Proceedings of the ASME 2013 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference
IDETC/CIE 2013
August 4-7, 2013, Portland, Oregon, USA

DETC2013-12696

DIRECT AND ADJOINT SENSITIVITY ANALYSIS OF MULTIBODY SYSTEMS USING MAGGI'S EQUATIONS

Daniel Dopico
Yitao Zhu

Advanced Vehicle Dynamics Laboratory
and Computer Science Laboratory
Department of Mechanical Engineering
and Department of Computer Science
Virginia Tech
Blacksburg, VA 24061, USA
Email: yitao7@vt.edu Email: ddopico@vt.edu

Adrian Sandu *

Computer Science Laboratory
Department of Computer Science
Virginia Tech
Blacksburg, VA 24061, USA
Email: asandu7@vt.edu

Corina Sandu

Advanced Vehicle Dynamics Laboratory
Department of Mechanical Engineering
Virginia Tech
Blacksburg, VA 24061, USA
Email: csandu@vt.edu

ABSTRACT

The importance of the sensitivity analysis of multibody systems for several applications is well known, concretely design optimization based on the dynamics of multibody systems usually requires the sensitivity analysis of the equations of motion. A broad range of methods for the dynamics of multibody systems include the state space formulations based on Maggis equations, nullspace methods or coordinate partitioning. Dynamic sensitivities, when needed, are often calculated by means of finite differences but, depending of the number of parameters involved, this procedure can be very demanding in terms of CPU time and the accuracy obtained can be very poor in many cases. In this paper, several ways to perform the sensitivity analysis are explored and analytical expressions for the direct and adjoint sensitivity analysis of multibody systems are presented, all of them based on Maggi's formulations. Moreover, two different approaches to the adjoint sensitivity analysis of multibody systems are presented.

Although particularized to one formulation, the general expressions provided in the paper, are intended to be easily generalized and applied to any other formulation that can be expressed as an ODE-like system of equations, including penalty formulations.

Besides, to check the validity and correctness of the proposed equations, the solutions of all the methods proposed are compared: 1) between them, 2) with the third party code FATODE and 3) with the numerical solution using real and complex perturbations.

Finally, all the techniques proposed are applied to the dynamical optimization of a multibody system.

INTRODUCTION

Multibody dynamics has become an essential tool for mechanical systems analysis and design. The evolution during the last decades makes possible not only to think about the analysis of mechanical systems, but also to develop tools that can help to improve the design of them. Thus, one interesting application of the state-of-the-art multibody models, is the optimal design for which sensitivity analysis is essential.

In general, the multibody dynamics equations, constitute an index-3 differential algebraic system of equations (DAE) that it is not is usually directly solved because of the numerical difficulties involved [1, 2]. Some of the most advanced families of formulations used nowadays are based on some ideas presented in the eighties and nineties. One of this families comprise state

*Address all correspondence to this author.

space formulations based on Maggis equations, nullspace methods or coordinate partitioning introduced in [3–5].

A lot of attention is being paid recently to the sensitivity analysis of multibody systems, for different applications. There are two different approaches to obtain the sensitivity equations of dynamical systems: the direct sensitivity approach and the adjoint sensitivity approach [6–8]. Different approaches were developed for different formulations of the equations of motion [9–13]. In this paper the two approaches are developed for one state space formulation based on the projection Matrix R [5] or Maggi's equations and the results are applied to the optimization of multibody systems.

To validate the expressions proposed for the sensitivities is essential, since small changes or omissions in the derivatives can lead to completely different results for the sensitivities obtained. For this reason, the validity of the theoretical results introduced in the paper to calculate the sensitivities is checked comparing the direct and the adjoint approaches, comparing with numerical results and comparing with the third party library for the sensitivity analysis of ODE systems, FATODE [14].

DESCRIPTION OF THE MULTIBODY FORMULATION

The formulation used here was described in [5] with the name of projection matrix R formulation.

The starting point for the formulation is the virtual power principle in dependent coordinates.

$$\delta \mathbf{q}^{*\mathrm{T}} (\mathbf{M}\ddot{\mathbf{q}} - \mathbf{Q}) = \mathbf{0} \tag{1}$$

Where $\delta \mathbf{q}^*$ constitutes a set of n dependent virtual displacements and the rest of the terms are the same as described before.

In addition to the equation (1), the actual positions, velocities and accelerations, and the virtual displacements, have to fulfill the constraint equations and their derivatives, that is the following kinematic relations.

$$\mathbf{\Phi} = \mathbf{0} \tag{2}$$
$$\mathbf{\Phi_q}\dot{\mathbf{q}} = -\mathbf{\Phi}_t = \mathbf{b} \tag{3}$$
$$\mathbf{\Phi_q}\ddot{\mathbf{q}} = -\dot{\mathbf{\Phi}}_\mathbf{q}\dot{\mathbf{q}} - \dot{\mathbf{\Phi}}_t = \mathbf{c} \tag{4}$$
$$\mathbf{\Phi_q}\delta\mathbf{q}^* = \mathbf{0} \tag{5}$$

Where, again, $\mathbf{\Phi}$ is the constraints vector, $\mathbf{\Phi_q} = \dfrac{\partial \mathbf{\Phi}}{\partial \mathbf{q}}$ is the Jacobian matrix of the constraints vector and $\mathbf{\Phi}_t = \dfrac{\partial \mathbf{\Phi}}{\partial t}$.

