stable scheme, then not only is stability maintained for $h\nu$ large but also accuracy *improves* as ν increases with h fixed, because there is only a fast, stable solution mode present and no slow ones. The transformation back from $\tilde{u}$ to $\mathbf{x}$ preserves this accuracy. Thus, the state space reduction performs here superbly.

In contrast, the same BDF discretization applied to the other formulations has a significant nonstiff behaviour. In Table 4.2 we display results using Baumgarte's technique with backward Euler and applying backward Euler directly to the original index-2 DAE. Tests are performed with $\nu = 1000$, $h = .01$.

A comparison between Tables 4.1 and 4.2 confirms that the practical control of the Baumgarte parameter may indeed be a nontrivial affair. Note also the excellent performance of the discretization of (4.7).

□

Example 3

This example is a linear model of a 'mechanical system'

$$\mathbf{p}' = \mathbf{v} \tag{4.13a}$$

$$M(t)\mathbf{v}' = \mathbf{f}(\mathbf{p}, \mathbf{v}, t) - C^T(t)\lambda + \mathbf{q}(t) \tag{4.13b}$$

$$\mathbf{0} = C(t)\mathbf{p} + r(t) \tag{4.13c}$$

Method	γ	Error	Drift
Baumgarte	0.	.19e-2	.85e-2
Baumgarte	1.	.56e-3	.49e-2
Baumgarte	10.	.91e-3	.29e-3
Baumgarte	100.	.27e-4	.93e-8
Baumgarte	1000.	.13e+42	.45e+39
Baumgarte	∞	.92e+74	.45e+58
Projected Invariant	NA	.14e-4	0
'State-space form' ($\tilde{R} = (1,0)$)	NA	.14e-4	0

Table 4.2: Behavior of Methods for Example 2

where $M(t)$ is symmetric positive definite. We choose

$$\mathbf{f} = \begin{pmatrix} 0 & 0 \\ 0 & \frac{2+t}{\nu} \end{pmatrix} \mathbf{v}, \qquad C = (1, t-2), \qquad M(t) = \frac{2+t}{\nu^2} \begin{pmatrix} \frac{\nu^2+(\nu-1)^2}{2-t} & -\nu(2\nu-1) \\ -\nu(2\nu-1) & 2(2-t)\nu^2 \end{pmatrix},$$

resulting in

$$B = M^{-1}C^T = \begin{pmatrix} (4-t^2)\nu \\ (\nu-1)(t+2) \end{pmatrix}$$

The inhomogeneous terms $\mathbf{q}(t)$, $r(t)$ and the initial conditions have been chosen so that the solutions for both components of $\mathbf{p}$ and $\mathbf{v}$ are e^t, and $\lambda = e^t/(2-t)$. This example is closely related to Example 2 and, in particular, has the same term $R'S$ with $R'' = 0$. We will consider its solution in two different formulations. In the first formulation, the twice-differentiated constraint is used to eliminate λ, and then the original constraint is reintroduced via a new Lagrange multiplier μ, to obtain

$$\begin{aligned} \mathbf{p}' &= \mathbf{v} + D\mu && \text{(4.14a)} \\ \mathbf{v}' &= HM^{-1}\mathbf{f} + HM^{-1}\mathbf{q}(t) - Fz - Fr''(t) && \text{(4.14b)} \\ \mathbf{0} &= C\mathbf{p} + r(t) && \text{(4.14c)} \end{aligned}$$

where $z = 2C'\mathbf{v} + C''\mathbf{p}$. This is the projected invariant formulation (3.27) with $m = 2, k = 1$.[4] We will consider various choices for the projection matrix $D(t)$ satisfying that CD is nonsingular for each t. The second formulation is the following stabilized index-2 system,

$$\begin{aligned} \mathbf{p}' &= \mathbf{v} + D\mu && \text{(4.15a)} \\ M\mathbf{v}' &= \mathbf{f} - C^T\lambda + \mathbf{q} && \text{(4.15b)} \\ \mathbf{0} &= C\mathbf{p} + r && \text{(4.15c)} \\ \mathbf{0} &= C\mathbf{v} + C'\mathbf{p} + r' && \text{(4.15d)} \end{aligned}$$

[4]We note that it is possible to construct another projected invariant formulation for mechanical systems, using the technique described in Section 3.5 for $m = 2, k = 2$, which enforces not only the position constraints but also the velocity constraints.

