Adjoint Sensitivity Analysis of Multibody System Equations in State-Space Representation obtained by QR Decomposition

Alexander Held[1], **R. Seifried**[1]

[1] Institute of Mechanics and Ocean Engineering
Hamburg University of Technology
Eißendorfer Straße 41, 21073 Hamburg, Germany
{alexander.held,robert.seifried}@tuhh.de

ABSTRACT

The paper is on applying the adjoint variable method in the sensitivity analysis of multibody systems, which are formulated in state-space representation using a QR decomposition. The Jacobian of the state equations, which is necessary to develop the adjoint dynamics, is derived. Moreover, an analytical way to determine the derivatives of the QR matrices with respect to the redundant position variables is presented, which avoids costly finite differences or differentiation of the direct QR algorithm. The developed method is tested by a spring torsional pendulum found in previous contributions to the subject. The obtained gradient is more precise compared to numerical differentiation using forward differences. Also, providing the analytical Jacobian of the state equations turns out to be beneficial regarding both the gradient accuracy and computational times.

Keywords: multibody systems, adjoint sensitivity analysis, state-space respresentation, QR decomposition

1 Introduction

Gradient determination is often an important step in the analysis and optimization of rigid and flexible multibody systems, which are schematically shown in Fig. 1. If the number of design variables is high, the adjoint variable method is typically the most efficient approach to sensitivity analysis. However, a set of adjoint differential equations have to be derived and solved first, whose structure depends on the structure of the multibody system equations. The focus in this work is on multibody systems, which are initially formulated in implicit differential-algebraic form as

$$
\begin{aligned}
\mathbf{x} \in \mathbb{R}^h \quad & \text{design variables,}\\
\mathbf{z}_\mathrm{I}(t,\mathbf{x}), \mathbf{z}_\mathrm{II}(t,\mathbf{x}) \in \mathbb{R}^r \quad & \text{redundant position and velocity variables,}\\
\boldsymbol{\lambda}(t,\mathbf{x}) \in \mathbb{R}^{n_\mathrm{c}} \quad & \text{Lagrange multipliers,}\\
\boldsymbol{\phi}^0(t^0,\mathbf{z}_\mathrm{I}^0,\mathbf{x}) = \mathbf{0} \quad & \text{initial conditions (position level),}\\
\dot{\boldsymbol{\phi}}^0(t^0,\mathbf{z}_\mathrm{I}^0,\mathbf{z}_\mathrm{II}^0,\mathbf{x}) = \mathbf{0} \quad & \text{initial conditions (velocity level),}\\
\dot{\mathbf{z}}_\mathrm{I} - \mathbf{Z}(\mathbf{z}_\mathrm{I})\mathbf{z}_\mathrm{II} = \mathbf{0} \quad & \text{kinematic relation,}\\
\mathbf{M}(\mathbf{z}_\mathrm{I},\mathbf{x})\dot{\mathbf{z}} - \mathbf{f}(t,\mathbf{z}_\mathrm{I},\mathbf{z}_\mathrm{II},\mathbf{x}) - \mathbf{C}^\mathrm{T}(\mathbf{z}_\mathrm{I},\mathbf{x})\boldsymbol{\lambda} = \mathbf{0} \quad & \text{kinetic equations,}\\
\mathbf{c}(t,\mathbf{z}_\mathrm{I},\mathbf{x}) = \mathbf{0} \quad & \text{constraint equations (position level).}
\end{aligned}
\tag{1}
$$

The system equations (1) include the initial conditions for positions and velocities, the kinematic relation with the kinematic matrix $\mathbf{Z}$, and the kinetic equations with the global mass matrix $\mathbf{M}$, and the right-hand-side vector $\mathbf{f}$ comprising the generalized inertia forces, elastic forces, and the applied loads. Kinematic constraints $\mathbf{c}$ are enforced by the Lagrange multiplier method, augmenting the system equations by the constraint equations $\mathbf{c}$ and the kinetic equations by the reactions $\mathbf{C}^\mathrm{T}\boldsymbol{\lambda}$. The Jacobian of the constraints $\mathbf{C}$ thereby contains the information on the direction of the reactions.

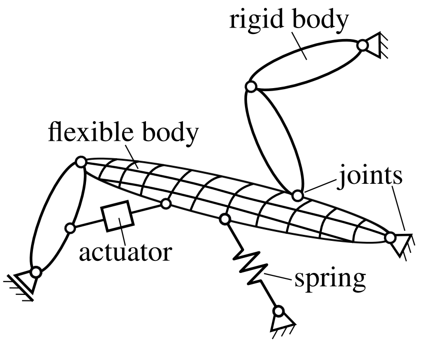


Figure 1: Schematic flexible multibody system

Considering the kinematic constraint at the position level leads to differential-algebraic equations of index-3, which are hard to solve numerically. Therefore, the index is often reduced by differentiating the kinematic constraints twice

$$\dot{\mathbf{c}}(t,\mathbf{z}_{\mathrm{I}},\mathbf{z}_{\mathrm{II}},\mathbf{x}) = \underbrace{\frac{\partial \mathbf{c}}{\partial \mathbf{z}_{\mathrm{I}}}}_{\mathbf{C}}\mathbf{Z}\,\mathbf{z}_{\mathrm{II}} + \underbrace{\frac{\partial \mathbf{c}}{\partial t}}_{\mathbf{c}_{\mathrm{t}}} = \mathbf{0}, \tag{2a}$$

$$\ddot{\mathbf{c}}(t,\mathbf{z}_{\mathrm{I}},\mathbf{z}_{\mathrm{II}},\mathbf{x}) = \mathbf{C}\dot{\mathbf{z}}_{\mathrm{II}} + \underbrace{\dot{\mathbf{C}}\mathbf{z}_{\mathrm{II}} + \frac{\partial^2 \mathbf{c}}{\partial t^2}}_{\mathbf{c}_{\mathrm{tt}}} = \mathbf{0}, \tag{2b}$$

and considering the constraints at the acceleration level in the time integration.

