A perturbation analysis for the dynamical simulation of mechanical multibody systems

M. Arnold*

Abstract

The dynamical behaviour of mechanical multibody systems with holonomic constraints is described by the Euler–Lagrange equations. In the present paper various approaches to the numerical integration of this differential-algebraic system are compared w. r. t. the sensitivity of the numerical solution to small errors (e. g. round-off errors that arise in the implementation on a computer). Position and velocity coordinates are proved to be much more robust against small perturbations than the Lagrangian multipliers. A comparison of systems with and without friction shows that small errors are substantially more amplified during integration if friction occurs.

Mathematics Subject Classification (1991): 65L05
Keywords: differential-algebraic systems, perturbation index, Euler–Lagrange equations

1 Introduction

The motion of constrained mechanical systems is described by the Euler–Lagrange equations

$$\begin{array}{rcl} M(p)\ddot{p} & = & f(p,\dot{p},\lambda) - G^T(p)\lambda \\ 0 & = & g(p) \end{array} \tag{1}$$

with the vector $p \in \mathbb{R}^{n_p}$ of position coordinates and the Lagrangian multipliers $\lambda \in \mathbb{R}^{n_\lambda}$, that couple the holonomic constraints $g(p) = 0$ to the dynamical equations ($G(p) := \frac{\partial}{\partial p} g(p)$, $\operatorname{rank} G = n_\lambda$). $M(p)$ denotes the symmetric positive definite mass matrix and $f(p,\dot{p},\lambda)$ the applied forces (see e. g. [HW91, Section VI.5]). The *descriptor form* (1) of the model equations can be rewritten as differential-algebraic system of index 3 in p, λ and the velocity coordinates $v := \dot{p}$.

It is well known that the numerical solution of initial value problems for (1) is non-trivial: order reduction may occur, the solution of the arising nonlinear systems of equations (in implicit Runge–Kutta methods and BDF) is difficult, stepsize control must be adapted ([HLR89], [BCP89]). Furthermore the numerical solution does not depend continuously on small perturbations. Small errors that arise in the implementation of the discretization method (e. g. errors caused by stopping the iterative solution of systems of nonlinear equations, round-off errors) can drastically be amplified during integration.

Example 1 Models for complex wheel-rail systems are composed of models for simpler subsystems (e. g. wheelsets). The Euler–Lagrange equations for a rigid wheelset are given in [NA93], [Arn94]. They consist of $n_p = 6$ dynamical equations (corresponding to the 6 degrees of freedom) and $n_\lambda = 2$ constraints that force both wheels to remain in contact with the rails.

*University of Rostock, Department of Mathematics, Postfach, D – 18051 Rostock, Germany,
E-mail: arnold@mathematik.uni-rostock.d400.de

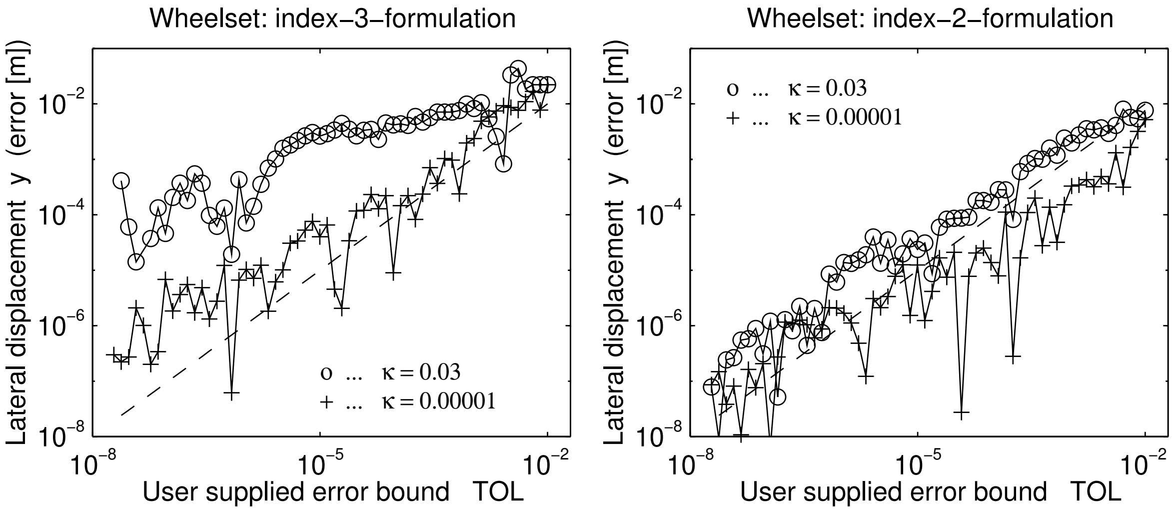


Figure 1: Errors in the dynamical simulation of a wheelset (RADAU5)

In the following numerical test the wheelset was at first shifted $4\,\mathrm{mm}$ in lateral direction and then moved with fixed speed $V = 30\,\mathrm{m/s}$ along a straight track. The solution approaches a limit cycle. The corresponding initial value problem for (1) was integrated with the Runge–Kutta code RADAU5 ([HLR89]) for various error bounds TOL. If the code is applied with standard settings for all parameters then errors of size $10^{-4} \dots 10^{-3}$ in *all* solution components are obtained independently of the tolerance TOL.
This large error in the numerical solution results from the amplification of errors that are caused by stopping Newton's method for the Runge–Kutta equations. In RADAU5 the stopping bound for Newton's method is $\kappa \cdot \mathrm{TOL}$ with $\kappa = 0.03$ as standard ([HW91, pp. 130f]). I. e., in contrast to the Runge–Kutta solution that satisfies $g(p_n) = 0$ exactly the actually computed solution satisfies $\|g(\hat{p}_n)\| \leq \kappa \cdot \mathrm{TOL}$, only. Following the suggestions in [HLR89, Section 9] we repeated the test with a much smaller error bound ($\kappa = 10^{-5}$) and got a numerical solution that has approximately the required accuracy. In the left plot of Figure 1 the error in one of the position coordinates (at $t_e = 5\,\mathrm{sec}$) is plotted for both values of κ.
The large error of RADAU5 with the standard setting $\kappa = 0.03$ is due to the "high" index of (1). This is illustrated applying RADAU5 to the *index-2 formulation* of the model equations: in (1) the constraints $g(p) = 0$ are substituted by their time derivatives $0 = \frac{d}{dt} g(p) = g_p(p)\dot{p} = G(p)v$. The resulting system is analytically equivalent to (1) but has index 2, only ([HW91, Section VI.5]). In the right plot of Figure 1 the errors of RADAU5 are given, again for $\kappa = 0.03$ (standard) and $\kappa = 10^{-5}$. The errors remain approximately within the given bounds TOL. In contrast to the left plot of Figure 1 the errors caused by stopping Newton's method are now already for $\kappa = 0.03$ negligible compared with the discretization error. (This is also the standard situation in ODE theory.)

Example 1 illustrates that a reliable integration of higher index differential-algebraic systems requires not only bounds for the discretization error but also bounds for errors that arise if the discretization method is implemented on a computer. In the present paper we give such error bounds for the model equations (1) of constraint mechanical systems with and without friction. With this perturbation analysis various analytically equivalent formulations of the model equations are compared w. r. t. the sensitivity to small perturbations. This comparison helps to select an appropriate formulation of the model equations for the dynamical simulation in multibody system packages.

The paper is organized as follows: in Sections 2 and 3 we restrict ourselves to systems (1) with $f = f(p, \dot{p})$, i. e. applied forces that are independent of the constraint forces $-G^T(p)\lambda$ (systems without friction). Then (1) may be decoupled into a differential-algebraic system of index 2 in p and $v = \dot{p}$ and an additional system of equations that determines the Lagrangian multipliers λ (Section 2). The perturbation analysis for systems of index 2 presented in [ASW93] can be carried over to these systems. The results are given in Section 2 for the analytical and in Section 3 for the numerical solution with backward Euler method.

The situation becomes more difficult if friction occurs. Then $f = f(p, \dot{p}, \lambda)$ includes the friction forces that depend on the constraint forces and thus on λ. In Section 4 error bounds for the analytical solution are given. The error in position and velocity coordinates is shown to be much larger than before. This is nicely illustrated by simulation results for a point mass moving on a parabola once without friction and once with Coulomb friction (Section 5).

2 Euler–Lagrange equations and differential-algebraic systems of index 2

To analyse the sensitivity of the analytical solution of system (1) the functions $p(t)$, $v(t)$, $\lambda(t)$ are compared with functions $\hat{p}(t)$, $\hat{v}(t) := \dot{\hat{p}}(t)$, $\hat{\lambda}(t)$ that cause (small) residuals $\delta(t)$, $\theta(t)$ when they are inserted into (1):

$$\begin{aligned} M(\hat{p})\ddot{\hat{p}} &= f(\hat{p}, \dot{\hat{p}}, \hat{\lambda}) - G^T(\hat{p})\hat{\lambda} + \delta(t) \\ \theta(t) &= g(\hat{p}) \end{aligned} \tag{2}$$

More precisely: let us consider an initial value problem for (1) on a bounded time interval $[t_0, t_e]$ and suppose that this initial value problem has a solution $p(t)$, $\lambda(t)$ and that the functions f, g and M are sufficiently differentiable in a neighbourhood of this solution. Then we want to give estimates for $\|\hat{p}(t) - p(t)\|$, $\|\hat{v}(t) - v(t)\|$, $\|\hat{\lambda}(t) - \lambda(t)\|$ in terms of the residuals δ and θ and of the perturbations $\|\hat{p}_0 - p_0\|$, $\|\hat{v}_0 - v_0\|$ in the initial values ($p_0 := p(t_0), \ldots$). In contrast to ODE theory these error bounds do *not* depend continuously on $\theta(\tau)$ but include $\|\dot{\theta}(\tau)\|$ and $\|\ddot{\theta}(\tau)\|$.

