

Enhancement of the Adjoint Method by Error Control of Accelerations for Parameter Identification in Multibody Dynamics

Karin Nachbagauer^{1,*}, Stefan Oberpeilsteiner², Wolfgang Steiner¹

¹Faculty of Engineering and Environmental Sciences, University of Applied Sciences Upper Austria, Stelzhamerstrasse 23, 4600 Wels, Austria

²Institute of Mechanics and Mechatronics, Vienna University of Technology, Getreidemarkt 9/325, 1060 Wien, Austria

Copyright ©2015 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract The present paper shows the embedding of the adjoint method in multibody dynamics and its broad applicability for examples for both, parameter identification and optimal control. Especially, in case of parameter identifications in engineering multibody applications, a theoretical enhancement of the proposed adjoint method by an error control of accelerations is inevitable in order to meet the circumstances of experimental studies using acceleration sensors in general.

Keywords Multibody Dynamics, Adjoint Method, Optimal Control, Parameter Identification

1 Introduction

In the explanations of Giles and Pierce [1], the adjoint method is seen as a special case of linear duality, in which the dual problem has to solve only a single linear system. Obviously, huge benefits can be achieved by solving the dual formulation. The adjoint method utilizes this powerful aspect of duality to dramatically improve the efficiency of the computation.

The adjoint method is probably the most efficient way to solve a wide range of optimization problems in engineering sciences. In the sensitivity analysis, the adjoint method has been utilized by various authors, e.g., [12, 8, 7], and in dynamic programming methods the adjoint approach has been used extensively for the computation of gradients as well [10].

The adjoint method is applied in, e.g., aerodynamic shape optimization by Jameson [13], in which the gradient of an objective function is determined indirectly by solving an adjoint equation which has coefficients determined by the solution of the multibody dynamics equations. This directly corresponds to the gradient technique for trajectory optimization pioneered by Bryson and Ho [9]. The fast computation of the gradients makes optimization computationally feasible even for complex three-dimensional multibody systems.

There are two strategies for this purpose, the equations of motion of the multibody system and adjoint equations may either be separately discretized from their representations as differential-algebraic equations, or, alternatively, the equations of motion of the multibody system may be discretized first, while the discrete adjoint equations are derived directly from the discrete multibody system equations [9].

Previous work on the adjoint method in multibody dynamics can be found in [14], where the solution of inverse dynamics and trajectory optimization problems for multibody systems is reflected by an indirect approach combining optimal control theory with control and adjoint equations with transversality conditions.

The design of the indirect method for solving optimal control problems for multibody systems presented in [18] seems to be similar to the idea to solve adjoint equations.

The work by Schaffer [15] proposes a numerical algorithm, the piecewise adjoint method, which formulates the coordinate partitioning underlying ordinary differential equations as a boundary value problem, which is solved by multiple shooting methods.

The group around Petzold, Cao, Li and Serban [17, 16] describe forward and adjoint methods for sensitivity analysis for differential-algebraic equations and partial differential equations and emphasizes that the results of sensitivity analysis have a wide range of applications, including model development, optimization, parameter estimation, model simplification, data assimilation, optimal control, uncertainty analysis and experimental design [17].

Eberhard [7] uses the adjoint method for sensitivity analysis in multibody systems interpreted as a continuous, hybrid form of automatic differentiation.

In case of an orientation parametrization of a body without using angles, as e.g. described in the absolute nodal coordinate formulation (ANCF) [22], a gradient-based optimization approach using adjoint equations has been proposed by Held and Seifried [19]. The framework of the ANCF combined with the adjoint method is as well used in the sensitivity analysis by Pi et al. [21] within a first order approach, while in [20] a second order adjoint sensitivity analysis of multi-

body systems has been presented within the classical floating frame of reference formulation for multibody systems. For a comparison of the computational costs between the adjoint method and the direct differentiation method see Table 4 in [20].

Despite of the dramatically decrease of computational costs and the resulting potential of the adjoint method in multibody dynamics, the adjoint method is rarely applied. The structure of the equations of motion in a multibody system is usually quite complicated, and the effort to obtain the set of adjoint equations seems tremendously high.

The main goal of the present paper is to show how the adjoint method can be embedded efficiently to a multibody system described by a system of differential-algebraic equations of index 3 for optimal control problems or parameter identification applications. The body description is, in contrast to the approach in [19], in which the flexible multibody system is displayed in minimal coordinates, defined by the position of the center of mass and the four redundant Euler parameters for the orientation parametrization. This choice includes an internal constraint for each body which has to be considered also in the constraint Jacobian of the whole system and of course also influences the optimal control equations. The present paper shows the potential of the adjoint method for solving classical optimization problems in multibody dynamics and enhances the governed equations by an error control of accelerations for parameter identification in order to meet the standards of experimental studies using acceleration sensors in general.