For the adjoint sensitivity analysis of rigid and flexible multibody systems in differential-algebraic form (1), interested readers are referred to, for example, [2, 7]. This work, however, is on the adjoint variable method for multibody formulations, which are transformed from the differential-algebraic form to a representation as ordinary differential equations for the numerical solution. In contrast to previous papers on the usage of the adjoint method for multibody systems in state-space representation, such as [1, 6], this contribution focuses on the application of the adjoint variable method to constrained multibody system equations, formulated using a QR decomposition. In particular, the (analytical) derivatives required for the formulation of the adjoint equations are presented.

The remainder of this paper is structured as follows. First, the multibody system equations (1) are formulated in state-space representation by applying a QR decomposition. Then, the adjoint problem is formulated and the necessary derivatives of the system equations with respect to the position and velocity variables are given. Different ways to determine the derivatives of the Q and R matrices are presented. Finally, the different ways are compared regarding their accuracy and computational efficiency in the adjoint sensitivity analysis using a simple example from the literature.

2 Multibody Systems in State-Space Representation

To avoid the solution of a differential-algebraic equation, the system equations (1) can be transformed into a set of $f = r - n_{\mathrm{c}}$ ordinary differential equations (ODEs) by, for instance, a QR decomposition as proposed in [4]. In the following, the properties of the full QR decomposition and its application to constrained mechanical systems are briefly described.

Every matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ $(m \geq n)$ can be expressed as

$$\mathbf{A} = \mathbf{Q}\mathbf{R} \tag{3}$$

by the product of an orthogonal matrix $\mathbf{Q}$ and an upper triangular matrix $\mathbf{R}$, see, for instance, [3]. If $\mathbf{Q}$ and $\mathbf{R}$ are of dimensions $(m\times m)$ and $(m\times n)$, the decomposition is called full QR decomposition. Accordingly, the transposed of the constraint Jacobian $\mathbf{C}^{\mathsf{T}} \in \mathbb{R}^{r\times n_{\mathrm{c}}}$ can be expressed as

$$\mathbf{C}^{\mathsf{T}} = \mathbf{Q}\mathbf{R} = \begin{bmatrix}\mathbf{Q}_1 & \mathbf{Q}_2\end{bmatrix}\begin{bmatrix}\mathbf{R}_1 \\ \mathbf{0}\end{bmatrix}, \quad \mathbf{Q}_1 \in \mathbb{R}^{r\times n_{\mathrm{c}}}, \mathbf{Q}_2 \in \mathbb{R}^{r\times f}, \mathbf{R}_1 \in \mathbb{R}^{n_{\mathrm{c}}\times n_{\mathrm{c}}}, \tag{4}$$

whereby the submatrices $\mathbf{Q}_1$ and $\mathbf{Q}_2$ represent the constrained and free motion directions of the multibody system. Since $\mathbf{Q}$ is orthogonal, it holds for the submatrices

$$\mathbf{Q}_1^{\mathsf{T}}\mathbf{Q}_1 = \mathbf{E}, \quad \mathbf{Q}_1^{\mathsf{T}}\mathbf{Q}_2 = \mathbf{0}, \quad \mathbf{Q}_2^{\mathsf{T}}\mathbf{Q}_2 = \mathbf{E}. \tag{5}$$

Moreover, since both $\mathbf{Q}_1$ and $\mathbf{C}^{\mathsf{T}} = \mathbf{Q}_1\mathbf{R}_1$ span the constrained motion directions, the product

$$\mathbf{C}\mathbf{Q}_2 = \mathbf{0} \tag{6}$$

is zero, too.

The matrices $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are used to partition the redundant velocities $\mathbf{z}_{\mathrm{II}}$ and accelerations $\dot{\mathbf{z}}_{\mathrm{II}}$ as

$$\mathbf{z}_{\mathrm{II}} = \mathbf{Q}_2\mathbf{z} + \mathbf{Q}_1\bar{\mathbf{z}}, \tag{7a}$$
$$\dot{\mathbf{z}}_{\mathrm{II}} = \mathbf{Q}_2\mathbf{a} + \mathbf{Q}_1\bar{\mathbf{a}}, \tag{7b}$$

into independent coordinates $\mathbf{z}, \mathbf{a} \in \mathbb{R}^f$ and dependent coordinates $\bar{\mathbf{z}}, \bar{\mathbf{a}} \in \mathbb{R}^{n_{\mathrm{c}}}$, respectively. Then the n_{c} dependent coordinates are substituted with the independent ones utilizing the constraint equations. Therefore, the partitioned redundant coordinates from Eq. (7) are incorporated into Eq. (2) and the equations are rearranged for the dependent coordinates as

$$\bar{\mathbf{z}} = -\left(\mathbf{C}\mathbf{Q}_1\right)^{-1}\left(\mathbf{C}\mathbf{Q}_2\mathbf{z} + \mathbf{c}_{\mathrm{t}}\right), \tag{8a}$$
$$\bar{\mathbf{a}} = -\left(\mathbf{C}\mathbf{Q}_1\right)^{-1}\left(\mathbf{C}\mathbf{Q}_2\mathbf{a} + \mathbf{c}_{\mathrm{tt}}\right). \tag{8b}$$

With the Eqs. (4) and (6), Eq. (8) simplifies to

$$\bar{\mathbf{z}} = -\mathbf{R}_1^{-\mathsf{T}}\mathbf{c}_{\mathrm{t}}, \tag{9a}$$
$$\bar{\mathbf{a}} = -\mathbf{R}_1^{-\mathsf{T}}\mathbf{c}_{\mathrm{tt}}. \tag{9b}$$

Substitute the dependent coordinates in Eq. (7) by (9), the redundant coordinates can be given in terms of the independent coordinates only as

$$\mathbf{z}_{\mathrm{II}} = \mathbf{Q}_2\mathbf{z} - \mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\mathbf{c}_{\mathrm{t}}, \tag{10a}$$
$$\dot{\mathbf{z}}_{\mathrm{II}} = \mathbf{Q}_2\mathbf{a} - \mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\mathbf{c}_{\mathrm{tt}}. \tag{10b}$$