In Table 4.3 we present the results for the projected invariant formulation (4.14) with projections D given by B, C^T, 'unit' $= (0,1)^T$, and for the unstabilized index-two ('U-2') formulation (i.e. where the constraints in (4.14) have been simply differentiated once), for values of $\nu = 1, 100$ and 1000. In Table 4.4 we present the results for the stabilized index-2 formulation (4.15) under the same conditions. All test results are with the backward Euler scheme on the interval $[0,1]$, with a uniform stepsize $h = .01$. The recorded errors are measured at $t = 1$ in the indicated variable. For the first 9 rows of Table 4.3 (and Tables 4.5 and 4.7 as well), 'Drift' indicates the drift in the velocity constraint (derivative of the original constraint) at the endpoint of the interval. For unstabilized index-2 formulation, where the drift in velocity constraint is 0 but the drift in position constraint is not, the latter is indicated. Since the drifts for the stabilized formulation (4.15) are essentially zero (except when everything blows up), they are not recorded in Table 4.4.

Projection	ν	μ	error(p)	error(v)	drift
B	1	.16e-2	.13e-1	.12e-1	.49e-2
B	100	.91e-4	.38e-2	.35e-2	.14e-3
B	1000	.34e+73	.93e+74	.36e-2	.93e+74
C^T	1	.16e-2	.12e-1	.12e-1	.35e-2
C^T	100	.18e-2	.74e-2	.35e-2	.37e-2
C^T	1000	.18e-2	.72e-2	.36e-2	.37e-2
unit	1	.24e-2	.11e-1	.12e-1	.26e-2
unit	100	.27e-2	.65e-2	.35e-2	.29e-2
unit	1000	.27e-2	.64e-2	.36e-2	.29e-2
U-2	1	NA	.13e-1	.12e-1	.17e-1
U-2	100	NA	.37e-2	.17e-1	.73e-2
U-2	1000	NA	.10e+71	.11e+73	.10e+71

Table 4.3: Example 3, Projected Invariant Formulation

Additional experiments were carried out for the problem (4.13) with $\mathbf{f} = \mathbf{0}$ and the same exact solution. The results are summarized in Tables 4.5 and 4.6, which are analogous to Tables 4.3 and 4.4, respectively.

We note that, as predicted, methods using the B-projection or involving B in a Hessenberg index-2 formulation (even when stabilized using other projections) can experience a serious error growth when $h\nu$ is large, due to the large size of the $R'S$-term. Only the projected invariant formulations using the "good" projections C^T and 'unit' yield acceptable results for both choices of $\mathbf{f}$ when $h\nu = 10$. The backward Euler scheme performs like a nonstiff integrator in these circumstances for the other methods. The good behaviour of the projected invariant formulation for the projections C^T and 'unit' follow directly from the discussion earlier in this section.

Let us calculate the EUODE for this example. Writing (4.13) with $\mathbf{f} = E(t)\mathbf{v}$ in the form (2.3), we have $m \leftarrow 2$, $\mathbf{x} \leftarrow \mathbf{p}$, $\mathbf{y} \leftarrow -\lambda$, $A_1 \leftarrow 0$, $A_2 \leftarrow M^{-1}E$. With R and