2 Adjoint gradient computation

The basic idea of the adjoint method, see e.g. [17, 16, 7], is summarized in brief here. Suppose we have the semi-explicit differential algebraic system

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, \mathbf{u}, t, \boldsymbol{\lambda}) \\ \mathbf{C}(\mathbf{x}) &= \mathbf{0} \\ \mathbf{x}(0) &= \mathbf{x}_0, \quad \text{satisfying } \mathbf{C}(\mathbf{x}_0) = \mathbf{0} \end{aligned} \quad (1)$$

where $\mathbf{x}(t) \in \mathbb{R}^{N_x}$ and $\boldsymbol{\lambda}(t) \in \mathbb{R}^{N_\lambda}$ are the vectors of state variables and algebraic variables of a dynamical system such as a multibody system in a redundant formulation, e.g., using Euler parameters for the parametrization of the orientation. Herein, $\mathbf{u} \in \mathbb{R}^{N_u}$ may either describe a vector of (constant) parameters or a vector of control signals. Our goal is to find \mathbf{u} such that a functional of the form

$$J(\mathbf{u}) = \int_0^T h(\mathbf{x}, \mathbf{u}, t) dt \quad (2)$$

is minimized. The focus is to determine a set of controls or parameters \mathbf{u} in a way that the measured signals $\bar{s}_i(t)$, $i = 1, \dots, N_s$ are best approximated by the system outputs $s_i(\mathbf{x})$. In this case one could define the root mean square (RMS) error as our functional to be minimized

$$J(\mathbf{u}) = \int_0^T \sum_{i=1}^{N_s} [s_i(\mathbf{x}) - \bar{s}_i(t)]^2 dt \quad (3)$$

Various methods are available to compute the argument for which a function or a functional attains a minimum, e.g., the method of the steepest descent, the conjugate gradient

method, the Gauss-Newton method or quasi Newton methods. In any cases the gradient of $J(\mathbf{u})$ has to be determined and for this purpose several strategies can be pursued again.

On the one hand, if \mathbf{u} is a vector of N_u parameters, the sensitivity equations for $\mathbf{x}_u = \partial \mathbf{x} / \partial \mathbf{u}$ are usually considered, see e.g. [11] or [23]. When choosing this sensitivity approach, the computational effort is equal to solving N_u linear sets of equations with the same dimension as Eq. (1). On the other hand, if $\mathbf{u}(t)$ is a vector of control signals, the problem could be transformed into a finite dimensional one by discretizing the control signals. The adjoint method is an efficient alternative to compute the gradient of $J(\mathbf{u})$ in both cases.

First, it has to be recognized that $J(\mathbf{u})$ does not change if we add Eq. (1) to the integrand

$$J(\mathbf{u}) = \int_0^T [h + \mathbf{y}^\top (\mathbf{f} - \dot{\mathbf{x}}) + \boldsymbol{\mu}^\top \mathbf{C}] dt \quad (4)$$

with arbitrary functions $\mathbf{y}(t)$ and $\boldsymbol{\mu}(t)$. The Hamiltonian

$$H(\mathbf{x}, \mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{u}, t) = h(\mathbf{x}, \mathbf{u}, t) + \mathbf{y}^\top \mathbf{f}(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{u}, t) + \boldsymbol{\mu}^\top \mathbf{C}(\mathbf{x}) \quad (5)$$

changes Eq. (4) to

$$J(\mathbf{u}) = \int_0^T [H - \mathbf{y}^\top \dot{\mathbf{x}}] dt \quad (6)$$

Let us now consider a forward solution $\mathbf{x}(t)$ and $\boldsymbol{\lambda}(t)$ of the system equations in Eq. (1) for a set of parameters or control variables \mathbf{u} . A variation of \mathbf{u} about $\delta \mathbf{u}$ will result in variations of $\mathbf{x}(t)$ and $\boldsymbol{\lambda}(t)$ about $\delta \mathbf{x}(t)$ and $\delta \boldsymbol{\lambda}(t)$ and, moreover, in a variation of the functional J about δJ . Up to a first order approximation δJ is given by

$$\delta J = \int_0^T [H_{\mathbf{u}} \delta \mathbf{u} + H_{\mathbf{x}} \delta \mathbf{x} + H_{\boldsymbol{\lambda}} \delta \boldsymbol{\lambda} - \mathbf{y}^\top \delta \dot{\mathbf{x}}] dt \quad (7)$$

where $H_{\mathbf{u}}$, $H_{\mathbf{x}}$ and $H_{\boldsymbol{\lambda}}$ denote the row vectors with the partial derivatives of H with respect to the components of \mathbf{u} , \mathbf{x} and $\boldsymbol{\lambda}$, e. g.,

$$H_{\mathbf{x}} = \left[\frac{\partial H}{\partial x_1} \quad \cdots \quad \frac{\partial H}{\partial x_n} \right]$$