These results are used twofold. On the one hand, Eq. (10a) is plugged into the kinematic relation of the system equations (1) yielding

$$\dot{\mathbf{z}}_{\mathrm{I}} - \mathbf{Z}\left(\mathbf{Q}_2\mathbf{z} - \mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\mathbf{c}_{\mathrm{t}}\right) = \mathbf{0}. \tag{11}$$

On the other hand, Eq. (10b) is used in the derivation of the kinetic equations. Using, for instance, Jourdain's principle, the virtual power of the multibody systems is given by

$$\delta\mathbf{z}_{\mathrm{II}}^{\mathsf{T}}\left\{\mathbf{M}\dot{\mathbf{z}}_{\mathrm{II}} - \mathbf{C}^{\mathsf{T}}\boldsymbol{\lambda} - \mathbf{f}\right\} = 0. \tag{12}$$

To eliminate the reaction forces and formulate a set of f generalized equations of motion, the variation $\delta\mathbf{z}_{\mathrm{II}} = \mathbf{Q}_2\delta\mathbf{z}$ of Eq. (10a), and Eq. (10b) are incorporated into Eq. (12) yielding

$$\delta\mathbf{z}^{\mathsf{T}}\mathbf{Q}_2^{\mathsf{T}}\left\{\mathbf{M}\left(\mathbf{Q}_2\mathbf{a} - \mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\mathbf{c}_{\mathrm{tt}}\right) - \mathbf{C}^{\mathsf{T}}\boldsymbol{\lambda} - \mathbf{f}\right\} = 0, \quad \forall\delta\mathbf{z}. \tag{13}$$

Since $\mathbf{Q}_2^\mathsf{T}\mathbf{C}^\mathsf{T} = \mathbf{0}$, the reaction forces in Eq. (12) vanish and the generalized equations of motion

$$\left(\mathbf{Q}_2^\mathsf{T}\mathbf{M}\mathbf{Q}_2\right)\mathbf{a} = \mathbf{Q}_2^\mathsf{T}\mathbf{M}\mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\mathbf{c}_{\mathrm{tt}} + \mathbf{Q}_2^\mathsf{T}\mathbf{f} \tag{14}$$

remain. Equation (14) can be solved for the independent accelerations $\mathbf{a}$, which are then substituted in Eq. (10b). The resulting equation, together with Eq. (11), represents the system equations in the state-space formulation

$$\begin{bmatrix}\dot{\mathbf{z}}_\mathrm{I}\\ \dot{\mathbf{z}}_\mathrm{II}\end{bmatrix} = \begin{bmatrix}\mathbf{v}(t,\mathbf{z}_\mathrm{I},\mathbf{z}_\mathrm{II},\mathbf{x})\\ \mathbf{w}(t,\mathbf{z}_\mathrm{I},\mathbf{z}_\mathrm{II},\mathbf{x})\end{bmatrix} = \begin{bmatrix}\mathbf{Z}\left(\mathbf{Q}_2\mathbf{Q}_2^\mathsf{T}\mathbf{z}_\mathrm{II} - \mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\mathbf{c}_\mathrm{t}\right)\\ \mathbf{Q}_2\left(\mathbf{Q}_2^\mathsf{T}\mathbf{M}\mathbf{Q}_2\right)^{-1}\mathbf{Q}_2^\mathsf{T}\left(\mathbf{M}\mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\mathbf{c}_{\mathrm{tt}} + \mathbf{f}\right) - \mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\mathbf{c}_{\mathrm{tt}}\end{bmatrix} \tag{15}$$

It is worth noting that the independent velocities $\mathbf{z}$ are substituted in the state equation (15) by $\mathbf{Q}_2^\mathsf{T}\mathbf{z}_\mathrm{II}$. This relation can be found by multiplying Eq. (10a) from the left with $\mathbf{Q}_2^\mathsf{T}$.

3 Adjoint Sensitivity Analysis

In this section, at first, a general integral-type criterion function is introduced, and the adjoint variable method to compute the first derivative of the criterion function is briefly given. For an extended derivation, interested readers are referred to [2, 6, 7]. Then the Jacobian of the state equations is derived in general, and the derivatives of two auxiliary matrices are obtained by the QR decomposition in particular.

3.1 General criterion function and adjoint sensitivity analysis

The performance of the dynamic system (15) shall be assessed with the comparatively simple but general integral criterion function

$$\psi(\mathbf{x}) = G^1(t^1,\mathbf{z}_\mathrm{I}^1,\mathbf{z}_\mathrm{II}^1,\mathbf{x}) + \int_{t^0}^{t^1} F(t,\mathbf{z}_\mathrm{I},\mathbf{z}_\mathrm{II},\dot{\mathbf{z}}_\mathrm{II},\mathbf{x})\,\mathrm{d}t. \tag{16}$$

To determine the gradient $\nabla\psi = \mathrm{d}\psi/\mathrm{d}\mathbf{x}$ with the adjoint variable method, at first, the adjoint variables $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ have to be determined at the final time t^1

$$\boldsymbol{\mu}^1 = \frac{\partial G^1}{\partial \mathbf{z}_\mathrm{I}^1}, \quad \boldsymbol{\nu}^1 = \frac{\partial G^1}{\partial \mathbf{z}_\mathrm{II}^1}. \tag{17}$$