Projection	ν	μ	error(p)	error(v)
B	1	.86e-4	.13e-1	.12e-1
B	100	.44e-6	.37e-2	.17e-1
B	1000	.28e-3	.35e-2	.78
C^T	1	.13e-3	.12e-1	.12e-1
C^T	100	.18e-7	.37e-2	.17e-1
C^T	1000	.99e+73	.18e+74	.20e+76
unit	1	.26e-3	.11e-1	.12e-1
unit	100	.42e-7	.38e-2	.17e-1
unit	1000	.16e+75	.15e+75	.16e+77

Table 4.4: Example 3, Stabilized Index-2 Formulation

Projection	ν	μ	error(p)	error(v)	drift
B	1	.16e-2	.12e-1	.86e-2	.49e-2
B	100	.91e-4	.38e-2	.35e-2	.14e-3
B	1000	.34e+73	.93e+74	.35e-2	.93e+74
C^T	1	.17e-2	.11e-1	.86e-2	.35e-2
C^T	100	.18e-2	.73e-2	.35e-2	.37e-2
C^T	1000	.81e-2	.72e-2	.35e-2	.37e-2
unit	1	.25e-2	.10e-1	.86e-2	.27e-2
unit	100	.27e-2	.65e-2	.35e-2	.28e-2
unit	1000	.27e-2	.64e-2	.35e-2	.28e-2
U-2	1	NA	.12e-1	.86e-2	.16e-1
U-2	100	NA	.62e-2	.13e-1	.98-2
U-2	1000	NA	.60e-2	.13e-1	.97e-2

Table 4.5: Example 3 with $\mathbf{f} = \mathbf{0}$, Projected Invariant Formulation

S as in Example 2 we obtain

$$(RA_1 + R'')S\mathbf{u} = 0$$

$$(RA_2 + 2R')(S\mathbf{u})' = \nu[RM^{-1}E + 2(0,-1)][Su' + S'u]$$

For E corresponding to Tables 4.3 and 4.4, $RM^{-1}E = (0,1)$, so by (2.13) the homogeneous EUODE is

$$u'' = -\frac{\nu}{2-t}u' - \frac{\nu}{(2-t)^2}u$$

For $E = 0$ corresponding to Tables 4.5 and 4.6, the homogeneous EUODE is

$$u'' = -\frac{2\nu}{2-t}u' - \frac{2\nu}{(2-t)^2}u$$

So the EUODE in both cases is stable for $\nu > 0$ and stiff for $\nu >> 1$.

Projection	ν	μ	error(p)	error(v)
B	1	.74e-4	.12e-1	.86e-2
B	100	.45e-4	.18	.19
B	1000	.11e+150	.32e+150	.28e+153
C^T	1	.11e-3	.11e-1	.86e-2
C^T	100	.34e-4	.53e-2	.12e-1
C^T	1000	.33e-4	.52e-2	.12e-1
unit	1	.22e-3	.99e-2	.86e-2
unit	100	.68e-4	.50e-2	.12e-1
unit	1000	.67e-4	.49e-2	.12e-1

Table 4.6: Example 3 with $\mathbf{f} = \mathbf{0}$, Stabilized Index-2 Formulation

If we now choose

$$E = \begin{pmatrix} 0 & 0 \\ 0 & \frac{2(2+t)}{\nu} \end{pmatrix}$$

then $RA_2 = -2R'$, so the EUODE is nonstiff. In Tables 4.7 and 4.8 we record results analogous to Tables 4.3 and 4.4 for this case where the EUODE is nonstiff.

Projection	ν	μ	error(p)	error(v)	drift
B	1	.52e-2	.11e-1	.18e-1	.10e-1
B	100	.91e-4	.39e-2	.37e-2	.14e-3
B	1000	.34e+73	.93e+74	.36e-2	.93e+74
C^T	1	.52e-2	.11e-1	.18e-1	.10e-1
C^T	100	.17e-2	.73e-2	.37e-2	.35e-2
C^T	1000	.18e-2	.72e-2	.36e-2	.37e-2
unit	1	.13e-1	.90e-2	.18e-1	.13e-1
unit	100	.25e-2	.64e-2	.37e-2	.27e-2
unit	1000	.27e-2	.64e-2	.36e-2	.28e-2
U-2	1	NA	.14e-1	.18e-1	.18e-1
U-2	100	NA	.30e+2	.30e+15	.30e+2
U-2	1000	NA	.17	.54	.17

Table 4.7: Example 3 with a nonstiff EUODE, Projected Invariant Formulation

We note with no surprise that the problem does not get easier when the large terms in the EUODE cancel one another.