Integrating by parts of the last term of Eq. (7) yields

$$\begin{aligned} \delta J &= \int_0^T [H_{\mathbf{u}} \delta \mathbf{u} + H_{\mathbf{x}} \delta \mathbf{x} + H_{\boldsymbol{\lambda}} \delta \boldsymbol{\lambda} + \dot{\mathbf{y}}^\top \delta \mathbf{x}] dt - \mathbf{y}^\top \delta \mathbf{x} \Big|_0^T \\ &= \int_0^T [H_{\mathbf{u}} \delta \mathbf{u} + (H_{\mathbf{x}} + \dot{\mathbf{y}}^\top) \delta \mathbf{x} + H_{\boldsymbol{\lambda}} \delta \boldsymbol{\lambda}] dt - \mathbf{y}^\top \delta \mathbf{x} \Big|_T \end{aligned} \quad (8)$$

since $\delta \mathbf{x}(0) = 0$ as the initial state is assumed to be prescribed. It is the key idea of the adjoint method that the computation of $\delta \mathbf{x}$ and $\delta \boldsymbol{\lambda}$ can be circumvented if the adjoint variables $\mathbf{y}(t)$ and $\boldsymbol{\mu}(t)$ are chosen such that

$$\dot{\mathbf{y}}^\top = -H_{\mathbf{x}}, \quad H_{\boldsymbol{\lambda}} = 0 \quad \text{and} \quad \mathbf{y}(T) = 0 \quad (9)$$

or including an other known fixed value for the boundary condition $\mathbf{y}(T)$. This set of semi-explicit differential algebraic equations is called the adjoint system of Eq. (1). It may be solved backwards in time starting at time $t = T$, once the original equations have been solved forward for $t \in [0, T]$.

Then, with $\mathbf{x}(t)$, $\boldsymbol{\lambda}(t)$, $\mathbf{y}(t)$ and $\boldsymbol{\mu}(t)$ from Eqs. (1) and (9), the variation of J according to Eq. (8) is readily given by

$$\delta J = \int_0^T H_{\mathbf{u}} \delta \mathbf{u} dt \quad (10)$$

In case \mathbf{u} is a control signal, the largest possible increase of δJ is achieved, if $\delta \mathbf{u}(t)$ is chosen in the direction of $H_{\mathbf{u}}^T$. Hence, $H_{\mathbf{u}}^T$ may be considered as the gradient of the functional $J(\mathbf{u})$.

It has to be mentioned here that in case \mathbf{u} is a control signal, the integral in Eq. (10) is computed as a sum including δu_i computed for the i -th time step and Δt is chosen as the time discretization of the control signal itself. In case \mathbf{u} is a vector of parameters, the gradient of the function $J(\mathbf{u})$ is given by $\nabla J = \int_0^T H_{\mathbf{u}}^T dt$. It has to be emphasized here, that only two systems of DAEs must be integrated for computing the direction of the gradient. The main difficulty with multibody dynamics results from the complexity of the original system (1). Hence, many authors focused on two-dimensional examples or rather general aspects, see e.g. [11], [6]. A former paper uses this key idea of the adjoint method and shows how these equations can be implemented in a MBS-code under the assumption that the rigid body rotations are described by Euler parameters, since a redundant formulation gives relatively simple adjoint equations, see [5]. The present paper extends the theory presented in [5] by enhancing the method by an error of accelerations in the cost functional.

3 The adjoint method enhanced by an error of accelerations for parameter identification

This section discusses the embedding of the adjoint gradient computation to multibody dynamics.

In case the measures included in the system are restricted to position and velocity level, the embedding of the adjoint method to multibody dynamics has been presented already in Nachbagauer et al. [5]. In case the error of accelerations has to be accounted for in the optimization functional, the adjoint method in multibody dynamics presented in [5] must be enhanced. Herein, the enhancement of the adjoint method by error control of accelerations for parameter identification is presented.

Firstly, the adjoint equations for the equations of motion of a multibody system are derived and, secondly, the structure of the boundary conditions is discussed.

3.1 Adjoint equations

A mechanical system consisting of rigid and flexible bodies, forces and constraints acting between these bodies can be described by equations of motion in the following form:

$$\begin{aligned} \mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} &= \mathbf{f}(\mathbf{q}, \dot{\mathbf{q}}, \mathbf{u}) - \mathbf{C}_{\mathbf{q}}^T(\mathbf{q})\boldsymbol{\lambda} \\ \mathbf{C}(\mathbf{q}) &= \mathbf{0} \end{aligned} \quad (11)$$

Here, \mathbf{q} denotes a vector of redundant generalized coordinates. They are subject to the holonomic constraints $\mathbf{C}(\mathbf{q}) = \mathbf{0}$, which enter the equations of motion via the Jacobian $\mathbf{C}_{\mathbf{q}}$ and the vector of Lagrange multipliers $\boldsymbol{\lambda}$ representing the

constraint forces in the system. Moreover, the system may incorporate a control $\mathbf{u}(t)$ or a vector of parameters \mathbf{u} . For simplicity, we suppose that \mathbf{u} appears only in the vector \mathbf{f} , which is the fact, if \mathbf{u} is an actuating force or a stiffness or damping parameter. Using the additional variables $\mathbf{v} = \dot{\mathbf{q}}$, the equations of motion can be reformulated as a first order system of equations as

$$\dot{\mathbf{q}} = \mathbf{v} \quad (12)$$

$$\mathbf{M}\dot{\mathbf{v}} = \mathbf{f}(\mathbf{q}, \mathbf{v}, \mathbf{u}) - \mathbf{C}_{\mathbf{q}}^T\boldsymbol{\lambda} \quad (13)$$