Then, the adjoint differential equations

$$\begin{aligned}\dot{\boldsymbol{\mu}} &= -\left(\frac{\partial \mathbf{v}}{\partial \mathbf{z}_\mathrm{I}}\right)^\mathsf{T}\boldsymbol{\mu} - \left(\frac{\partial \mathbf{w}}{\partial \mathbf{z}_\mathrm{I}}\right)^\mathsf{T}\left(\boldsymbol{\nu} + \frac{\partial F}{\partial \dot{\mathbf{z}}_\mathrm{II}}\right) - \frac{\partial F}{\partial \mathbf{z}_\mathrm{I}},\\ \dot{\boldsymbol{\nu}} &= -\left(\frac{\partial \mathbf{v}}{\partial \mathbf{z}_\mathrm{II}}\right)^\mathsf{T}\boldsymbol{\mu} - \left(\frac{\partial \mathbf{w}}{\partial \mathbf{z}_\mathrm{II}}\right)^\mathsf{T}\left(\boldsymbol{\nu} + \frac{\partial F}{\partial \dot{\mathbf{z}}_\mathrm{II}}\right) - \frac{\partial F}{\partial \mathbf{z}_\mathrm{II}}.\end{aligned} \tag{18}$$

have to be developed and solved by a backward time integration. For the formulation of the adjoint equations (18), the Jacobian of the state equations (15) is required since it contains the derivatives of the kinematic and kinetic function $\mathbf{v}$ and $\mathbf{w}$ with respect to the redundant position and velocity coordinates $\mathbf{z}_\mathrm{I}$ and $\mathbf{z}_\mathrm{II}$. It is worth noting that this also holds for the discrete adjoint method described, for instance, in [5].

With the solution of the adjoint equations at hand, the adjoint variables, which enforce the initial conditions at position and velocity level $\boldsymbol{\zeta}^0$ and $\boldsymbol{\eta}^0$ are determined by

$$\begin{aligned}\left(\frac{\partial \dot{\boldsymbol{\phi}}^0}{\partial \mathbf{z}_\mathrm{II}^0}\right)^\mathsf{T}\boldsymbol{\eta}^0 &= \boldsymbol{\nu}^0,\\ \left(\frac{\partial \boldsymbol{\phi}^0}{\partial \mathbf{z}_\mathrm{I}^0}\right)^\mathsf{T}\boldsymbol{\zeta}^0 &= \boldsymbol{\mu}^0 - \left(\frac{\partial \dot{\boldsymbol{\phi}}^0}{\partial \mathbf{z}_\mathrm{I}^0}\right)^\mathsf{T}\boldsymbol{\eta}^0.\end{aligned} \tag{19}$$

The sought gradient is finally calculated by evaluating the equation

$$\nabla\psi = \frac{\partial G^1}{\partial \mathbf{x}} - \left(\frac{\partial \boldsymbol{\phi}^0}{\partial \mathbf{x}}\right)^{\mathsf{T}} \boldsymbol{\zeta}^0 - \left(\frac{\partial \dot{\boldsymbol{\phi}}^0}{\partial \mathbf{x}}\right)^{\mathsf{T}} \boldsymbol{\eta}^0 + \int\limits_{t^0}^{t^1} \left[\frac{\partial F}{\partial \mathbf{x}} + \left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}}\right)^{\mathsf{T}} + \left(\frac{\partial \mathbf{w}}{\partial \mathbf{x}}\right)^{\mathsf{T}} \left(\boldsymbol{\nu} + \frac{\partial F}{\partial \dot{\mathbf{z}}_{\mathrm{II}}}\right)\right] \mathrm{d}t, \tag{20}$$

whereby the derivatives of the initial conditions $\boldsymbol{\phi}^0$ and $\dot{\boldsymbol{\phi}}^0$, of the elements G^1 and F of the criterion function, and of the state equation with respect to the design variables $\mathbf{x}$ are required.

3.2 Jacobian of state equations

For the formulation and solution of the adjoint equations (18), among others, the Jacobian

$$\mathbf{J}(t, \mathbf{z}_{\mathrm{I}}, \mathbf{z}_{\mathrm{II}}, \mathbf{x}) = \begin{bmatrix} \dfrac{\mathrm{d}\mathbf{v}}{\mathrm{d}\mathbf{z}_{\mathrm{I}}} & \dfrac{\mathrm{d}\mathbf{v}}{\mathrm{d}\mathbf{z}_{\mathrm{II}}} \\ \dfrac{\mathrm{d}\mathbf{w}}{\mathrm{d}\mathbf{z}_{\mathrm{I}}} & \dfrac{\mathrm{d}\mathbf{w}}{\mathrm{d}\mathbf{z}_{\mathrm{II}}} \end{bmatrix} \tag{21}$$

of the state equation (15) is required. The derivatives with respect to the velocities $\mathrm{d}\mathbf{v}/\mathrm{d}\mathbf{z}_{\mathrm{II}}$ and $\mathrm{d}\mathbf{w}/\mathrm{d}\mathbf{z}_{\mathrm{II}}$ can be determined comparatively simply. In contrast, providing the derivatives of the kinematics with respect to the position variables

$$\begin{aligned} \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}z_{\mathrm{I},k}} = \frac{\mathrm{d}\mathbf{Z}}{\mathrm{d}z_{\mathrm{I},k}} \left(\mathbf{Q}_2\mathbf{Q}_2^{\mathsf{T}}\mathbf{z}_{\mathrm{II}} - \mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\mathbf{c}_{\mathrm{t}}\right) + \mathbf{Z}\left(\frac{\mathrm{d}\mathbf{Q}_2}{\mathrm{d}z_{\mathrm{I},k}}\mathbf{Q}_2^{\mathsf{T}}\mathbf{z}_{\mathrm{II}} + \mathbf{Q}_2\frac{\mathrm{d}\mathbf{Q}_2^{\mathsf{T}}}{\mathrm{d}z_{\mathrm{I},k}}\mathbf{z}_{\mathrm{II}}\right) \\ - \mathbf{Z}\left(\frac{\partial \mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}}{\partial z_{\mathrm{I},k}}\mathbf{c}_{\mathrm{t}} + \mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\frac{\partial \mathbf{c}_{\mathrm{t}}}{\partial z_{\mathrm{I},k}}\right) \end{aligned} \tag{22}$$