Such a perturbation analysis is given in great detail for differential-algebraic systems

$$\begin{aligned} \dot{y} &= \varphi(y, z) \\ 0 &= \gamma(y) \end{aligned} \tag{3}$$

of index 2 in [ASW93]. Following the concept of perturbation index ([HLR89, pp. 1ff]) the problem is studied first for the analytical solution. It is proved that the differential components y are in general less affected by small perturbations than the algebraic components z. If appropriate discretization methods are used then the influence of small errors on the numerical solution is bounded in terms that are the discrete analogue of the estimates for the analytical solution. This is discussed in detail for backward Euler method in [ASW93, Sections 3 and 5], similar estimates can be proved for implicit Runge–Kutta methods, that satisfy a contractivity condition ([Arn93]), and for BDF.

Example 2 The index-2 formulation of the model equations in Example 1 is of the form (3) with $y = (p^T, v^T)^T$, $z = \lambda$. Simulations based on this formulation sometimes suffer from the *drift-off effect*: since the original constraints $g(p) = 0$ are not present in the index-2 formulation they may be violated by the numerical solution – especially in long-time simulations. This is avoided if both $g(p) = 0$ and $\frac{d}{dt}g(p) = 0$ are included in the system. Instead of $\dot{p} = v$ function v is defined by $\frac{d}{dt}g(p) = G(p)v = 0$ and $\dot{p} = v - G^T(p)\mu$ with auxiliary variables μ that vanish identically for the analytical solution (*Gear/Gupta/Leimkuhler formulation* [GLG85], [HW91, Section VI.5]). The GGL formulation has form (3) with $y = (\,p^T\,,\,v^T\,)^T\,,\; z = (\,\mu^T\,,\,\lambda^T\,)^T$. Both GGL formulation and index-2 formulation satisfy the *index-2 condition* " $\gamma_y\varphi_z$ non-singular" in (3) that guarantees that system (3) is of (differential and perturbation) index 2.

If there is no friction in the mechanical system and the applied forces $f(p,\dot{p})$ do not depend on the Lagrangian multipliers then also the original Euler–Lagrange equations (1) are closely related to (3): The solution $p(t)$ remains in the manifold $\mathcal{M} := \{\,p\,:\,g(p) = 0\,\}$ that is defined by the holonomic constraints. At any fixed time t^* the velocity vector $v(t^*)$ lies in the tangential space of $\mathcal{M}$ at $p(t^*)$: $[G(p)v](t^*) = 0$. Thus $v(t^*)$ coincides with its projection $[S(p)v](t^*)$ into this tangential space. The projector $S(p)$ is defined by

$$S(p) := I - [M^{-1}G^T(GM^{-1}G^T)^{-1}G](p) \tag{4}$$

(note, that $GM^{-1}G^T$ is non-singular since M is positive definite and G has full row rank). It holds $[SM^{-1}G^T](p) \equiv 0$ such that the time derivative of the projection $S(p)v$ is *independent of the Lagrangian multipliers*: with the notation $\frac{\partial}{\partial p}S(p)(v,w) := \frac{\partial}{\partial p}(S(p)v)\cdot w$ we get

$$\frac{d}{dt}(S(p)v) = S(p)\dot{v} + (\frac{d}{dt}S(p))v = [SM^{-1}f](p,v) + \frac{\partial}{\partial p}S(p)(v,v)\ . \tag{5}$$

Since $v = S(p)v$ equations (5) together with $\dot{p} = v$ and $g(p) = 0$ are analytically equivalent to system (1). These equations form a system (3) of index 2 with $y = (\,p^T\,,\,(S(p)v)^T\,)^T\,,\; z = G(p)v$ and

$$\begin{aligned} \frac{d}{dt}y &= \frac{d}{dt}\begin{pmatrix} p \\ S(p)v \end{pmatrix} = \varphi(y,z) := \begin{pmatrix} S(p)v + [M^{-1}G^T(GM^{-1}G^T)^{-1}](p)\cdot G(p)v \\ [SM^{-1}f](p,v) + \frac{\partial}{\partial p}S(p)(v,v) \end{pmatrix}\,, \\ 0 &= \gamma(y) := g(p)\ . \end{aligned} \tag{6}$$

Because of $\gamma_y\varphi_z = [GM^{-1}G^T(GM^{-1}G^T)^{-1}](p) \equiv I$ the index-2 condition is trivially satisfied. The Lagrangian multipliers λ can be computed from the 2nd time derivative of the constraints

$$\begin{aligned} 0 &= \frac{d^2}{dt^2}g(p) = \frac{d}{dt}(G(p)v) = g_{pp}(p)(v,\dot{p}) + G(p)\dot{v} \\ &= g_{pp}(p)(v,v) + [GM^{-1}f](p,v,\lambda) - [GM^{-1}G^T](p)\cdot\lambda\ , \end{aligned} \tag{7}$$

a system of equations that is linear in λ if no friction occurs. The same transformation can be performed for the perturbed system (2). Instead of $v = S(p)v$ we now have $G(\hat{p})\hat{v} = \frac{d}{dt}g(\hat{p}) = \dot{\theta}(t)$ and

$$\dot{\hat{p}} = \hat{v} = S(\hat{p})\hat{v} + [M^{-1}G^T(GM^{-1}G^T)^{-1}](\hat{p})\cdot G(\hat{p})\hat{v}$$

and instead of $\dfrac{d^2}{dt^2}g(p) = 0$ in (7) the 2nd time derivative of θ appears: $\dfrac{d^2}{dt^2}g(\hat{p}) = \ddot{\theta}(t)$. Since (7) can be solved w. r. t. λ the perturbation analysis for index-2 systems of [ASW93, Section 2] can straightforwardly be extended to (1):

Theorem 1 *Let a system (1) be given with $f = f(p,\dot{p})$ independent of the Lagrangian multipliers λ. There is a constant C_0 so that for all functions $\hat{p}(t)$, $\hat{\lambda}(t)$ with (2), $\delta \in \mathcal{C}[t_0,t_e]$, $\theta \in \mathcal{C}^2[t_0,t_e]$ and $\|\delta(\tau)\| \le C_\delta$, $\tau \in [t_0,t_e]$ the estimates*

$$\begin{aligned} \|\hat{p}(t) - p(t)\| + \|[S(\hat{p})\hat{v}](t) - [S(p)v](t)\| \le C_0\Big(\Delta_0 + \max_{\tau\in[t_0,t]}\|\int_{t_0}^{\tau}[SM^{-1}](p(w))\delta(w)\,dw\| + \\ + \max_{\tau\in[t_0,t]}\|\theta(\tau)\| + \max_{\tau\in[t_0,t]}(\|[SM^{-1}](p(\tau))\delta(\tau)\| + \|\dot{\theta}(\tau)\|)^2\Big) \end{aligned} \tag{8}$$

$$\|\hat{\lambda}(t) - \lambda(t)\| \leq C_0(\Delta_0 + \max_{\tau\in[t_0,t]} \|\delta(\tau)\| + \max_{\tau\in[t_0,t]} \|\theta(\tau)\| + \max_{\tau\in[t_0,t]} \|\dot{\theta}(\tau)\| + \max_{\tau\in[t_0,t]} \|\ddot{\theta}(\tau)\|) \qquad (9)$$

with $\Delta_0 := \|\hat{p}_0 - p_0\| + \|S(\hat{p}_0)\hat{v}_0 - S(p_0)v_0\|$ *are satisfied for all* $t \in [t_0, t_e]$ *provided that the expression at the right hand side of (8) is sufficiently small (the constant* C_0 *may depend on* C_δ*). Furthermore it holds* $G(p)v = 0$ *and* $G(\hat{p})\hat{v} = \dot{\theta}(t)$ *.*

Proof Estimate (8) is an immediate consequence of [ASW93, Theorem 1a] since

$$\dot{\hat{y}} = \varphi(\hat{y}, \hat{z}) + \begin{pmatrix} 0 \\ [SM^{-1}](\hat{p})\delta(t) \end{pmatrix} = \varphi(\hat{y}, \hat{z}) + \begin{pmatrix} 0 \\ [SM^{-1}](p)\delta(t) + \mathcal{O}(1)\|\hat{p} - p\| \end{pmatrix} .$$

Bound (9) follows from (7). ■

Remark 1

a) The expressions at the right hand side of (8) and (9) can be bounded in terms of $\|\delta(\tau)\|$, $\|\theta(\tau)\|$, $\|\dot{\theta}(\tau)\|$ and $\|\ddot{\theta}(\tau)\|$, i. e. system (1) with $f = f(p, \dot{p})$ is of perturbation index 3 ([HLR89, p. 8]). Nevertheless position and velocity coordinates are not affected by $\|\ddot{\theta}(\tau)\|$, they are *index-2 variables*. In general, errors in p and v are much smaller than in the Lagrangian multipliers λ.

b) Example 3 below illustrates that the bounds of Theorem 1 are sharp for all solution components. In general neither position nor velocity coordinates depend continuously on perturbations $\theta(t)$ in the algebraic constraints. As an exception consider linear systems (1) with constant coefficients, i. e. $f = Dv + Kp$, $g = Cp$ with constant matrices C, D, K, M: for these systems $\|\hat{p} - p\|$ and $\|S(\hat{p})\hat{v} - S(p)v\|$ can be bounded in terms of $\max_\tau \|SM^{-1}\delta(\tau)\|$ and $\max_\tau \|\theta(\tau)\|$ ([ASW93, Theorem 1c], however $G(\hat{p})\hat{v} = \dot{\theta}(t)$).

c) Theorem 1b of [ASW93] gives the corresponding error bounds for index-2 and GGL formulation (note, that these systems are *linear* w. r. t. the algebraic components z if $f = f(p, \dot{p})$). In the present paper the results are given in detail for the numerical solution, only (Remark 4b below).