$$\mathbf{C}(\mathbf{q}) = \mathbf{0} \quad (14)$$

The state vector \mathbf{x} defined in the previous section now consists of \mathbf{q} and \mathbf{v} . The goal of the optimization is to find the parameter (or control) \mathbf{u} which minimizes a cost functional of the form

$$J = \int_0^T h(\mathbf{q}, \mathbf{v}, \dot{\mathbf{v}}, \mathbf{u}) dt + S(\mathbf{q}, \mathbf{v}) \Big|_T \quad (15)$$

in which $S(\mathbf{q}, \mathbf{v})|_T$ introduces an additional end point term, usually denoted as scrap function. For example, S could describe an end point RMS-error of a sensor variable similar to Eq. (3). The cost functional J is augmented by the system equations as follows

$$\begin{aligned} J = \int_0^T [& h(\mathbf{q}, \mathbf{v}, \dot{\mathbf{v}}, \mathbf{u}, t) + \mathbf{p}^T(\dot{\mathbf{q}} - \mathbf{v}) \\ & + \mathbf{w}^T(\mathbf{M}\dot{\mathbf{v}} - \mathbf{f} + \mathbf{C}_{\mathbf{q}}^T\boldsymbol{\lambda}) + \boldsymbol{\mu}^T\mathbf{C}] dt \\ & + S(\mathbf{q}, \mathbf{v}) \Big|_{t=T} \end{aligned} \quad (16)$$

The variables $\mathbf{p}(t)$, $\mathbf{w}(t)$ and $\boldsymbol{\mu}(t)$ may be chosen arbitrarily here. The variation of the cost functional J is derived by

$$\begin{aligned} \delta J = \int_0^T \left\{ & h_{\mathbf{q}}\delta\mathbf{q} + h_{\mathbf{v}}\delta\mathbf{v} + h_{\dot{\mathbf{v}}}\delta\dot{\mathbf{v}} + h_{\mathbf{u}}\delta\mathbf{u} \right. \\ & + \mathbf{p}^T(\delta\dot{\mathbf{q}} - \delta\mathbf{v}) + \\ & \left. + \mathbf{w}^T\mathbf{D} + \boldsymbol{\mu}^T\mathbf{C}_{\mathbf{q}}\delta\mathbf{q} \right\} dt \\ & + \mathbf{S}_{\mathbf{q}}\delta\mathbf{q} \Big|_{t=T} + \mathbf{S}_{\mathbf{v}}\delta\mathbf{v} \Big|_{t=T} \end{aligned} \quad (17)$$

with

$$\begin{aligned} \mathbf{D} = & (\mathbf{M}\dot{\mathbf{v}})_{\mathbf{q}} \delta\mathbf{q} + \mathbf{M}\delta\dot{\mathbf{v}} - \mathbf{f}_{\mathbf{q}}\delta\mathbf{q} - \mathbf{f}_{\mathbf{v}}\delta\mathbf{v} - \mathbf{f}_{\mathbf{u}}\delta\mathbf{u} \\ & + (\mathbf{C}_{\mathbf{q}}^T\boldsymbol{\lambda})_{\mathbf{q}} \delta\mathbf{q} + \mathbf{C}_{\mathbf{q}}^T\delta\boldsymbol{\lambda} \end{aligned} \quad (18)$$

Integration by parts for the terms including $\delta\dot{\mathbf{q}}$ and $\delta\dot{\mathbf{v}}$ is computed by

$$\begin{aligned} \int_0^T \mathbf{p}^T \delta\dot{\mathbf{q}} dt &= - \int_0^T \dot{\mathbf{p}}^T \delta\mathbf{q} dt + \mathbf{p}^T \delta\mathbf{q} \Big|_{t=T} \\ \int_0^T \mathbf{w}^T \mathbf{M} \delta\dot{\mathbf{v}} dt &= - \int_0^T \frac{d}{dt} (\mathbf{w}^T \mathbf{M}) \delta\mathbf{v} dt + \mathbf{w}^T \mathbf{M} \delta\mathbf{v} \Big|_{t=T} \\ \int_0^T h_{\dot{\mathbf{v}}} \delta\dot{\mathbf{v}} dt &= - \int_0^T \frac{d}{dt} (h_{\dot{\mathbf{v}}}) \delta\mathbf{v} dt + h_{\dot{\mathbf{v}}} \delta\mathbf{v} \Big|_{t=T} \end{aligned} \quad (19)$$

with $\delta\mathbf{q}(0) = \delta\mathbf{v}(0) = 0$ since initial conditions for the adjoint variables are prescribed as $\mathbf{q}(0) = \mathbf{q}_0$ and $\mathbf{v}(0) = \mathbf{v}_0$.