and of the kinetics with respect to the position variables

$$\begin{aligned} \frac{\partial \mathbf{w}}{\partial z_{\mathrm{I},k}} &= \frac{\partial \mathbf{Q}_2}{\partial z_{\mathrm{I},k}}\left(\mathbf{Q}_2^{\mathsf{T}}\mathbf{M}\mathbf{Q}_2\right)^{-1}\mathbf{Q}_2^{\mathsf{T}}\left(\mathbf{M}\mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\mathbf{c}_{\mathrm{tt}} + \mathbf{f}\right) + \mathbf{Q}_2\frac{\partial\left(\mathbf{Q}_2^{\mathsf{T}}\mathbf{M}\mathbf{Q}_2\right)^{-1}}{\partial z_{\mathrm{I},k}}\mathbf{Q}_2^{\mathsf{T}}\left(\mathbf{M}\mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\mathbf{c}_{\mathrm{tt}} + \mathbf{f}\right) \\ &\quad + \mathbf{Q}_2\left(\mathbf{Q}_2^{\mathsf{T}}\mathbf{M}\mathbf{Q}_2\right)^{-1}\frac{\partial \mathbf{Q}_2^{\mathsf{T}}}{\partial z_{\mathrm{I},k}}\left(\mathbf{M}\mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\mathbf{c}_{\mathrm{tt}} + \mathbf{f}\right) + \mathbf{Q}_2\left(\mathbf{Q}_2^{\mathsf{T}}\mathbf{M}\mathbf{Q}_2\right)^{-1}\mathbf{Q}_2^{\mathsf{T}}\frac{\partial\left(\mathbf{M}\mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\mathbf{c}_{\mathrm{tt}} + \mathbf{f}\right)}{\partial z_{\mathrm{I},k}} \\ &\quad - \frac{\partial \mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}}{\partial z_{\mathrm{I},k}}\mathbf{c}_{\mathrm{tt}} - \mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\frac{\partial \mathbf{c}_{\mathrm{tt}}}{\partial z_{\mathrm{I},k}}, \\ \text{with} \quad & \frac{\partial\left(\mathbf{Q}_2^{\mathsf{T}}\mathbf{M}\mathbf{Q}_2\right)^{-1}}{\partial z_{\mathrm{I},k}} = -\left(\mathbf{Q}_2^{\mathsf{T}}\mathbf{M}\mathbf{Q}_2\right)^{-1}\frac{\partial\left(\mathbf{Q}_2^{\mathsf{T}}\mathbf{M}\mathbf{Q}_2\right)}{\partial z_{\mathrm{I},k}}\left(\mathbf{Q}_2^{\mathsf{T}}\mathbf{M}\mathbf{Q}_2\right)^{-1} \\ \text{and} \quad & \frac{\partial\left(\mathbf{M}\mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\mathbf{c}_{\mathrm{tt}} + \mathbf{f}\right)}{\partial z_{\mathrm{I},k}} = \frac{\partial \mathbf{M}}{\partial z_{\mathrm{I},k}}\mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\mathbf{c}_{\mathrm{tt}} + \mathbf{M}\frac{\partial \mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}}{\partial z_{\mathrm{I},k}}\mathbf{c}_{\mathrm{tt}} + \mathbf{M}\mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\frac{\partial \mathbf{c}_{\mathrm{tt}}}{\partial z_{\mathrm{I},k}} + \frac{\partial \mathbf{f}}{\partial z_{\mathrm{I},k}} \end{aligned} \tag{23}$$

is noticeably more difficult. This is because the matrices $\mathbf{Q}_1$, $\mathbf{Q}_2$, and $\mathbf{R}_1$ depend on the position coordinates.

In summary, it can be stated that for Eq. (22) and (23) the derivatives of the system equations

$$\frac{\partial \mathbf{M}}{\partial z_{\mathrm{I},k}}, \frac{\partial \mathbf{f}}{\partial z_{\mathrm{I},k}}, \frac{\partial \mathbf{c}_{\mathrm{t}}}{\partial z_{\mathrm{I},k}}, \frac{\partial \mathbf{c}_{\mathrm{tt}}}{\partial z_{\mathrm{I},k}} \tag{24}$$

are required on the one hand and the derivatives of the auxiliary matrices

$$\frac{\partial \mathbf{Q}_2}{\partial z_{\mathrm{I},k}}, \frac{\partial \mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}}{\partial z_{\mathrm{I},k}} \tag{25}$$

obtained using the QR decomposition on the other hand.

3.3 Derivatives of QR Decomposition

The derivatives of $\mathbf{Q}_2$ and $\mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}$ with respect to position variables can be determined by numerical differentiation, by direct differentiation of the QR decomposition algorithm, or by utilizing the properties of the QR decomposition in the differentiation. The latter is most interesting since it promises to be the most efficient and accurate approach and will be discussed in the following.

To determine $\partial \mathbf{Q}_2/\partial z_{\mathrm{I}k}$, Eq. (6) is differentiated with respect to the k-th element of the position vector $\mathbf{z}_\mathrm{I}$ yielding

$$\frac{\partial \mathbf{C}}{\partial z_{\mathrm{I}k}}\mathbf{Q}_2 + \mathbf{C}\frac{\partial \mathbf{Q}_2}{\partial z_{\mathrm{I}k}} = \mathbf{0}. \tag{26}$$

Since the constraint Jacobian $\mathbf{C}$ is usually rank-deficient, it is not possible to find its inverse and solve Eq. (26) for the derivatives of $\mathbf{Q}_2$. Instead, the first term in Eq. (26) is augmented by

$$\mathbf{C}\mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}} = \mathbf{I} \in \mathbb{R}^{n_c \times n_c}, \tag{27}$$

which can be found from Eq. (4), yielding

$$\mathbf{C}\mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\frac{\partial \mathbf{C}}{\partial z_{\mathrm{I}k}}\mathbf{Q}_2 + \mathbf{C}\frac{\partial \mathbf{Q}_2}{\partial z_{\mathrm{I}k}} = \mathbf{C}\left(\mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\frac{\partial \mathbf{C}}{\partial z_{\mathrm{I}k}}\mathbf{Q}_2 + \frac{\partial \mathbf{Q}_2}{\partial z_{\mathrm{I}k}}\right) = \mathbf{0} \tag{28}$$

By factoring out $\mathbf{C}$, an equation for calculating the sought derivatives of $\mathbf{Q}_2$ is obtained

$$\frac{\partial \mathbf{Q}_2}{\partial z_{\mathrm{I}k}} = -\mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\frac{\partial \mathbf{C}}{\partial z_{\mathrm{I}k}}\mathbf{Q}_2, \tag{29}$$

in which, next to the results of the QR decomposition, only the derivative of the Jacobian with respect to the position variable $z_{\mathrm{I}k}$ is required.