3 Perturbations in the numerical solution of the Euler–Lagrange equations

The efficient numerical solution of the model equations for multibody systems is very important in simulation packages. The convergence of discretization methods for the *original* Euler–Lagrange equations (1) (i. e. for the system of index 3) is studied e. g. in [HLR89] (implicit Runge–Kutta methods) and [LP86] (BDF for systems with $f = f(p, \dot{p})$). In this section we analyse for systems without friction the sensitivity of the numerical solution to small perturbations. Similar to systems of index 2 ([ASW93]) it can be expected that the influence of perturbations on the numerical solution can be bounded in terms that look like a discretization of (8) and (9): substitute $\|\delta(\tau)\| \to \|\delta_n\|$, $\|\theta(\tau)\| \to \|\theta_n\|$, $\|\dot{\theta}(\tau)\| \to \frac{1}{h}\|\theta_n\|$, $\|\ddot{\theta}(\tau)\| \to \frac{1}{h^2}\|\theta_n\|$. A complete analysis of the error propagation is, however, technically complicated. So we restrict ourselves to backward Euler method

$$\begin{aligned} \frac{p_{n+1} - p_n}{h} &= v_{n+1} \\ M(p_{n+1})\frac{v_{n+1} - v_n}{h} &= f(p_{n+1}, v_{n+1}) - G^T(p_{n+1})\lambda_{n+1} \\ 0 &= g(p_{n+1}) \end{aligned} \qquad (10)$$

While backward Euler method is not well-defined for general systems of index 3 in Hessenberg form (the system of nonlinear equations that has to be solved in each step of integration may have no solution, [HLR89, p. 75]) it can be applied to systems of the special form (1) with $f = f(p, v)$. For these systems backward Euler method is convergent with order 1 (see e. g. Theorem 3 below with $\hat{p}_n = p(t_n)$, $\hat{v}_n = v(t_n)$, $\hat{\lambda}_n = \lambda(t_n)$).

As for the analytical solution the perturbation analysis for (10) is based on the equivalence of (1) and (6). The discretization (10) is obviously not equivalent to backward Euler method for (6) but with $S(p_{n+1})v_n = S(p_n + hv_{n+1})v_n$ it can be transformed to

$$\begin{aligned} \frac{p_{n+1} - p_n}{h} &= v_{n+1} \\ \frac{S(p_{n+1})v_{n+1} - S(p_n)v_n}{h} &= [SM^{-1}f](p_{n+1}, v_{n+1}) + \frac{S(p_n + hv_{n+1}) - S(p_n)}{h} v_n \\ 0 &= g(p_{n+1}) \end{aligned} \tag{11}$$

$$\lambda_{n+1} = [(GM^{-1}G^T)^{-1}](p_{n+1})([GM^{-1}f](p_{n+1}, v_{n+1}) - G(p_{n+1})\frac{v_{n+1} - v_n}{h}) \tag{12}$$

Backward Euler method is well-defined for (3) ([HLR89, Theorem 4.1], [ASW93, Theorem 2]). Using the transformation (11) this can be generalized to

Theorem 2 *The (nonlinear) system of equations (10) is for stepsizes $h \leq h_0$ (locally) uniquely solvable if (p_n, v_n) satisfies*

$$\|p_n - p(t_n)\| + \|v_n - v(t_n)\| \leq \gamma_0 \;\; , \;\; \Delta := \|g(p_n)\| + h\|G(p_n)v_n\| \leq \Delta_0 h$$

with sufficiently small positive constants h_0, γ_0 and Δ_0. It holds

$$\|p_{n+1} - p_n\| + \|S(p_{n+1})v_{n+1} - S(p_n)v_n\| \leq C_{h\Delta}(h + \Delta) \;\; , \;\; \|G(p_{n+1})v_{n+1}\| \leq C_{h\Delta}(h + \frac{\Delta}{h}) \tag{13}$$

with a constant $C_{h\Delta}$ that is independent of h, Δ and (p_n, v_n).

Proof Substituting (11) system (10) is seen to be equivalent to

$$\begin{aligned} \frac{p - p_n}{h} &= v \\ \frac{q - S(p_n)v_n}{h} &= [SM^{-1}f](p, v) + \int_0^1 \frac{\partial S}{\partial p}(p_n + \vartheta(p - p_n))(v_n, v)\, d\vartheta \\ 0 &= g(p) \end{aligned} \tag{14}$$

with $v = q + [M^{-1}G^T(GM^{-1}G^T)^{-1}](p) \cdot z$ and the solution $p = p_{n+1}$, $q = S(p_{n+1})v_{n+1}$, $z = G(p_{n+1})v_{n+1}$, $v = v_{n+1}$ (apply $\Phi(1) - \Phi(0) = \int_0^1 \Phi'(\vartheta)\, d\vartheta$ to $\Phi(\vartheta) := S(p_n + \vartheta hv)v_n$). The only difference to backward Euler method for system (6) is that instead of $\frac{\partial S}{\partial p}(p_{n+1})(v_{n+1}, v_{n+1})$ a similar term containing p_{n+1}, v_{n+1} *and* p_n, v_n appears in the discretization of $\varphi(y, z)$.[1] It is verified straightforwardly that this does not affect the solvability of the system of nonlinear equations since $[\gamma_y \varphi](y, z) = (\, G(p) \,,\, 0 \,) \begin{pmatrix} v \\ \ldots \end{pmatrix} = G(p)v$ is discretized in standard way. Thus the statement of the theorem follows from the corresponding result for systems of index 2 ([ASW93, Theorem 2]). Equations (12) define λ_{n+1}. ■

[1] In [AP93] an analogous phenomenon is studied for the application of backward Euler to linear differential-algebraic systems of index 2 with time-varying coefficients. Backward Euler is said to discretize these systems *in part explicitly*.

Remark 2
The solution of (10) can be computed applying simplified Newton method with matrix

$$J = \begin{pmatrix} \frac{1}{h}I & 0 & -[M^{-1}G^T(GM^{-1}G^T)^{-1}](p_n) \\ 0 & \frac{1}{h}I & 0 \\ G(p_n) & 0 & 0 \end{pmatrix}$$

to (11) (initial guess: $p = p_n$, $q = S(p_n)v_n$, $z = G(p_n)v_n$). Using the ideas of [ASW93, Theorem 2] it is shown that in a neighbourhood of the exact solution this method induces a contractive mapping w. r. t. the scaled norm $\|p\| + \|q\| + \sigma\|z\|$ with $\sigma = 1/(L + \frac{\Delta}{h^2})$ and a sufficiently large constant $L > 0$.

While the sequence (p_n) defined by backward Euler method remains in $\mathcal{M} := \{\, p : g(p) = 0 \,\}$ the really computed sequences $(\hat{p}_n)$, $(\hat{v}_n)$, $(\hat{\lambda}_n)$ do not satisfy equations (10) exactly (cf. Example 1) but

$$\begin{aligned} \frac{\hat{p}_{n+1} - \hat{p}_n}{h} &= \hat{v}_{n+1} \\ M(\hat{p}_{n+1})\frac{\hat{v}_{n+1} - \hat{v}_n}{h} &= f(\hat{p}_{n+1}, \hat{v}_{n+1}) - G^T(\hat{p}_{n+1})\hat{\lambda}_{n+1} + \delta_{n+1} \\ \theta_{n+1} &= g(\hat{p}_{n+1}) \end{aligned} \tag{15}$$

with (small) residuals δ_{n+1}, θ_{n+1}. (The kinematic equations $\dot{p} = v$ are linear, such that errors σ_{n+1} in these equations are small. We neglect them for simplicity, otherwise $\|\theta_{n+1}\|$ should have been substituted by $\|\theta_{n+1}\| + h\|\sigma_{n+1}\|$ in the following error bounds.)

We aim at bounds for $\|\hat{p}_n - p_n\|$, $\|\hat{v}_n - v_n\|$, $\|\hat{\lambda}_n - \lambda_n\|$. Obviously it must be guaranteed that the sequences (p_n, v_n, λ_n) and $(\hat{p}_n, \hat{v}_n, \hat{\lambda}_n)$ remain in the neighbourhood of one and the same analytical solution of (1). So the perturbation analysis is restricted to sequences with

$$\begin{aligned} \|\hat{p}_{n+1} - \hat{p}_n\| + \|S(\hat{p}_{n+1})\hat{v}_{n+1} - S(\hat{p}_n)\hat{v}_n\| &\le C^*_{h\Delta}(h + \|g(\hat{p}_n)\| + h\|G(\hat{p}_n)\hat{v}_n\| + h\delta + \theta) \\ \|G(\hat{p}_{n+1})\hat{v}_{n+1}\| &\le C^*_{h\Delta}(h + \frac{1}{h}\|g(\hat{p}_n)\| + \|G(\hat{p}_n)\hat{v}_n\| + \delta + \frac{\theta}{h}) \end{aligned} \tag{16}$$

$$\delta := \max_m \|\delta_m\| \quad , \quad \theta := \max_m \|\theta_m\| + \|g(\hat{p}_0)\| + h\|G(\hat{p}_0)\hat{v}_0\|$$

and a constant $C^*_{h\Delta}$ that is independent of h, δ and θ. Similar to Theorem 2 it can be proved that there is always at least one sequence that satisfies both (15) and (16) with given (δ_n, θ_n) if δ and θ are sufficiently small ($\delta + \frac{1}{h}\theta \le \Delta_0$).