Using Eqs. (19) and extracting the terms multiplied with $\delta\mathbf{q}$, $\delta\mathbf{v}$, $\delta\boldsymbol{\lambda}$ and $\delta\mathbf{u}$, the variation of the functional J given by Eq. (17) can be rewritten as

$$\begin{aligned} \delta J = \int_0^T & \left\{ [h_{\mathbf{q}} - \dot{\mathbf{p}}^\top + \mathbf{w}^\top ((\mathbf{M}\dot{\mathbf{v}})_{\mathbf{q}} - \mathbf{f}_{\mathbf{q}} + (\mathbf{C}_{\mathbf{q}}^\top \boldsymbol{\lambda})_{\mathbf{q}}) \right. \\ & + \boldsymbol{\mu}^\top \mathbf{C}_{\mathbf{q}}] \delta \mathbf{q} \\ & + \left[h_{\mathbf{v}} - \mathbf{p}^\top - \mathbf{w}^\top \mathbf{f}_{\mathbf{v}} - \frac{d}{dt} (\mathbf{w}^\top \mathbf{M}) - \frac{d}{dt} h_{\dot{\mathbf{v}}} \right] \delta \mathbf{v} \\ & + [\mathbf{w}^\top \mathbf{C}_{\mathbf{q}}^\top] \delta \boldsymbol{\lambda} + [h_{\mathbf{u}} - \mathbf{w}^\top \mathbf{f}_{\mathbf{u}}] \delta \mathbf{u} \left. \right\} dt \\ & + [\mathbf{S}_{\mathbf{q}} + \mathbf{p}^\top] \delta \mathbf{q} \Big|_{t=T} \\ & + [\mathbf{S}_{\mathbf{v}} + \mathbf{w}^\top \mathbf{M} + h_{\dot{\mathbf{v}}}] \delta \mathbf{v} \Big|_{t=T} \end{aligned} \quad (20)$$

To eliminate the terms involving $\delta\mathbf{q}$, $\delta\mathbf{v}$ the adjoint variables \mathbf{p} , \mathbf{w} and $\boldsymbol{\mu}$ may now be defined by equating the respective expressions in square brackets to zero. After transposing these expressions one obtains the following system of adjoint equations:

$$\frac{d\mathbf{p}}{dt} = h_{\mathbf{q}}^\top + \mathbf{A}\mathbf{w} + \mathbf{C}_{\mathbf{q}}^\top \boldsymbol{\mu} \quad (21)$$

$$\frac{d}{dt} (\mathbf{M}\mathbf{w}) = h_{\dot{\mathbf{v}}}^\top - \mathbf{p} - \mathbf{B}\mathbf{w} - \frac{d}{dt} (h_{\dot{\mathbf{v}}})^\top \quad (22)$$

$$\mathbf{0} = \mathbf{C}_{\mathbf{q}} \mathbf{w} \quad (23)$$

$$\mathbf{0} = \mathbf{S}_{\mathbf{q}}^\top + \mathbf{p}(T) \quad (24)$$

$$\mathbf{0} = \mathbf{S}_{\mathbf{v}}^\top + \mathbf{M}(T)\mathbf{w}(T) + h_{\dot{\mathbf{v}}}^\top \quad (25)$$

where the abbreviations

$$\mathbf{A} = (\mathbf{M}\dot{\mathbf{v}})_{\mathbf{q}}^\top - \mathbf{f}_{\mathbf{q}}^\top + (\mathbf{C}_{\mathbf{q}}^\top \boldsymbol{\lambda})_{\mathbf{q}}^\top \quad (26)$$

$$\mathbf{B} = \mathbf{f}_{\dot{\mathbf{v}}}^\top \quad (27)$$

and simplifications due to the symmetry of the mass matrix $\mathbf{M} = \mathbf{M}^\top$ have been used. Note that in case a root mean square error has to be minimized, as similar to Eq. (3),

$$h = \frac{1}{2} (s(\mathbf{q}, \mathbf{v}, \dot{\mathbf{v}}) - \bar{s}(t))^2 \quad (28)$$

in which $\bar{s}(t)$ is a measured acceleration signal which only depends on time, the derivative $h_{\dot{\mathbf{v}}}$ can be simply computed by

$$h_{\dot{\mathbf{v}}} = (s - \bar{s}) s_{\dot{\mathbf{v}}} \quad (29)$$

in which the abbreviations $s = s(\mathbf{q}, \mathbf{v}, \dot{\mathbf{v}})$ and $\bar{s} = \bar{s}(t)$ are used for sake of simpler reading. The time derivative $\frac{d}{dt} (h_{\dot{\mathbf{v}}})$ in Eq. (22) can be computed by

$$\frac{d}{dt} (h_{\dot{\mathbf{v}}}) = \frac{d}{dt} (s - \bar{s}) s_{\dot{\mathbf{v}}} + (s - \bar{s}) \frac{d}{dt} s_{\dot{\mathbf{v}}} \quad (30)$$

$$\begin{aligned} &= (s_{\mathbf{q}} \dot{\mathbf{q}} + s_{\mathbf{v}} \dot{\mathbf{v}} + s_{\dot{\mathbf{v}}} \ddot{\mathbf{v}} - \dot{\bar{s}}) s_{\dot{\mathbf{v}}} \\ &+ (s - \bar{s}) (s_{\dot{\mathbf{v}}\mathbf{q}} \dot{\mathbf{q}} + s_{\dot{\mathbf{v}}\mathbf{v}} \dot{\mathbf{v}} + s_{\dot{\mathbf{v}}\dot{\mathbf{v}}} \ddot{\mathbf{v}}) \end{aligned} \quad (31)$$