□

A number of methods have been proposed in the literature (see [6] and references therein, [1], [9], [12], [13]), where at each step in t, an integration step for the ODE (3.18a),(3.22) or another form of the DAE, is followed by a projection using a weighted least squares norm to satisfy the constraints (3.25) at the end of the step. Thus, using e.g. backward Euler for the unstabilized (4.14a) (i.e. with $D = 0$) and (4.14b) we

Projection	ν	μ	error(p)	error(v)
B	1	.11e-3	.15e-1	.18e-1
B	100	.47e-4	.37	.35
B	1000	.54e-6	.36	.34
C^T	1	.16e-3	.13e-1	.18e-1
C^T	100	.11e+2	.23	.22e+4
C^T	1000	.18e-2	.60e-1	.34
unit	1	.32e-3	.12e-1	.18e-1
unit	100	.13e+4	.14	.13e+6
unit	1000	.28e-2	.26e-1	.27

Table 4.8: Example 3 with a nonstiff EUODE, Stabilized Index-2 Formulation

have at the n^{th} step

$$h^{-1}(\tilde{\mathbf{p}}_n - \mathbf{p}_{n-1}) = \mathbf{v}_n \tag{4.16a}$$

$$h^{-1}(\mathbf{v}_n - \mathbf{v}_{n-1}) = H_n M_n^{-1}(\mathbf{f}_n + \mathbf{q}_n) - F_n \mathbf{z}_n - F_n \mathbf{r}_n'' \tag{4.16b}$$

and then we find $\mathbf{p}_n$ which satisfies

$$C_n \mathbf{p}_n + \mathbf{r}_n = \mathbf{0} \tag{4.17}$$

and minimizes

$$(\mathbf{p}_n - \tilde{\mathbf{p}}_n)^T W_n (\mathbf{p}_n - \tilde{\mathbf{p}}_n) \tag{4.18}$$

where W_n is a symmetric positive definite matrix. This idea can clearly be written down in the generality of Section 3.5 and gives another variant for resolving overdetermination.

The necessary conditions for the constrained minimization (4.17),(4.18) are

$$W_n(\mathbf{p}_n - \tilde{\mathbf{p}}_n) = C_n^T \mu_n \tag{4.19}$$

where μ_n is a Lagrange multiplier. Therefore

$$\mathbf{p}_n = \tilde{\mathbf{p}}_n + W_n^{-1} C_n^T \mu_n = \mathbf{p}_{n-1} + h\mathbf{v}_n + W_n^{-1} C_n^T \mu_n \tag{4.20}$$

However, an unfortunate choice of W_n may again produce a nonstiff behaviour out of a BDF scheme , because (4.20) is in essence a backward Euler discretization of (4.14a) with $D = W^{-1}C^T$. For Example 3, in particular, the choice $W = M$ is not advisable.

For some schemes, though, the choice of W is not sufficiently arbitrary. For instance in [1] the integration step is an implicit Runge-Kutta (or collocation) step applied directly to an index-2 DAE (4.1). This necessitates in the following projection the choice of W so that $W^{-1}C^T = B$. Therefore, that method applied to Example 2 also behaves like a nonstiff integrator. A way to remedy this is to transform (4.1) into (4.7) before applying the projected Runge-Kutta method. For Example 2 this works very well.