The estimates of the following theorem are the key result of the perturbation analysis: the global error at a (fixed) time t_n is bounded in terms of the errors δ_m, θ_m in one single step of integration.

Theorem 3
There is a constant C_0 so that for all sequences $(\hat{p}_n, \hat{v}_n, \hat{\lambda}_n)$ with (15) and (16) the estimates

$$\begin{aligned} \|\hat{p}_n - p_n\| + \|S(\hat{p}_n)\hat{v}_n - S(p_n)v_n\| &\le C_0(\Delta_0 + \int_h SM^{-1}\delta + h\delta + \theta + (\delta + \frac{1}{h}\theta)^2) \\ \|G(\hat{p}_n)\hat{v}_n - G(p_n)v_n\| &\le C_0(\Delta_0 + \delta + \frac{1}{h}\theta) \\ \|\hat{\lambda}_n - \lambda_n\| &\le C_0(\Delta_0 + \delta + \frac{1}{h^2}\theta) \end{aligned} \tag{17}$$

with $\Delta_0 := \|\hat{p}_0 - p_0\| + \|S(\hat{p}_0)\hat{v}_0 - S(p_0)v_0\|$ *and*

$$\int_h SM^{-1}\delta := \sum_{m=0}^{n-1} h\|[SM^{-1}](p(t_m))\delta_{m+1}\| \tag{18}$$

are satisfied for all $n > 0$, $t_0 + nh \le t_e$ *provided that the stepsize* h *and the expression at the right hand side of (17) are sufficiently small.*

Proof Most parts of the proof follow the lines of the proof of Theorem 3a in [ASW93], so we will skip minor interesting details. E. g. the estimates are proved under the additional assumption

$$\|p_m - p(t_m)\| + \|v_m - v(t_m)\| \le \gamma_0 \ , \quad \|\hat{p}_m - p(t_m)\| + \|\hat{v}_m - v(t_m)\| \le \gamma_0 \ , \quad m = 0, 1, \ldots, n-1$$

with a sufficiently small constant $\gamma_0 > 0$, that follows from the assumptions of the theorem (this can be verified by standard techniques [ASW93]).
1. As for the analytical solution the components $G(p)v$ and λ of the numerical solution are completely defined by p and $S(p)v$: Apply $\Phi(1) - \Phi(0) = \int_0^1 \Phi'(\vartheta)\,d\vartheta$ to the functions $\Phi(\vartheta_1) := g(p_n + \vartheta_1(p_{n-1} - p_n))$ and $\Phi(\vartheta_2) := G(p_n - h\vartheta_2\vartheta_1 v_n)v_n$ to get

$$\begin{aligned} G(p_n)v_n &= \int_0^1 (G(p_n) - G(p_n - h\vartheta_1 v_n))v_n\,d\vartheta_1 + \int_0^1 G(p_n + \vartheta_1(p_{n-1} - p_n))\frac{p_n - p_{n-1}}{h}\,d\vartheta_1 \\ &= -h\int_0^1\int_0^1 g_{pp}(p_n - h\vartheta_2\vartheta_1 v_n)(v_n, \vartheta_1 v_n)\,d\vartheta_2 d\vartheta_1 + \frac{g(p_n) - g(p_{n-1})}{h} \end{aligned}$$

and the corresponding equation for $G(\hat{p}_n)\hat{v}_n$. Thus

$$G(\hat{p}_n)\hat{v}_n - G(p_n)v_n = \frac{\theta_n - \theta_{n-1}}{h} + \mathcal{O}(h)(\|\hat{p}_n - p_n\| + \|\hat{v}_n - v_n\|)$$

and

$$\|G(\hat{p}_n)\hat{v}_n - G(p_n)v_n\| \le 4(\|\hat{p}_n - p_n\| + \|S(\hat{p}_n)\hat{v}_n - S(p_n)v_n\| + \frac{1}{h}\theta)$$

is obtained if h is sufficiently small. In the same way

$$\begin{aligned} \hat{\lambda}_n - \lambda_n = -[(GM^{-1}G^T)^{-1}](\hat{p}_n)\frac{\theta_n - 2\theta_{n-1} + \theta_{n-2}}{h^2} &+ \\ + \mathcal{O}(1)(\|\hat{p}_n - p_n\| + \|\hat{p}_{n-1} - p_{n-1}\| + \|\hat{v}_n - v_n\| + \|\hat{v}_{n-1} - v_{n-1}\|)& \end{aligned} \tag{19}$$

is derived from (12).
2. The error bounds for components p, $S(p)v$ are substantially smaller than that for $G(p)v$ (e. g. $(\frac{1}{h}\theta)^2$ instead of $\frac{1}{h}\theta$). As in the proof of convergence of implicit Runge–Kutta methods ([HLR89]) the error propagation can not be studied in p and v directly to prove these estimates. Therefore the sequence $u_n := (\,(\Delta_n^p)^T\,,\,(\Delta_n^v)^T\,)^T$ with

$$\begin{aligned} \Delta_n^p &:= S(p_n)(\hat{p}_n - p_n) \\ \Delta_n^v &:= S(\hat{p}_n)\hat{v}_n - S(p_n)v_n - [SM^{-1}f_v](p(t_n), v(t_n))(\hat{p}_n - p_n) - \\ &\qquad - \frac{\partial S}{\partial p}(p(t_n))(v(t_n), \hat{p}_n - p_n) - \frac{\partial S}{\partial p}(p(t_n))(\hat{p}_{n-1} - p_{n-1}, v(t_n)) \end{aligned}$$

is defined (set $p_{-1} := p_0 - hv_0$, $\hat{p}_{-1} := \hat{p}_0 - h\hat{v}_0$). As in part 1 of the proof we get

$$\begin{aligned} \|\hat{p}_n - p_n\| + \|S(\hat{p}_n)\hat{v}_n - S(p_n)v_n\| &\le \mathcal{O}(1)(\|u_n\| + \|u_{n-1}\| + \theta) \\ \|G(\hat{p}_n)\hat{v}_n - G(p_n)v_n\| &\le \mathcal{O}(1)(\|u_n\| + \|u_{n-1}\| + \frac{1}{h}\theta) \end{aligned} \tag{20}$$

with $\|u_n\| := \|\Delta_n^p\| + \|\Delta_n^v\|$. Using (14) and the corresponding equations for $\hat{p}_n$, $\hat{v}_n$ the estimates

$$\begin{aligned} \|u_0\| + \|u_1\| &\leq \mathcal{O}(1)(\Delta_0 + h\delta + \theta) \\ \|u_{n+1}\| &\leq (1+\mathcal{O}(h))\|u_n\| + \mathcal{O}(h)\|u_{n-1}\| + h\|[SM^{-1}](p(t_n))\delta_{n+1}\| + \\ &\quad + \mathcal{O}(h)\|p_n - p(t_n)\| \cdot \delta + \mathcal{O}(h)\cdot(h\delta+\theta) + \mathcal{O}(h)\cdot d_n \cdot \frac{1}{h}\theta \end{aligned} \tag{21}$$

with

$$\begin{aligned} d_n := \|p_{n+1} - p(t_n)\| + \|\hat{p}_{n+1} - p_{n+1}\| + \|p_n - p(t_n)\| + \|\hat{p}_n - p_n\| + \\ + \|v_{n+1} - v(t_n)\| + \|\hat{v}_{n+1} - v_{n+1}\| + \|v_n - v(t_n)\| + \|\hat{v}_n - v_n\| \end{aligned} \tag{22}$$

are obtained. Recursive application of (21) results in

$$\|u_n\| \leq C_u(\Delta_0 + \int_h SM^{-1}\delta + \max_{0\leq m<n} \|p_m - p(t_m)\|\cdot\delta + h\delta + \theta + d\cdot\frac{1}{h}\theta) \tag{23}$$

with $d := \max_m d_m$.