in which $\ddot{\mathbf{v}} = \ddot{\mathbf{q}}$ can be approximated by Taylor's theorem when interpolating \mathbf{q} up to order three while ignoring higher order terms. In case an end point error must be included in the computation and the scrap function $S(\mathbf{q}, \mathbf{v})$ in Eq. (15) depends as well on $\dot{\mathbf{v}}$, the derivations in Eqs. (28)–(30) have

to be performed similarly for $S(\mathbf{q}, \mathbf{v}, \dot{\mathbf{v}})$ as well.

Concluding this derivation, in case Eqs. (21)–(25) are satisfied the variation of the cost functional in Eq. (20) reduces to

$$\delta J = \int_0^T [h_{\mathbf{u}} - \mathbf{w}^\top \mathbf{f}_{\mathbf{u}}] \delta \mathbf{u} dt \quad (32)$$

which directly relates the independent variation $\delta\mathbf{u}$ to the variation of the objective function.

3.2 Consistent boundary conditions for the adjoint system

It has to be noticed that the boundary condition (25) for the adjoint variable \mathbf{w} is generally incompatible with the constraint equation (23) at time $t = T$. Only in case $S_{\dot{\mathbf{v}}} = 0$, i. e. when the scrap function does not depend on $\dot{\mathbf{v}}$, all equations are satisfied by setting $\mathbf{p}(T) = -\mathbf{S}_{\mathbf{q}}^\top$ and $\mathbf{w}(T) = \mathbf{0}$. However, we run into trouble if $S_{\dot{\mathbf{v}}} \neq 0$. To circumvent this issue an approach similar to the idea of Gear-Gupta-Leimkuhler [3] is applied by considering the constraint equations $\mathbf{C}(\mathbf{q}) = \mathbf{0}$ of the original system Eq. (11) at velocity level and its variation with respect to \mathbf{v} and \mathbf{q} results in

$$\mathbf{C}_{\mathbf{q}} \dot{\mathbf{q}} = \mathbf{C}_{\mathbf{q}}(\mathbf{q})\mathbf{v} = \mathbf{0} \quad (33)$$

$$\mathbf{C}_{\mathbf{q}} \delta \mathbf{v} + (\mathbf{C}_{\mathbf{q}\mathbf{v}})_{\mathbf{q}} \delta \mathbf{q} = \mathbf{0} \quad (34)$$

This relation is considered at time $t = T$ and is multiplied by an arbitrary vector of numbers $\boldsymbol{\xi}$ leading to

$$\boldsymbol{\xi}^\top \left(\mathbf{C}_{\mathbf{q}} \delta \mathbf{v} + (\mathbf{C}_{\mathbf{q}\mathbf{v}})_{\mathbf{q}} \delta \mathbf{q} \right) \Big|_T = \mathbf{0} \quad (35)$$

Since the expression on the left side is always zero one could add it to Eq. (20) without modifying the variation of our objective functional. Hence, two additional boundary terms in Eq. (20) arise transforming Eqs. (24) and (25) into

$$\mathbf{S}_{\mathbf{q}}^\top + \mathbf{p} + (\mathbf{C}_{\mathbf{q}\mathbf{v}})_{\mathbf{q}}^\top \boldsymbol{\xi} = \mathbf{0} \quad \dots \text{at } t = T \quad (36)$$

$$\mathbf{S}_{\mathbf{v}}^\top + \mathbf{M}\mathbf{w} + (h_{\dot{\mathbf{v}}})^\top + \mathbf{C}_{\mathbf{q}}^\top \boldsymbol{\xi} = \mathbf{0} \quad \dots \text{at } t = T \quad (37)$$

Using the still undetermined variable $\boldsymbol{\xi}$, the boundary conditions $\mathbf{p}(T)$, $\mathbf{w}(T)$ can be computed in such a way that also the constraint equation (23) is satisfied at time $t = T$. For that purpose, the following system of equations has to be solved for $\mathbf{p}(T)$, $\mathbf{w}(T)$ and $\boldsymbol{\xi}$ at time $t = T$:

$$\begin{bmatrix} \mathbf{I} & \mathbf{0} & (\mathbf{C}_{\mathbf{q}\mathbf{v}})_{\mathbf{q}}^\top \\ \mathbf{0} & \mathbf{M} & \mathbf{C}_{\mathbf{q}}^\top \\ \mathbf{0} & \mathbf{C}_{\mathbf{q}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{p} \\ \mathbf{w} \\ \boldsymbol{\xi} \end{bmatrix} = \begin{bmatrix} -\mathbf{S}_{\mathbf{q}}^\top \\ -\mathbf{S}_{\mathbf{v}}^\top - (h_{\dot{\mathbf{v}}})^\top \\ \mathbf{0} \end{bmatrix} \quad (38)$$