The derivatives $\partial \mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}/\partial z_{\mathrm{I}k}$ can be determined similarly. The starting point is Eq. (27), which is differentiated with respect to $z_{\mathrm{I}k}$

$$\frac{\partial \mathbf{C}}{\partial z_{\mathrm{I}k}}\mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}} + \mathbf{C}\frac{\partial \mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}}{\partial z_{\mathrm{I}k}} = \mathbf{0}. \tag{30}$$

Equation (30) represents an underdetermined system of linear equations for the sought derivatives $\partial \mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}/\partial z_{\mathrm{I}k}$ with f missing equations. They can be found exploiting again that $\mathbf{Q}_2$ is the null space of $\mathbf{Q}_1$. Multiplying the basic equation of the QR decomposition in the formulation

$$\mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\mathbf{C} = \mathbf{Q}_1\mathbf{Q}_1^{\mathsf{T}} \tag{31}$$

from the left with $\mathbf{Q}_2^{\mathsf{T}}$ yields the zero term

$$\mathbf{Q}_2^{\mathsf{T}}\mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\mathbf{C} = \mathbf{Q}_2^{\mathsf{T}}\mathbf{Q}_1\mathbf{Q}_1^{\mathsf{T}} = \mathbf{0}. \tag{32}$$

However, differentiating Eq. (32) with respect to $z_{\mathrm{I}k}$ gives

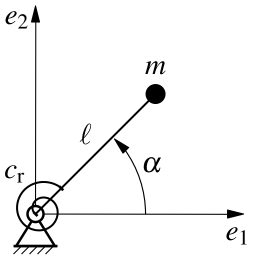
$$\frac{\partial \mathbf{Q}_2^{\mathsf{T}}}{\partial z_{\mathrm{I}k}}\mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\mathbf{C} + \mathbf{Q}_2^{\mathsf{T}}\frac{\partial \mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}}{\partial z_{\mathrm{I}k}}\mathbf{C} + \underbrace{\mathbf{Q}_2^{\mathsf{T}}\mathbf{Q}_1}_{\mathbf{0}}\mathbf{R}_1^{-\mathsf{T}}\frac{\partial \mathbf{C}}{\partial z_{\mathrm{I}k}} = \mathbf{0}, \tag{33}$$

whereby the last term vanishes. Factoring out matrix $\mathbf{C}$ yields

$$\left(\frac{\partial \mathbf{Q}_2^{\mathsf{T}}}{\partial z_{\mathrm{I}k}}\mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}} + \mathbf{Q}_2^{\mathsf{T}}\frac{\partial \mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}}{\partial z_{\mathrm{I}k}}\right)\mathbf{C} = \mathbf{0} \tag{34}$$

and the missing equations to compute $\partial\left(\mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}\right)/\partial z_{\mathrm{I}k}$ are found as

$$\mathbf{Q}_2^{\mathsf{T}}\frac{\partial \mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}}{\partial z_{\mathrm{I}k}} = -\frac{\partial \mathbf{Q}_2^{\mathsf{T}}}{\partial z_{\mathrm{I}k}}\mathbf{Q}_1\mathbf{R}_1^{-\mathsf{T}}. \tag{35}$$



parameter	value
mass m	$2\,\mathrm{kg}$
length ℓ	$2\,\mathrm{m}$
torsional stiffness c_r	$100\,\mathrm{Nm}$
initial deflection α_0	0.5
final time t^1	$10\,\mathrm{s}$

Figure 2: Application example torsional spring pendulum

Thus, combining Eq. (30) and (35), the system of linear equations

$$\begin{bmatrix} \mathbf{C} \\ \mathbf{Q}_2^\mathsf{T} \end{bmatrix} \frac{\partial \mathbf{Q}_1 \mathbf{R}_1^{-\mathsf{T}}}{\partial z_{\mathrm{I}k}} = \begin{bmatrix} -\dfrac{\partial \mathbf{C}}{\partial z_{\mathrm{I}k}} \mathbf{Q}_1 \mathbf{R}_1^{-\mathsf{T}} \\ -\dfrac{\partial \mathbf{Q}_2^\mathsf{T}}{\partial z_{\mathrm{I}k}} \mathbf{Q}_1 \mathbf{R}_1^{-\mathsf{T}} \end{bmatrix} \tag{36}$$

can be formulated, to efficiently determine the sought derivatives $\partial \left(\mathbf{Q}_1 \mathbf{R}_1^{-\mathsf{T}}\right)/\partial z_{\mathrm{I}k}$. Equation (36) must be solved for r redundant position variables, where the coefficient matrix, however, remains constant.

4 Application example

The results of the analytical Jacobian and their influence on the accuracy and computational efficiency in the adjoint sensitivity analysis are checked and investigated. As a testing and application example, the same torsional spring pendulum as in [2] is used. It is described in the following. Thereafter, the results for the Jacobian and gradient evaluation are presented.