3. Estimates (19), (20) and (23) prove convergence with order 1 in all solution components if the special sequence $\hat{p}_n = p(t_n)$, $\hat{v}_n = v(t_n)$, $\hat{\lambda}_n = \lambda(t_n)$ is considered, i. e. if the analytical solution is inserted into (10). (Then δ and θ represent the local discretization error, $\delta = \mathcal{O}(h)$, $\theta = 0$.) Therefore (22) reduces to

$$d = \mathcal{O}(1)\cdot\max_m \|u_m\| + \mathcal{O}(h) + \mathcal{O}(\frac{1}{h}\theta)$$

and the theorem can be proved combining (19), (20) and (23). (Note, that $\frac{1}{h}\theta$ as part of the right hand side in error bound (17) is supposed to be small.) ∎

Remark 3

a) The main difference of the proofs of Theorem 3 and [ASW93, Theorem 3a] (that gives the corresponding result for backward Euler method applied to systems of index 2) is in the term $\frac{\partial S}{\partial p}(p)(v,v)$ of (6): If backward Euler is applied to (1) directly then the discretization of this term contains not only p_{n+1}, v_{n+1} but also p_n, v_n.

b) Theorems 1 and 3 illustrate nicely the motivation of the perturbation analysis for the analytical solution: it is technically simpler than that for the numerical solution but gives already much insight into the sensitivity of the numerical solution to small errors: the bounds of Theorem 3 are obtained if in Theorem 1 the term $\max_\tau \|\theta(\tau)\|$ is substituted by $\max_m \|\theta_m\|$ and $\int SM^{-1}\delta := \max_\tau \|\int_{t_0}^\tau [SM^{-1}](p(w))\delta(w)\,dw\|$ by $\int_h SM^{-1}\delta + h\delta$. The perturbation analysis for (10) is simpler than for more complex discretization methods (implicit Runge–Kutta methods, BDF) since backward Euler method is a one-step one-stage method and the numerical solution satisfies the holonomic constraints $g(p) = 0$.

c) For simplicity we restricted ourselves to fixed stepsizes h but the results can be carried over to the variable stepsize case by standard techniques ([HNW93, Section III.5]).

Remark 4

a) The global error of an approximation $(\hat{p}_n, \hat{v}_n, \hat{\lambda}_n)$ to the solution of (1) consists of the *global discretization error* of the numerical method (that is $\mathcal{O}(h)$ for backward Euler method) and an *additional error term* that is caused by the amplification of small errors that arise in the implementation of the method. Theorem 3 shows that this additional error term is large in the Lagrangian multipliers ($\mathcal{O}(\frac{1}{h^2}\theta)$) and smaller in the components $G(p)v$ that vanish identically for the analytical solution ($\mathcal{O}(\frac{1}{h}\theta)$). Position coordinates p and the projection $S(p)v$ of the velocity vector are less affected by perturbations (note, that in Theorem 3 the right hand side of (17) and thus also $\frac{1}{h}\theta$ are supposed to be small, i. e. $\frac{1}{h^2}\theta \ll \frac{1}{h}\theta \ll \frac{1}{h^2}\theta^2$). The additional error term is negligible if the errors θ in the algebraic constraints are very small (i. e. if the systems of nonlinear equations are solved with very high accuracy in each step of integration) and the stepsize is not too small. The amplification of errors θ causes problems if the stepsize h is small, e. g. at the beginning of the integration and in time intervalls where the solution changes rapidly.

b) To avoid difficulties in the numerical solution of (1) often the time derivative $G(p)v = 0$ of the holonomic constraints is added to the system (Example 2). A drawback of these additional constraints might be the increasing dimension of the system (GGL formulation) or a possible drift off the manifold $\{\, p : g(p) = 0 \,\}$ (index-2 formulation). But compared with the original system of index 3 the numerical solution of these systems of index 2 is much more robust against small perturbations. In Table 1 the essential parts of the error

<table>
<tr><th></th><th>index-3 formulation</th><th>GGL formulation</th><th>index-2 formulation</th></tr>
<tr><td>$\|\hat{p}_n - p_n\|$</td><td rowspan="2">$\left.\right\} \int_h SM^{-1}\delta + h\delta + \theta + (\delta + \frac{1}{h}\theta)^2$</td><td rowspan="3">$\left.\right\} \int_h SM^{-1}\delta + h\delta + \theta + \theta(\delta + \frac{1}{h}\theta)$</td><td rowspan="3">$\left.\right\} \int_h SM^{-1}\delta + h\delta + \theta$</td></tr>
<tr><td>$\|S(\hat{p}_n)\hat{v}_n - S(p_n)v_n\|$</td></tr>
<tr><td>$\|G(\hat{p}_n)\hat{v}_n - G(p_n)v_n\|$</td><td>$\delta + \frac{1}{h}\theta$</td></tr>
<tr><td>$\|\hat{\lambda}_n - \lambda_n\|$</td><td>$\delta + \frac{1}{h^2}\theta$</td><td>$\delta + \frac{1}{h}\theta$</td><td>$\delta + \frac{1}{h}\theta$</td></tr>
</table>

Table 1: Error bounds for the numerical solution (see (18))

bounds for backward Euler method are summarized (the perturbations in the initial values have been omitted for simplicity). These are the results of Theorem 3 (index-3 formulation) and [ASW93, Theorem 3b] (GGL formulation: note, that the system is linear in the algebraic variables λ and μ). Position and velocity variables in the index-2 formulation depend *continuously* on the perturbations. (This can be proved extending the proof of [ASW93, Theorem 3c] to $M(p)\ddot{p} = f(p,\dot{p}) - G^T(p)\lambda\,,\;\; 0 = G(p)v$.) In Table 1 the variable θ denotes the errors in $g(p) = 0$ (index-3), in $g(p) = 0$ and $G(p)v = 0$ (GGL) and in $G(p)v = 0$ (index-2), respectively. Compared with the index-3 formulation the error bounds for GGL formulation and index-2 formulation are smaller by factor h in all solution components. If the perturbations δ and θ are small then the additional error terms in position and velocity coordinates of GGL and index-2 formulation are negligible compared with the discretization error (if e. g. $\theta = \mathcal{O}(h)$ then $\frac{1}{h}\theta^2 = \mathcal{O}(\theta)$ in the error bound for GGL formulation). There are various codes that integrate (1) with $f = f(p,\dot{p})$ in GGL formulation or in index-2 formulation with a stepsize control that is based on position and velocity coordinates. They are robust against small errors and allow an efficient dynamical simulation of multibody systems.

c) Extending the ideas of [GLG85] the *stabilized index-1 formulation* ([FL91]) couples not only the first but in addition the second time derivative of the holonomic constraints $g(p) = 0$ to system (1). The resulting system is of index 2. In the notation of Table 1 the error in p, v and λ can be bounded by $\int_h SM^{-1}\delta + h\delta + \theta + \theta(\delta + \frac{1}{h}\theta)$, i. e. by the error bound for p, v in the GGL formulation. Compared with the GGL formulation the stabilized index-1 formulation increases the dimension of the system by n_λ equations and decreases the error in λ. Therefore it should be prefered only if $n_\lambda \ll n_p$ and if the Lagrangian multipliers and/or the constraint forces have to be computed with very high accuracy.

4 The dynamical simulation of multibody systems with friction

The very special structure of the Euler–Lagrange equations (1) for systems without friction ($f = f(p, \dot{p})$) simplifies the dynamical simulation of the corresponding multibody systems. Both in the original system (1) and in the index reduced systems of Example 2 only the Lagrangian multipliers λ show the typical behaviour of general differential-algebraic systems of index 3 or 2 in Hessenberg form. The error bounds for position and velocity coordinates are substantially smaller than the usual ones.

In this section we illustrate that this does not hold if the applied forces depend on the Lagrangian multipliers, e. g. if friction occurs. The perturbation analysis is extended to the analytical solution of these systems. Again a solution (p, v, λ) of an initial value problem for (1) and functions $(\hat{p}, \hat{v}, \hat{\lambda})$ with (2) are considered. To guarantee that (1) is of index 3 we suppose that $[GM^{-1}\Gamma](p, v, \lambda)$ with $\Gamma(p, v, \lambda) := f_\lambda(p, v, \lambda) - G^T(p)$ is non-singular in a neighbourhood of the analytical solution. This condition is satisfied in most practical applications. (If $f = f(p, \dot{p})$ then $[GM^{-1}\Gamma]$ is non-singular since $f_\lambda \equiv 0$ and $\operatorname{rank} G = n_\lambda$.) The definition (4) of the projector on the tangential space of $\{\, p \,:\, g(p) = 0 \,\}$ is generalized to

$$S(p, v, \lambda) := I - [M^{-1}\Gamma(GM^{-1}\Gamma)^{-1}G](p, v, \lambda) \ .$$

The error bounds of Theorem 4 depend on the structure of Γ. It is straightforwardly verified that the analysis of Sections 2 and 3 can be extended to systems (1) with $f = f_0(p, v) + f_\lambda(p)\lambda$, i. e. $\Gamma = \Gamma(p)$. From a practical point of view this is not very interesting since models for friction usually include constraint forces *and* the velocity vector. In contrast to Theorem 1 the error in p and v can in general not be bounded by $\|\delta\|$, $\|\theta\|$ and $\|\dot{\theta}\|$ even if f depends *linearly* on λ: $f = f_0(p, v) + f_\lambda(p, v)\lambda$. The largest errors have to be expected if (1) is non-linear w. r. t. λ.