This means that one may first solve at time $t = T$:

$$\begin{bmatrix} \mathbf{M} & \mathbf{C}_{\mathbf{q}}^\top \\ \mathbf{C}_{\mathbf{q}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\xi} \end{bmatrix} = \begin{bmatrix} -\mathbf{S}_{\mathbf{v}}^\top - (h_{\dot{\mathbf{v}}})^\top \\ \mathbf{0} \end{bmatrix} \quad (39)$$

yielding $\boldsymbol{\xi}$ and $\mathbf{w}(T)$, and compute then $\mathbf{p}(T)$ at time $t = T$ from

$$\mathbf{p} = -(\mathbf{C}_{\mathbf{q}\mathbf{v}})_{\mathbf{q}}^\top \boldsymbol{\xi} - \mathbf{S}_{\mathbf{q}}^\top \quad (40)$$

Once $\mathbf{w}(T)$ and $\mathbf{p}(T)$ has been determined in the correct way, the differential-algebraic set of adjoint equations

(21)–(23) may be solved to obtain \mathbf{w} , \mathbf{p} and $\boldsymbol{\mu}$ at every time instant in the interval $[0, T]$.

For the adjoint equations of a multibody system the matrices \mathbf{M} , \mathbf{C}_q , \mathbf{A} and \mathbf{B} from Eqs. (26) and (27) are required along a forward simulation of the equations of motion as well. The matrices \mathbf{M} , \mathbf{C}_q also appear in the equations of motion and are therefore available, while the determination of the Jacobian matrices \mathbf{A} and \mathbf{B} requires additional computational effort.

In [5], more details on the different computation strategies for the Jacobians have been presented. Beside exact computation of the Jacobian matrices by implementing explicit formulas in the MBS-code, one could also compute the derivatives numerically by substituting the derivatives by finite difference quotients or investigate the technique of automatic differentiation. Despite the explicit formulas seems expensive and susceptible to programming errors, it becomes more attractive in case a highly redundant formulation of the equations of motion is used. The use of Euler parameters for the parametrization of rotational motions simplifies the system matrices such that programming explicit formulas for \mathbf{A} and \mathbf{B} becomes the most efficient strategy. It should also be noted, that the Jacobian matrices $(\mathbf{M}\dot{\mathbf{v}})_q$, \mathbf{f}_q , \mathbf{f}_v and $(\mathbf{C}_q^T \boldsymbol{\lambda})_q$ may be required already for the simulation of the multibody system if an implicit integration scheme such as, e. g., the HHT-algorithm [2, 4] is applied. Hence, an efficient computation of these matrices is crucial also for solving the equations of motion in forward direction.

Equations (21)–(23) must be solved backwards in the physical time $t \in [0, T]$ choosing a proper time step size depending on the numerical stiffness of the problem. For the integration scheme backwards in time, it is advantageous to introduce a new time coordinate τ running from $\tau = 0$ to $\tau = T$ as well.

The transformation

$$\tau = T - t, \quad \tau \in [0, T], \quad \frac{d}{dt} = \frac{d}{d\tau} \frac{d\tau}{dt} = -\frac{d}{d\tau}$$

converts the adjoint equations in Eqs. (21-23) into

$$\begin{aligned} \frac{d\mathbf{p}}{d\tau} &= -h_q^T(\tau) - \mathbf{A}(\tau)\mathbf{w} - \mathbf{C}_q^T(\tau)\boldsymbol{\mu} \\ \frac{d}{d\tau} (\mathbf{M}(\tau)\mathbf{w}) &= -h_v^T(\tau) + \mathbf{p} + \mathbf{B}(\tau)\mathbf{w} + \frac{d}{d\tau} (h_v(\tau))^T \\ \mathbf{0} &= \mathbf{C}_q(\tau)\mathbf{w} \end{aligned} \quad (41)$$

in which $\mathbf{M}(\tau)$, $\mathbf{A}(\tau)$, $\mathbf{B}(\tau)$, $\mathbf{C}_q(\tau)$, $h_q(\tau)$ and $h_v(\tau)$ have to be computed for $\mathbf{q}(\tau) = \mathbf{q}(T - t)$, $\mathbf{v}(\tau) = \mathbf{v}(T - t)$ and $\boldsymbol{\lambda}(\tau) = \boldsymbol{\lambda}(T - t)$. The boundary condition for \mathbf{p} and \mathbf{w} is given by Eqs. (39) and (40) at time $\tau = 0$.

4 Conclusions

The proposed method shows the embedding of the adjoint method in multibody system dynamics and its potential for parameter identification and optimal control on account of its linear structure. The enhanced adjoint method includes the error of accelerations in the optimization functional and therefore delivers an elegant and efficient way to incorporate inverse dynamics for parameter identification in flexible multibody system applications arising from modern engineering problems including acceleration sensors.