5 Conclusions

We have considered various problem formulations and their discretizations for higher-order, higher-index DAEs such as those which arise in the numerical integration of Lagrange's equations of the first kind for multibody dynamics. A linearized form of the equations was considered, allowing a methodical examination of a number of methods which are in use in practice with respect to stability. This yields a number of conclusions, based on the methods considered.

1. All reasonable problem reformulations used in practice are stable under certain mild assumptions. The exception is an unstabilized index reduction which has an algebraic instability of degree $m-1$. Thus, for holonomic constraints, two unstabilized constraint differentiations yield a linear instability. Applying only one unstabilized differentiation is still stable, though not asymptotically stable. (Note however that an asymptotic stability of $\mathbf{v}$ in (3.9b) does not yield a similar statement for $\mathbf{u}$ in (3.9a).)

2. Applying the same discretization to two stable problem formulations does not necessarily yield similar method characteristics.

3. For simple, slowly varying nonstiff problems, a Baumgarte stabilization coupled with an explicit discretization is recommended. (Note though that other good alternatives exist; one such is proposed in [13].)

4. For problems with rapidly varying constraints, especially if the frozen coefficient problem is stiff, a BDF (or other stiff) discretization applied to a stable index-2 reduction is recommended.

5. For heterogeneous problems, where the mass matrix has widely varying eigenvalues (or, when C is much better behaved than B in (2.3)), the projected invariants stabilization (with the projection based on C^T) coupled with a BDF discretization or other suitable stiff method is recommended.

References

[1] U. Ascher and L. Petzold, *Projected implicit Runge-Kutta methods for differential-algebraic equations*, SIAM J. Numer. Anal., to appear, August 1991.

[2] U. Ascher, R. Mattheij and R. Russell, *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*, Prentice-Hall, 1988.

[3] J. Baumgarte, *Stabilization of constraints and integrals of motion in dynamical systems*, Comp. Math. Appl. Mech. Eng. 1 (1976), 1-16.

[4] K. Brenan, S. Campbell and L. Petzold, *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*, North-Holland, 1989.

[5] L. Dieci, M. Osborne and R. Russell, *A Riccati transformation method for solving linear BVPs: II. Computational aspects*, SIAM J. Num. Anal. 25 (1988), 1074-1092.

[6] E. Eich, K. Führer, B. Leimkuhler and S. Reich, *Stabilization and projection methods for multibody dynamics*, Research report, Inst. Math., Helsinki Univ. of Technology, 1990.

[7] K. Führer and B. Leimkuhler, *Formulation and numerical solution of the equations of constrained mechanical motion*, DFVLR-FB 89-08, Munich, 1989.

[8] C.W. Gear, *Differential-algebraic equation index transformations*, SIAM J. Sci. Stat. Comput. 9 (1988), 39-47.

[9] C. W. Gear, *Maintaining solution invariants in the numerical solution of ODEs*, SIAM J. Sci. Stat. Comp., 7 (1986), 734-743.

[10] H. -O. Kreiss, *Difference methods for stiff ordinary differential equations*, SIAM J. Numer. Anal. 15 (1978), 21-58.

[11] H. -O. Kreiss, N.K. Nichols and D.L. Brown, *Numerical methods for stiff two-point boundary value problems*, SIAM J. Numer. Anal. 23 (1986), 325-368.

[12] Ch. Lubich, *h^2-Extrapolation methods for differential-algebraic systems of index 2*, IMPACT 1 (1989), 260-268.

[13] Ch. Lubich, *Extrapolation integrators for constrained multibody systems*, Technical Report, Universität Innsbruck, 1990.

[14] F. A. Potra and W. C. Rheinboldt, *On the numerical solution of the Euler-Lagrange equations*, J. Mechanics of Structures and Machines, to appear.

[15] W. Schiehlen (Editor), *Multibody Systems Handbook*, Springer-Verlag, 1990.

[16] R. A. Wehage and E. J. Haug, *Generalized coordinate partitioning for dimension reduction in analysis of constrained dynamic systems*, J. of Mechanical Design 104 (1982), 247-255.