Theorem 4

a) There is a constant C_0 so that for all functions $\hat{p}(t)$, $\hat{\lambda}(t)$ with (2), $\delta \in \mathcal{C}[t_0, t_e]$, $\theta \in \mathcal{C}^2[t_0, t_e]$ and $\|\delta(\tau)\| \leq C_\delta$, $\tau \in [t_0, t_e]$ the estimates

$$\begin{aligned} \|\hat{p}(t) - p(t)\| + \|\hat{v}(t) - v(t)\| \leq C_0 \Big(\Delta_0 + \max_{\tau \in [t_0, t]} \| \int_{t_0}^{\tau} [SM^{-1}](w)\delta(w)\, dw \| + \\ + \max_{\tau \in [t_0, t]} (\|\theta(\tau)\| + \|\dot{\theta}(\tau)\|) + \max_{\tau \in [t_0, t]} D(\tau) \Big) \end{aligned} \tag{24}$$

$$\|\hat{\lambda}(t) - \lambda(t)\| \leq C_0(\Delta_0 + \max_{\tau \in [t_0, t]} \|\delta(\tau)\| + \max_{\tau \in [t_0, t]} \|\theta(\tau)\| + \max_{\tau \in [t_0, t]} \|\dot{\theta}(\tau)\| + \max_{\tau \in [t_0, t]} \|\ddot{\theta}(\tau)\|)$$

are satisfied for all $t \in [t_0, t_e]$ with $[SM^{-1}](w) := [SM^{-1}](p(w), v(w), \lambda(w))$, $\Delta_0 := \|\hat{p}_0 - p_0\| + \|\hat{v}_0 - v_0\|$ and $D(\tau) := (\|\delta(\tau)\| + \|\ddot{\theta}(\tau)\|)^2$ provided that the expression at the right hand side of (24) is sufficiently small (the constant C_0 may depend on C_δ).

b) In the special case $f = f_0(p, v) + f_\lambda(p, v)\lambda$ (i. e. the system is linear in the Lagrangian multipliers λ) estimate (24) holds with $D(\tau) := \|\dot{\theta}(\tau)\|(\|\delta(\tau)\| + \|\ddot{\theta}(\tau)\|)$.

Proof As for systems without friction errors are propagated in p and v, only. The Lagrangian multipliers λ are completely defined by (7). The corresponding equation for $\hat{\lambda}$ includes $\dfrac{d^2}{dt^2} g(p(t)) = \ddot{\theta}(t)$ such that

$$\|\hat{\lambda}(t) - \lambda(t)\| \leq \mathcal{O}(1)(\|\hat{p}(t) - p(t)\| + \|\hat{v}(t) - v(t)\| + \|\delta(t)\| + \|\ddot{\theta}(t)\|) \ . \tag{25}$$

a) Differentiating (7) the time derivative $\dot{\lambda}$ may be expressed in terms of p, v, $\dot{p}$ and $\dot{v}$, i. e. for sufficiently differentiable functions p also the projector $S(t) := S(p(t), v(t), \lambda(t))$ has a bounded time derivative $\frac{d}{dt}S(t)$ (chain rule). Thus we have

$$\begin{aligned}\frac{d}{dt}(S(t)(\hat{v}(t)-v(t))) &= S(t)(\dot{\hat{v}}(t)-\dot{v}(t))+\mathcal{O}(\|\hat{v}(t)-v(t)\|)\\ &= S(t)(([M^{-1}f](\hat{p},\hat{v},\hat{\lambda})-[M^{-1}G^T](\hat{p})\cdot\hat{\lambda})-([M^{-1}f](p,v,\lambda)-[M^{-1}G^T](p)\cdot\lambda))+\\ &\qquad +[SM^{-1}](t)\delta(t)+\mathcal{O}(\|\hat{p}-p\|+\|\hat{v}-v\|)\end{aligned}$$

and with $S(t)[M^{-1}\Gamma](p(t),v(t),\lambda(t))\equiv 0$ and

$$\begin{aligned}&([M^{-1}f](\hat{p},\hat{v},\hat{\lambda})-[M^{-1}G^T](\hat{p})\cdot\hat{\lambda})-([M^{-1}f](p,v,\lambda)-[M^{-1}G^T](p)\cdot\lambda)=\\ &\qquad =[M^{-1}\Gamma](p,v,\lambda)\cdot(\hat{\lambda}-\lambda)+\mathcal{O}(\|\hat{p}-p\|+\|\hat{v}-v\|)\cdot(1+\|\hat{\lambda}-\lambda\|)+\mathcal{O}(\|\hat{\lambda}-\lambda\|^2)\\ &\frac{d}{dt}(S(t)(\hat{v}(t)-v(t)))=[SM^{-1}](t)\delta(t)+\mathcal{O}(D(t))+\mathcal{O}(\|\hat{p}-p\|+\|\hat{v}-v\|)\ .\end{aligned}$$

This estimate is used to prove part a) of the theorem applying Gronwall's Lemma to

$$u(t):=\begin{pmatrix}\hat{p}(t)-p(t)\\ S(t)(\hat{v}(t)-v(t))\end{pmatrix}\ .$$

(Note, that

$$\hat{v}(t)-v(t)=S(t)(\hat{v}(t)-v(t))+\mathcal{O}(\|[G(\hat{p})\hat{v}](t)-[G(p)v](t)\|)+\mathcal{O}(\|\hat{p}(t)-p(t)\|)=\mathcal{O}(\|u(t)\|+\|\dot{\theta}(t)\|)$$

since $[G(p)v](t)\equiv 0$ and $[G(\hat{p})\hat{v}](t)=\frac{d}{dt}g(\hat{p})=\dot{\theta}(t)$.)

b) If f is linear w. r. t. λ then $S(p,v)$ and $S(p,v)\dot{v}=[SM^{-1}f_0](p,v)$ do not depend on λ, the estimate

$$S(\hat{p},\hat{v})\dot{\hat{v}}-S(p,v)\dot{v}=[SM^{-1}](t)\delta(t)+\mathcal{O}(\|\hat{p}-p\|+\|\hat{v}-v\|) \tag{26}$$

follows straightforwardly. Nevertheless $\hat{p}$ and $\hat{v}$ are influenced by $\ddot{\theta}(t)$ since $\frac{d}{dt}(S(\hat{p},\hat{v})\cdot\hat{v})$ includes $\frac{d}{dt}S(\hat{p},\hat{v})$, i. e. $\dot{\hat{v}}(t)$ and thus $\hat{\lambda}(t)$. Because of

$$(\frac{d}{dt}S(\hat{p},\hat{v}))\cdot\hat{v}(t)=\frac{\partial}{\partial p}S(\hat{p},\hat{v})(\hat{v},\hat{v})-\frac{\partial}{\partial v}[M^{-1}\Gamma(GM^{-1}\Gamma)^{-1}](\hat{p},\hat{v})(G(\hat{p})\hat{v},\dot{\hat{v}})\ ,$$

$[G(\hat{p})\hat{v}](t)=\dot{\theta}(t)$, the corresponding equations for p and v and (25) we get

$$(\frac{d}{dt}S(\hat{p},\hat{v}))\cdot\hat{v}(t)-(\frac{d}{dt}S(p,v))\cdot v(t)=\mathcal{O}(\|\hat{p}-p\|+\|\hat{v}-v\|)+\mathcal{O}(\|\dot{\theta}(t)\|)+\mathcal{O}(D(t))\ . \tag{27}$$

Combining (26) and (27) part b) of the theorem is proved applying Gronwall's Lemma to

$$u(t):=\begin{pmatrix}\hat{p}(t)-p(t)\\ [S(\hat{p},\hat{v})\hat{v}](t)-[S(p,v)v](t)\end{pmatrix}\ .\quad\blacksquare$$

Remark 5

a) Theorem 4 is proved following the basic lines of a corresponding result for systems of index 2 in [ASW93] were not only the ideas but also all technical details are given.

b) In contrast to systems without friction p and v are now affected by $\ddot{\theta}(t)$ but nevertheless the error bounds for p and v are again substantially smaller than that for the Lagrangian multipliers λ. This result is not restricted to systems of the special form (1): it can be proved in the same way for the differential components in arbitrary systems of index 3 in Hessenberg form.

c) As for systems without friction error bounds for the numerical solution can be given (but not for backward Euler method, see the remark after (10)). In [HLR89, Section 6] such bounds are derived in the proof of convergence for implicit Runge–Kutta methods. There the perturbation analysis is simplified by the restrictive condition $\theta = \mathcal{O}(h^3)$ that guarantees $\frac{1}{h^2}\theta = \mathcal{O}(h)$, i. e. the discretization of $\|\ddot{\theta}(\tau)\|$ in Theorem 4 is $\mathcal{O}(h)$.

d) For GGL formulation and index-2 formulation the error in p, v, λ is bounded in terms of $\|\delta(\tau)\|$, $\|\theta(\tau)\|$, $\|\dot{\theta}(\tau)\|$ with smaller bounds for p and v ([ASW93, Theorem 1a]). If the applied forces depend non-linearly on the constraint forces then the systems are non-linear in the algebraic variables and the error bounds are larger than for systems without friction.

Error bounds for the analytical and numerical solution of higher index differential-algebraic systems have been given before in the literature (e. g. [HLR89], [Mär89]). The aim of the perturbation analysis in [ASW93] and in the present paper are realistic error bounds for *all* solution components that exploit the structure of the differential-algebraic system and need as less assumptions to the size of the perturbations as possible. Thus examples have to be given to illustrate that the error bounds can in general not be improved. This is non-trivial for systems that have the very special structure of the GGL formulation ([ASW93, Section 4]).