Acknowledgements

K. Nachbagauer acknowledges support from the Austrian Science Fund (FWF): T733-N30.

REFERENCES

- [1] M.B. Giles, N.A. Pierce. An Introduction to the Adjoint Approach to Design, *Journal of Flow, Turbulence and Combustion*, Vol.65, 393-415, 2000.
- [2] H.M. Hilbert, T.J.R. Hughes, R.L. Taylor. Improved numerical dissipation for time integration algorithms in structural dynamics, *Earthquake Engineering and Structural Dynamics*, Vol.5, 283-292, 1977.
- [3] C.W. Gear, G. Gupta, B. Leimkuhler. Automatic integration of the Euler-Lagrange equations with constraints, *J. Comp. Appl. Math.*, Vol.12/13, 77-90, 1985.
- [4] D. Negrut, R. Rampalli, G. Ottarsson, A. Sajdak. On the Use of the HHT Method in the Context of Index 3 Differential Algebraic Equations of Multibody Dynamics, *Int. J. Numer. Meth. Engng.*, 1-24, 2000.
- [5] K. Nachbagauer, S. Oberpeilsteiner, K. Sherif, W. Steiner. The Use of the Adjoint Method for Solving Typical Optimization Problems in Multibody Dynamics, *Journal for Computational and Nonlinear Dynamics*, doi: 10.1115/1.4028417, 2014.
- [6] W. Steiner and S. Reichl. The optimal control approach to dynamical inverse problems, *Journal of Dynamical Systems, Measurement and Control*, Vol.134 (2), doi: 10.1115/1.4005365, 2012.
- [7] P. Eberhard. Adjoint Variable Method for Sensitivity Analysis of Multibody Systems Interpreted as a Continuous, Hybrid Form of Automatic Differentiation, *Proc. of the 2nd Int. Workshop on Computational Differentiation*, Santa Fe. Philadelphia: SIAM, 319-328, 1996.
- [8] D. Bestle, P. Eberhard. Analyzing and Optimizing Multibody Systems, *Mech. Struc. Mach.*, Vol.20, 67-92, 1992.
- [9] A. E. Bryson, Y. C. Ho. *Applied Optimal Control*, Hemisphere, Washington, DC, 1975.
- [10] J.P. Lions. *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, New York, 1971.
- [11] J.-Y. Ding, Z.-K. Pan, L.-Q. Chen. Parameter identification of multibody systems based on second order sensitivity analysis, *International Journal of Non-Linear Mechanics*, Vol.47, 1105-1110, 2012.
- [12] E.J. Haug, R.A. Wehage, N.K. Mani. Design sensitivity analysis of large-scaled constrained dynamic mechanical systems, *Trans ASME*, Vol.106, 156-162, 1984.
- [13] A. Jameson. *Aerodynamic Shape Optimization Using the Adjoint Method*, VKI Lecture Series on Aerodynamic Drag Prediction and Reduction, von Karman Institute of Fluid Dynamics, Rhode St Genese, 3-7, 2003.
- [14] C.L. Bottasso, A. Croce, L. Ghezzi, P. Faure. On the Solution of Inverse Dynamics and Trajectory Optimization Problems for Multibody Systems, *Multibody System Dynamics*, Vol.11, 1-22, 2004.

- [15] A.S. Schaffer. On the adjoint formulation of design sensitivity analysis of multibody dynamics, Dissertation, University of Iowa, <http://ir.uiowa.edu/etd/93>, 2005.
- [16] Y. Cao, S. Li, L. Petzold, R. Serban. Adjoint sensitivity analysis for differential-algebraic equations: the adjoint DAE system and its numerical solution, *SIAM Journal on Scientific Computing*, Vol.24 (3), 1076-108, 2003.
- [17] L. Petzold, S. Li, Y. Cao, R. Serban. Sensitivity analysis for differential-algebraic equations and partial differential equations, *Computers and Chemical Engineering*, Vol.30, 1553-1559, 2006.
- [18] E. Bertolazzi, F. Biral, M. Da Lio. Symbolic-numeric indirect method for solving optimal control problems for large multibody systems, *Multibody System Dynamics*, Vol.13, 233-252, 2005.
- [19] A. Held, R. Seifried. Gradient-Based Optimization of Flexible Multibody Systems Using the Absolute Nodal Coordinate Formulation, *Proceedings of the ECCOMAS Thematic Conference Multibody Dynamics 2013*, Zagreb, Croatia, 1-4 July 2013.
- [20] J.-Y. Ding, Z.-K. Pan, L.-Q. Chen. Second order adjoint sensitivity analysis of multibody systems described by differential-algebraic equations, *Multibody System Dynamics*, Vol.18, 599-617, doi:10.1007/s11044-007-9080-4, 2007.
- [21] T. Pi, Y. Zhang, L. Chen. First order sensitivity analysis of flexible multibody systems using the absolute nodal coordinate formulation, *Multibody System Dynamics*, Vol.27 (2), doi: 10.1007/s11044-011-9269-4, 153-171, 2012.
- [22] A.A. Shabana. Definition of the slopes and the finite element absolute nodal coordinate formulation, *Multibody System Dynamics*, Vol.1 (3), 339-348, 1997.
- [23] C.P. Vyasarayani, T. Uchida, J. McPhee. Nonlinear parameter identification in multibody systems using homotopy continuation, *J. Comput. Nonlinear Dyn.*, Vol.7 (1), 011012, 2012.