4.1 Torsional spring pendulum

In Fig. 2, the system and its parameters are shown. If the position of the point mass m is described by the redundant position coordinates $\mathbf{z}_\mathrm{I} = [z_{\mathrm{I},1}, z_{\mathrm{I},2}]^\mathsf{T}$, the following implicit differential-algebraic system equations can be formulated

$$\begin{aligned}
\boldsymbol{\phi}^0(\mathbf{z}_\mathrm{I}^0, \mathbf{x}) = \begin{bmatrix} z_{\mathrm{I},1}^0 \\ z_{\mathrm{I},2}^0 \end{bmatrix} - \begin{bmatrix} \ell\cos(\alpha_0) \\ \ell\sin(\alpha_0) \end{bmatrix} &= \mathbf{0}, \quad \text{(initial positions)} \\
\dot{\boldsymbol{\phi}}^0(\mathbf{z}_\mathrm{II}^0) = \begin{bmatrix} z_{\mathrm{II},1}^0 \\ z_{\mathrm{II},2}^0 \end{bmatrix} &= \mathbf{0}, \quad \text{(initial velocities)} \\
\begin{bmatrix} \dot{z}_{\mathrm{I},1}^2 \\ \dot{z}_{\mathrm{I},2}^2 \end{bmatrix} - \begin{bmatrix} z_{\mathrm{II},1}^2 \\ z_{\mathrm{II},2}^2 \end{bmatrix} &= \mathbf{0}, \quad \text{(kinematics)} \\
\begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix} \begin{bmatrix} \dot{z}_{\mathrm{II},1}^2 \\ \dot{z}_{\mathrm{II},2}^2 \end{bmatrix} - \begin{bmatrix} 2\,z_{\mathrm{I},1} \\ 2\,z_{\mathrm{I},2} \end{bmatrix} \lambda - \begin{bmatrix} \dfrac{c_\mathrm{r}\alpha}{\ell^2} z_{\mathrm{I},2} \\ -\dfrac{c_\mathrm{r}\alpha}{\ell^2} z_{\mathrm{I},1} \end{bmatrix} &= \mathbf{0}, \quad \text{(kinetics)} \\
\mathbf{c}(\mathbf{z}_\mathrm{I}, \mathbf{x}) = z_{\mathrm{I},1}^2 + z_{\mathrm{I},2}^2 - \ell^2 &= \mathbf{0}, \quad \text{(constraints)}
\end{aligned} \tag{37}$$

with the auxiliary variable $\alpha(\mathbf{z}_\mathrm{I}) = \mathrm{atan2}(z_{\mathrm{I},2}, z_{\mathrm{I},1})$. As in [2], the pendulum length ℓ is chosen as the only design variable $\mathbf{x} = [\ell]$. The system performance is also assessed by the position of mass m in e_2-direction at the final time $t^1 = 10\,\mathrm{s}$. Thus, the objective function is defined as $\psi = G^1 = z_{\mathrm{I},2}^1$.

Even though the number of design variables is minimal and, thus, the adjoint variable method is unlikely the most efficient approach to sensitivity analysis, this application example possesses a major advantage: the analytical solution to the problem and, hence, for the objective function

$$\psi_\mathrm{ana} = \ell \sin\left[\alpha_0 \cos\left(\sqrt{\frac{c_\mathrm{r}}{m\ell^2}}\, t^1\right)\right] \tag{38}$$

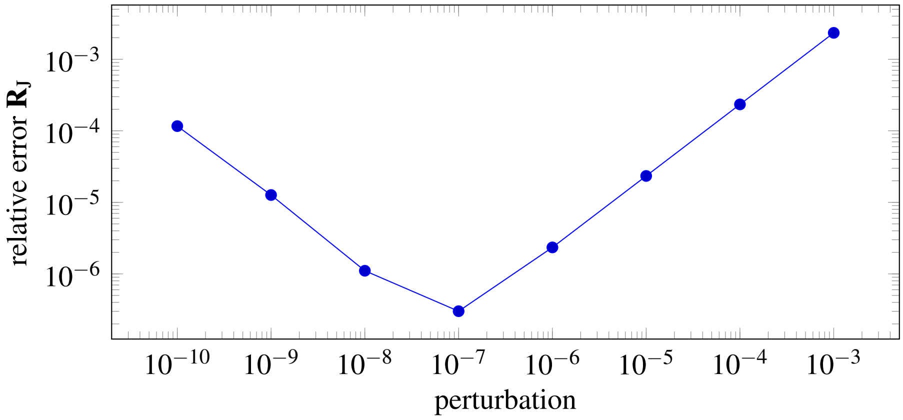


Figure 3: Relative error of Jacobian for torsional spring pendulum at time t^1

and its gradient

$$\nabla\psi_{\mathrm{ana}} = \sin\left[\alpha_0 \cos\left(\sqrt{\frac{c_\mathrm{r}}{m\ell^2}}t^1\right)\right] + \frac{\alpha_0 t^1}{\ell}\sqrt{\frac{c_\mathrm{r}}{m}}\cos\left[\alpha_0\cos\left(\sqrt{\frac{c_\mathrm{r}}{m\ell^2}}t^1\right)\right]\sin\left(\sqrt{\frac{c_\mathrm{r}}{m\ell^2}}t^1\right) \tag{39}$$

are available.

4.2 Jacobian and gradient results

At first, the analytical derivatives for the Jacobian (21) are checked. Therefore, $\mathbf{J}$ is evaluated at time t^1 analytically as shown in the sections 3.2 and 3.3 and numerically using simple forward finite differences with a uniform perturbation. The relative error between the analytical Jacobian $\mathbf{J}_{\mathrm{ana}}$ and the approximated Jacobian $\mathbf{J}_{\mathrm{fd}}$ is computed as

$$\mathbf{R}_\mathrm{J} = \sqrt{\sum_i\sum_j\left(\frac{J_{ij}^{\mathrm{ana}} - J_{ij}^{\mathrm{fd}}}{J_{ij}^{\mathrm{ana}}}\right)}. \tag{40}$$

The results obtained for different perturbations of the state variables in the computation of the finite differences are given in Fig. 3. In general, there is a good agreement with the analytical results, whereby the lowest relative error is obtained using a perturbation of 10^{-7}. More importantly, the ratio of the evaluation times for the approximated and analytical Jacobian $t_{\mathrm{fd}}/t_{\mathrm{ana}}$ is about 2.2. Thus, for this application example, the computation time reduces even though the accuracy improves.