	Theorem 1	Theorem 4a	Theorem 4b
$f_0(p,v,\lambda)$	$v_1^2 + v_2^2$	$\lambda_1^2 + \lambda_2^2$	$v_1\lambda_2 - v_2\lambda_1$
$\hat{p}_3(t)$	$\frac{1}{2}\frac{\Theta^2}{\varepsilon^2}(t-t_0)^2$	$\frac{1}{2}\frac{\Theta^2}{\varepsilon^4}(t-t_0)^2$	$\frac{1}{2}\frac{\Theta^2}{\varepsilon^3}(t-t_0)^2$
$\|\hat{p}_3 - p_3\|$	$\mathcal{O}(\frac{\Theta^2}{\varepsilon^2}) = \mathcal{O}(\|\dot{\theta}\|^2)$	$\mathcal{O}(\frac{\Theta^2}{\varepsilon^4}) = \mathcal{O}(\|\ddot{\theta}\|^2)$	$\mathcal{O}(\frac{\Theta^2}{\varepsilon^3}) = \mathcal{O}(\|\dot{\theta}\| \cdot \|\ddot{\theta}\|)$
$\|\hat{v}_3 - v_3\|$	$\mathcal{O}(\frac{\Theta^2}{\varepsilon^2}) = \mathcal{O}(\|\dot{\theta}\|^2)$	$\mathcal{O}(\frac{\Theta^2}{\varepsilon^4}) = \mathcal{O}(\|\ddot{\theta}\|^2)$	$\mathcal{O}(\frac{\Theta^2}{\varepsilon^3}) = \mathcal{O}(\|\dot{\theta}\| \cdot \|\ddot{\theta}\|)$
$\|\hat{v}_1 - v_1\|$	$\mathcal{O}(\frac{\Theta}{\varepsilon}) = \mathcal{O}(\|\dot{\theta}\|)$	$\mathcal{O}(\frac{\Theta}{\varepsilon}) = \mathcal{O}(\|\dot{\theta}\|)$	$\mathcal{O}(\frac{\Theta}{\varepsilon}) = \mathcal{O}(\|\dot{\theta}\|)$
$\|\hat{\lambda}_1 - \lambda_1\|$	$\mathcal{O}(\frac{\Theta}{\varepsilon^2}) = \mathcal{O}(\|\ddot{\theta}\|)$	$\mathcal{O}(\frac{\Theta}{\varepsilon^2}) = \mathcal{O}(\|\ddot{\theta}\|)$	$\mathcal{O}(\frac{\Theta}{\varepsilon^2}) = \mathcal{O}(\|\ddot{\theta}\|)$

Table 2: Examples illustrating the error bounds of Theorems 1 and 4

Example 3 To prove that the bounds of Theorems 1 and 4 are optimal, systems (1) with $n_p = 3$, $n_\lambda = 2$ and

$$M(p) \equiv I \ , \ f(p,v,\lambda) = \begin{pmatrix} 0 \\ 0 \\ f_0(p,v,\lambda) \end{pmatrix} \ , \ g(p) = g(p_1,p_2,p_3) = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = 0 \ , \ p(t_0) = \dot{p}(t_0) = 0$$

with various functions $f_0(p,v,\lambda)$ are considered (Table 2). Let functions $\hat{p}(t)$, $\hat{\lambda}(t)$ be given with $\hat{p}_3$ from Table 2 and

$$\hat{p}_1(t) := \Theta \sin\frac{t}{\varepsilon} \ , \ \hat{p}_2(t) := \Theta\cos\frac{t}{\varepsilon} \ , \ \hat{\lambda}_1 = -\ddot{\hat{p}}_1 \ , \ \hat{\lambda}_2 = -\ddot{\hat{p}}_2$$

where Θ and ε denote (small) positive parameters. Inserting $\hat{p}$, $\hat{\lambda}$ into (2) residuals $\delta(t) \equiv 0$, $\theta(t) = (\,\Theta \sin t/\varepsilon\,,\,\Theta\cos t/\varepsilon\,)^T$ are obtained, i. e. $\|\theta(t)\| = \mathcal{O}(\Theta)$, $\|\dot{\theta}(t)\| = \mathcal{O}(\frac{1}{\varepsilon}\Theta)$, $\|\ddot{\theta}(t)\| = \mathcal{O}(\frac{1}{\varepsilon^2}\Theta)$. (Similar examples for systems of index 2 are given in [Lub89], [ASW93, Example 2].)

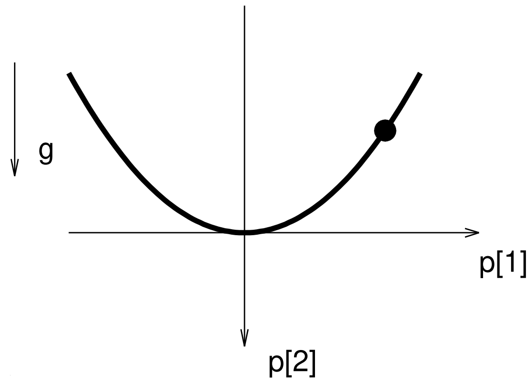
The analytical solutions $p(t)$, $\lambda(t)$ of these systems vanish identically such that the estimates in the lower part of Table 2 follow from the definition of $\hat{p}_3(t)$. In the example for Theorem 1 we have $S(p)v = (\,0\,,\,0\,,\,v_3\,)^T$ and $G(p)v = (\,v_1\,,\,v_2\,)^T$. The errors in all solution components are of the size that has to be expected from Theorems 1 and 4.
For simplicity this example is restricted to perturbations in the algebraic part. With a modified definition of $\hat{p}_3(t)$ and $\hat{\lambda}(t)$ it can be proved that the error contains also the terms $\int SM^{-1}\delta$, $(\max_\tau \|\delta(\tau)\|)^2$, ... Because of their very simple structure these examples can be extended to illustrate the errors in the numerical solution. Backward Euler method integrates these systems exactly: $p_n = v_n = \lambda_n = 0$. The discrete analogue of Table 2 ($\|\theta(\tau)\| \to \|\theta_n\|$, $\|\dot\theta(\tau)\| \to \frac{1}{h}\|\theta_n\|$, ...) is obtained for $\theta_n := (\,\Theta \sin n\frac{\pi}{2}\,,\,\Theta\cos n\frac{\pi}{2}\,)^T$.

5 Errors in the dynamical simulation: An example

In Sections 2, 3 and 4 bounds for the error in the analytical and numerical solution are given. In practical computations the *global* error is of interest, i. e. the additional error term (see Remark 4a) has to be compared with the discretization error. Tests with variable stepsize codes (Example 1) underline the importance of a perturbation analysis but show only the qualitative behaviour of the additional error term. For a quantitative analysis the error amplification during integration has to be isolated from other effects that are caused e. g. by details of the stepsize control algorithm. In the following numerical test system (1) is integrated with fixed stepsize h by backward Euler method. To illustrate the additional error term a "randomly" chosen artificial error of size 10^{-6} is introduced into the algebraic constraints.

Example 4
The motion of a point mass m on the parabola $p_2 + p_1^2 = 0$ with the force of gravitation acting in p_2-direction is described by

$$m\ddot{p} = \begin{pmatrix} 0 \\ mg \end{pmatrix} + F_R(p,\dot{p},\lambda) - \lambda \begin{pmatrix} 2p_1 \\ 1 \end{pmatrix}$$



with friction force $F_R(p,\dot{p},\lambda)$ and constraint force $F_N := -\lambda(\,2p_1\,,\,1\,)^T$. The system is modelled with Coulomb friction, i. e. F_R acts opposite to the velocity vector $\dot{p}$ and $|F_R| = \mu|F_N|$ with μ the coefficient of friction ($\mu = 0.03\ldots0.09$ for contact steel-steel, [Bud78, I.§11]):

$$F_R(p,\dot{p},\lambda) := -\mu\frac{|\lambda|\sqrt{1+4p_1^2}}{\sqrt{\dot{p}_1^2+\dot{p}_2^2}}\begin{pmatrix} \dot{p}_1 \\ \dot{p}_2 \end{pmatrix} \quad .$$

Initial values are set to $p(0) = (\,1\,,\,-1\,)^T$, $\dot{p}(0) = 0$, ($m = 1\,\mathrm{kg}$). The initial value problem was solved by fixed stepsize backward Euler method on a PC with machine eps $\approx 10^{-19}$, the stopping bound for Newton's method is set to 10^{-10}. We added an artificial perturbation to the algebraic constraints: $g(p_n) = 30\cdot\theta(t_n)$ with $\theta(t)$ the difference between $\sin t$ in `extended`-arithmetic (10 bytes) and $\sin t$ in `single`-arithmetic (4 bytes), $\|30\cdot\theta\| \approx 10^{-6}$.
On a fixed time grid $t_i = i\cdot 0.01\,\mathrm{sec}$, ($i = 1,\ldots,500$) this (perturbed) numerical solution is compared with the exact solution of the initial value problem (that is approximated by a numerical solution of high accuracy). Figure 2 shows the errors in position and velocity coordinates, in the projection $S(p,v)v$ and in the Lagrangian multipliers for various values of h. Each diagramm summarizes the results for a simulation without friction ($\mu = 0$, marked with crosses) and for a

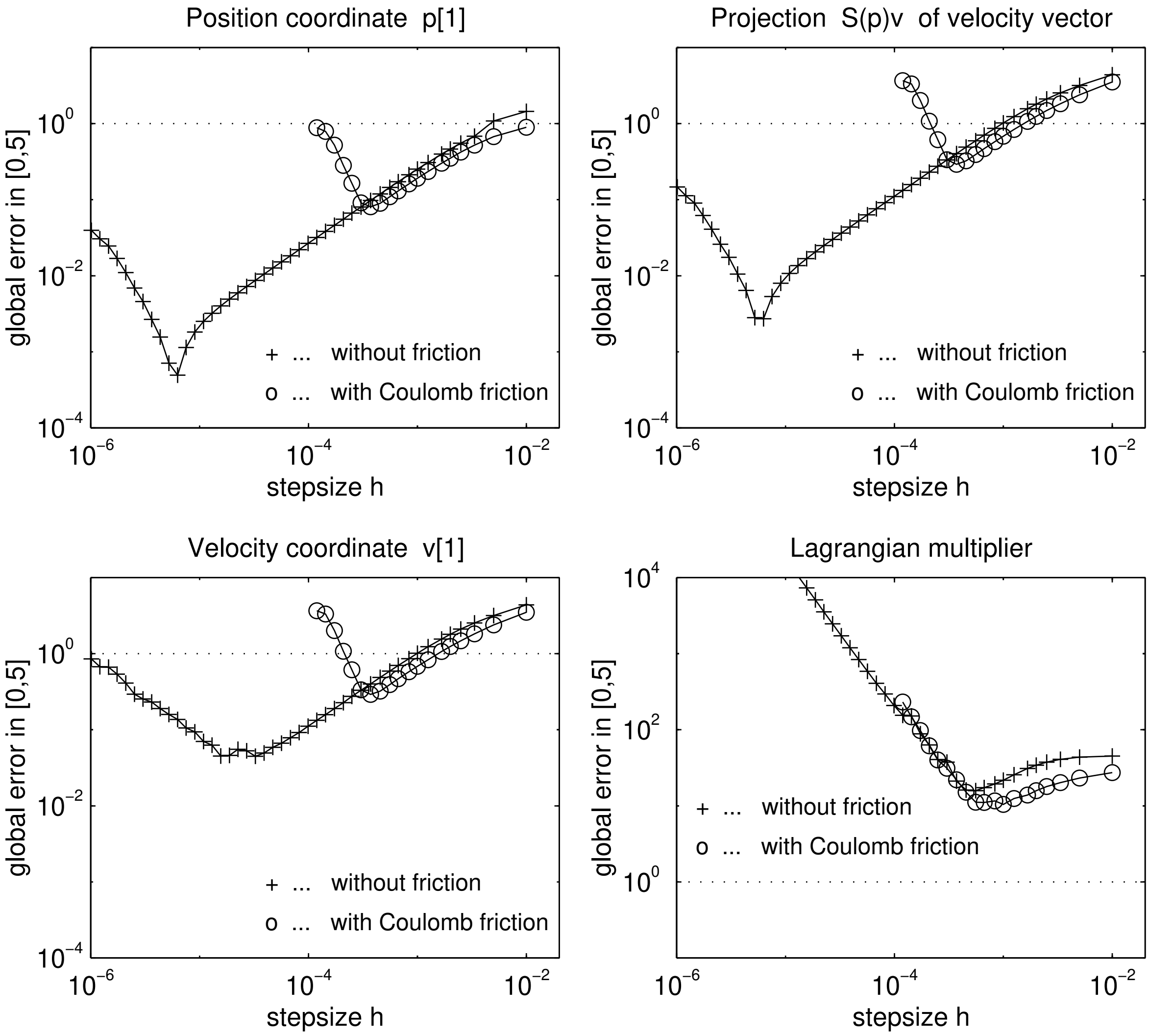


Figure 2: Global error in the dynamical simulation of a point mass moving on a parabola

simulation with Coulomb friction ($\mu = 0.03$, marked with circles).
For larger stepsizes the discretization error dominates in components p and v, the additional error term is negligible, slope $+1$ of the graphs in double logarithmic scale corresponds to the discretization error $\mathcal{O}(h)$ of backward Euler method. The error is substantially larger in the Lagrangian multipliers: it grows with $1/h^2$, the slope of the graph is -2 (note, that in this plot the scale ranges from 10^{-1} to 10^4 but in all other plots from 10^{-4} to 10^1). For smaller stepsizes the error in p and v is not further decreased, it is now dominated by the additional error term that tends to ∞ if $h \to 0$. If no friction occurs then errors of size θ/h in $G(p)v$ and of size θ^2/h^2 in p and $S(p)v$ have to be expected (Theorem 3). The numerical test is in very good accordance with this theoretical result: the errors in p and $S(p)v$ are much smaller than in v (coefficient θ^2 instead of θ) and the graphs have slope -2 in p_1, $S(p)v$ (corresponding to $1/h^2$) and slope -1 in v_1 (corresponding to $1/h$), respectively.
The situation changes drastically if friction occurs: the errors in λ are of the same size as before but p and v are much more affected by the perturbations, furthermore the errors in position and velocity coordinates are now approximately of the same size.

Example 4 underlines that the error bounds of Theorems 1, 3 and 4 reflect the influence of small perturbations correctly for *all* solution components. If the errors in all solution components would be considered simultaneously (as in the definition of the perturbation index) then there would be no great difference between systems with and without friction since the overall error is dominated by the error in λ. Variable stepsize codes base stepsize control on estimates for the local discretization error in the differential components (i. e. p and v, [HLR89, Section 8]). Thus they are robust against errors in the Lagrangian multipliers but (very) sensitive to errors in p, v. The different results for $\mu = 0$ and $\mu = 0.03$ in Example 4 illustrate why the dynamical simulation of constrained mechanical systems causes in general much more trouble if friction occurs.

In practical computations the situation is not as bad as suggested by Figure 2 where the perturbations are always of size 10^{-6}. The additional error term is substantially smaller if the perturbations are kept very small during integration. In the code RADAU5 (see Example 1) the influence of perturbations on the stepsize control is reduced furthermore since the scaled norm $\|(p, v, \lambda)\| := \|p\| + h\|v\| + h^2\|\lambda\|$ is used. If, however, the stepsize h is very small and the error in p increases if h is decreased ($h \leq 10^{-5}$ in Figure 2, $\mu = 0$), then any variable stepsize integrator fails and the integration breaks down.

Summary

The amplification of small errors causes problems in the numerical solution of the model equations (1) for constrained mechanical systems. The perturbation analysis gives bounds for this additional error term that grows if the discretization stepsize is decreased. The error in position and velocity coordinates is substantially smaller than that in the Lagrangian multipliers. If the applied forces in the system (1) of *index 3* do not depend on the constraint forces ("systems without friction") then position and velocity coordinates are affected by perturbations in the same way as the solution of *index-2* systems in Hessenberg form. The typical error term $\frac{1}{h^2}\theta$ in systems of index 3 (that corresponds to $\|\ddot{\theta}(t)\|$ in the definition of perturbation index) appears in the Lagrangian multipliers, only. Theoretical results and a numerical test illustrate that this can not be generalized to systems with friction.

An efficient and reliable dynamical simulation of constrained mechanical systems requires that the additional error term is kept small without large additional computational effort. The perturbation analysis suggests two ways to do this: If (1) is discretized as system of index 3 then stepsize control should be based on components p, v. The systems of nonlinear equations have to be solved with very high accuracy. The alternative are discretization methods based on index reduced formulations of the model equations, especially the GGL formulation.

References

[Arn93] M. Arnold. Stability of numerical methods for differential–algebraic equations of higher index. *APNUM*, 13:5–14, 1993.

[Arn94] M. Arnold. Numerische Probleme in der dynamischen Simulation von Rad–Schiene–Systemen. *submitted to: ZAMM*, 1994.

[ASW93] M. Arnold, K. Strehmel, and R. Weiner. Small perturbations in differential-algebraic systems of index 2. Preprint 93/1, University of Rostock, Department of Mathematics, 1993, *submitted to: BIT*.

[AP93] U.M. Ascher and L.R. Petzold. Stability of computational methods for constrained dynamics systems. *SIAM J. Sci. Comput.*, 14:95–120, 1993.

[BCP89] K.E. Brenan, S.L. Campbell, and L.R. Petzold. *Numerical solution of initial–value problems in differential–algebraic equations*. North-Holland Publ. Co., Amsterdam, 1989.

[Bud78] A. Budó. *Theoretische Mechanik*. Deutscher Verlag der Wissenschaften, Berlin, 9th edition, 1978.

[FL91] C. Führer and B. Leimkuhler. Numerical solution of differential–algebraic equations for constrained mechanical motion. *Numer. Math.*, 59:55–69, 1991.

[GLG85] C.W. Gear, B. Leimkuhler, and G.K. Gupta. Automatic integration of Euler-Lagrange equations with constraints. *J. Comp. Appl. Math.*, 12&13:77–90, 1985.

[HLR89] E. Hairer, Ch. Lubich, and M. Roche. *The numerical solution of differential-algebraic systems by Runge–Kutta methods*. Lecture Notes in Mathematics, 1409. Springer-Verlag, Berlin Heidelberg New York, 1989.

[HNW93] E. Hairer, S.P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations.I. Nonstiff Problems*. Springer-Verlag, Berlin Heidelberg New York, 2nd edition, 1993.

[HW91] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations.II. Stiff and Differential-Algebraic Problems*. Springer-Verlag, Berlin Heidelberg New York, 1991.

[LP86] P. Lötstedt and L.R. Petzold. Numerical solution of nonlinear differential equations with algebraic constraints I: convergence results for backward differentiation formulas. *Math. of Comp.*, 46:491–516, 1986.

[Lub89] Ch. Lubich. Linearly implicit extrapolation methods for differential-algebraic systems. *Numer. Math.*, 55:197–211, 1989.

[Mär89] R. März. Analysis and numerical treatment of differential-algebraic systems. In R.E. Bank, R. Bulirsch, and K. Merten, editors, *Mathematical Modelling and Simulation of Electrical Circuits and Semiconductor Devices. Proc. of a Conference held at the Math. Forschungsinstitut, Oberwolfach, Oct. 30 – Nov. 5, 1988*, pages 27–43, Basel, 1989. Birkhäuser-Verlag.

[NA93] H. Netter and M. Arnold. Geometrie und Dynamik eines Rad-Schiene-Modells in Deskriptorform mit unstetigen Zustandsgrößen. Technical Report IB 515–93–02, DLR, D-5000 Köln 90, 1993.