After the Jacobian has been validated, the gradient of objective $\nabla\psi$ is determined using the adjoint method described in section 3.1. Therefore, at first, the system equations are solved forward in time (MATLAB solver ODE45, AbsTol = 10^{-14}, RelTol = 10^{-10}). Then the adjoint equations (21) are solved backward in time (MATLAB solver ODE45, AbsTol = 10^{-13}, RelTol = 10^{-9}). Finally, the gradient equation (20) is numerically evaluated (MATLAB function INTEGRAL, AbsTol = 10^{-12}).

The results of the adjoint variable method are compared with two other methods. On the one hand, a second adjoint approach is implemented for the system equations of the torsional spring pendulum formulated in minimal coordinates. On the other hand, the gradient is determined with different perturbations of the pendulum length ℓ.

The relative and absolute gradient errors for both adjoint approaches are summarized in Tab. 1, whereas the relative and absolute gradient errors using forward finite differences are shown in Fig. (4). It can be seen that the gradient errors obtained with the two adjoint approaches are in the same order of magnitude and in good agreement. In contrast, the errors obtained with finite differences are about two orders of magnitude higher.

system equations obtained by	QR decomposition	minimal coordinates
absolute error $\lvert\nabla\psi - \nabla\psi_{\text{ana}}\rvert$	$5.1{\cdot}10^{-09}$	$9.9{\cdot}10^{-09}$
relative error $\lvert(\nabla\psi - \nabla\psi_{\text{ana}})/\nabla\psi_{\text{ana}}\rvert$	$4.2{\cdot}10^{-10}$	$8.1{\cdot}10^{-10}$

Table 1: Gradient errors using the adjoint variable method

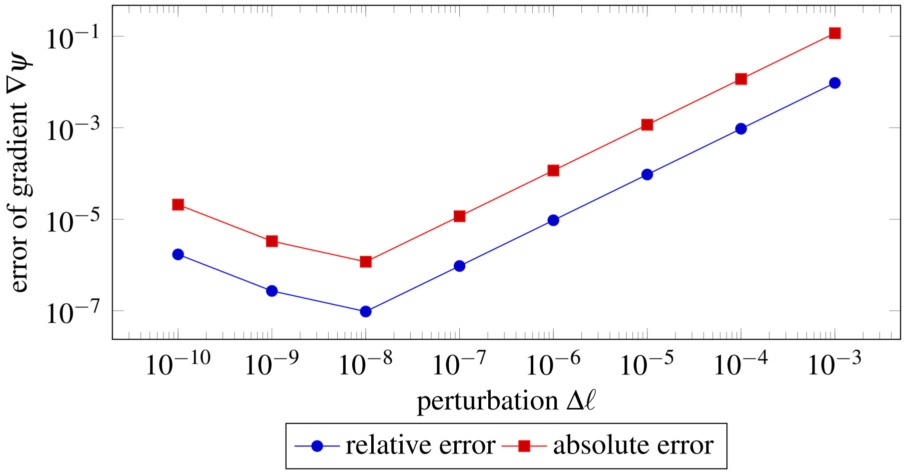


Figure 4: Gradient error using the finite difference method

Finally, it is investigate whether the effort of the analytical calculation of the Jacobian of the state equation including the derivatives of the $\mathbf{Q}_2$ and $\mathbf{Q}_1\mathbf{R}_1$ is justified in terms of computational effort and accuracy of the gradient. To this end, the analytical Jacobian is replaced by a numerical one, which is obtained using simple forward differences (uniform state perturbations of 10^{-7}). With the numerical Jacobian, the absolute and relative gradient error increase to $5.4{\cdot}10^{-08}$ and $4.4{\cdot}10^{-09}$, respectively. At the same time, the total time of the sensitivity analysis increases by a factor of about 1.6 due to a significantly higher number of failed steps in the numerical solution of the adjoint dynamics. As implied by Fig. 3, the results are worsening for smaller and bigger perturbations of the state variables.

5 Summary and Conclusion

In the analysis and optimization of rigid and flexible multibody systems, gradient information is often required, and thus their efficient and accurate computation is of great interest. The paper presents how the adjoint variable method can be applied to multibody systems, whose system equations are initially set up in differential-algebraic form but solved in minimal coordinates obtained using a QR decomposition.

In contrast to previous works, the dependency of the projection matrices on the redundant position variables must be taken into account when developing the adjoint differential equations. In this work, it is shown that the costly usage of the finite difference method and the derivation of the direction QR decomposition algorithm can be avoided. Instead, it is sufficient to compute the derivatives of two auxiliary matrices solving only systems of linear equations.

The developed method is tested by means of a spring torsional pendulum described in the literature. The results show the high accuracy of the adjoint variable method compared to numerical differentiation using forward finite differences on the one hand. On the other hand, it turns out that it is beneficial to provide the analytical Jacobian, including the analytical derivatives of the auxiliary matrices compared to finite differencing, regarding both the accuracy of the gradient and the computational efficiency of the method.

REFERENCES

[1] D. Bestle, P. Eberhard. Analyzing and optimizing multibody systems. Mechanics of structures and machines, 20(1), 67-92, 1992.

[2] Bestle, D. J. Seybold. Sensitivity analysis of constrained multibody systems. Archive of Applied Mechanics 62(3), 181–190, 1992.

[3] Karpfinger, C. The QR-Decomposition of a Matrix. In: Calculus and Linear Algebra in Recipes. Springer, Berlin, Heidelberg, 2022.

[4] S. S. Kim, M. J. Vanderploeg. QR decomposition for state space representation of constrained mechanical dynamic systems, Journal of Mechanisms, Transmissions, and Automation in Design, 1986.

[5] Lauß, T., S. Oberpeilsteiner, W. Steiner, K. Nachbagauer. The discrete adjoint method for parameter identification in multibody system dynamics. Multibody Syst Dyn 42, 397–410, 2018.

[6] D. Dopico, Y. Zhu, A. Sandu, C. Sandu. Direct and adjoint sensitivity analysis of ordinary differential equation multibody formulations. Journal of Computational and Nonlinear Dynamics, 10(1), 2015.

[7] A. Azari Nejat, A. Moghadasi, A. Held. Adjoint sensitivity analysis of flexible multibody systems in differential-algebraic form. Computers & Structures, 2020.