Let $\mathbf{R} \in \mathbb{R}^{n\mathrm{x}(n-m)}$ be a matrix of size $n\mathrm{x}(n-m)$ being n the number of variables and m the number of independent constraints of the system, including the reonomous ones, and let $\mathbf{S} \in \mathbb{R}^{n\mathrm{x}m}$ be another matrix of size $n\mathrm{x}m$. It is possible to write the dependent velocities and accelerations of the system in terms of the selected degrees of freedom by means of this matrices, without worrying about how to calculate them yet.

$$\dot{\mathbf{q}} = \mathbf{R}\dot{\mathbf{z}} + \mathbf{S}\mathbf{b} \tag{6}$$
$$\ddot{\mathbf{q}} = \mathbf{R}\ddot{\mathbf{z}} + \mathbf{S}\mathbf{c} \tag{7}$$

In a similar way the virtual dependent displacements can be expressed in function of the virtual degrees of freedom, or independent virtual displacements.

$$\delta \mathbf{q}^* = \mathbf{R}\delta\mathbf{z}^* \tag{8}$$

Using (7) and (8) in (1), and taking into account that the virtual velocities $\mathbf{z}$ are independent.

$$\left(\mathbf{R}^\mathrm{T}\mathbf{M}\mathbf{R}\right)\ddot{\mathbf{z}} = \mathbf{R}^\mathrm{T}(\mathbf{Q} - \mathbf{M}\mathbf{S}\mathbf{c}) \tag{9}$$

Or more compactly:

$$\bar{\mathbf{M}}(\mathbf{z})\ddot{\mathbf{z}} = \bar{\mathbf{Q}}(t, \mathbf{z}, \dot{\mathbf{z}}) \tag{10}$$
$$\bar{\mathbf{M}}(\mathbf{z}) = \mathbf{R}^\mathrm{T}\mathbf{M}\mathbf{R} \tag{11}$$
$$\bar{\mathbf{Q}}(t, \mathbf{z}, \dot{\mathbf{z}}) = \mathbf{R}^\mathrm{T}(\mathbf{Q} - \mathbf{M}\mathbf{S}\mathbf{c}) \tag{12}$$

Equations (10)-(12) constitute a second order state-space ODE.

One might be wondering now, how to calculate the matrices $\mathbf{R}$ and $\mathbf{S}$. There are different ways to calculate them. One possibility is to write the following kinematic systems of equations.

$$\begin{bmatrix} \mathbf{\Phi_q} \\ \mathbf{B} \end{bmatrix} \dot{\mathbf{q}} = \begin{bmatrix} \mathbf{b} \\ \dot{\mathbf{z}} \end{bmatrix} \tag{13}$$
$$\begin{bmatrix} \mathbf{\Phi_q} \\ \mathbf{B} \end{bmatrix} \ddot{\mathbf{q}} = \begin{bmatrix} \mathbf{c} \\ \ddot{\mathbf{z}} \end{bmatrix} \tag{14}$$

Where $\mathbf{B}$ is a very simple matrix composed of zeros and ones, thus choosing the degrees of freedom $\mathbf{z}$ as a subset of the dependent coordinates $\mathbf{q}$. From (13) and (14), it can be followed that.

$$\dot{\mathbf{q}} = \begin{bmatrix} \mathbf{\Phi_q} \\ \mathbf{B} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{b} \\ \dot{\mathbf{z}} \end{bmatrix} = \mathbf{R}\dot{\mathbf{z}} + \mathbf{S}\mathbf{b} \tag{15}$$
$$\ddot{\mathbf{q}} = \begin{bmatrix} \mathbf{\Phi_q} \\ \mathbf{B} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{c} \\ \ddot{\mathbf{z}} \end{bmatrix} = \mathbf{R}\ddot{\mathbf{z}} + \mathbf{S}\mathbf{c} \tag{16}$$

So the matrices $\mathbf{R}$ and $\mathbf{S}$ are not but the columns of the inverse of the leading matrix in (13) or (14), and the equations (15) and (16) are not but the relations (6) and (7) introduced before without explaining how to obtain them. Nevertheless the explicit calculation of the matrix $\mathbf{S}$ is rarely needed, so from the computational point of view, the calculation of the inverse in (15) and (16) can be avoided most of times. Nevertheless, in this case as will be shown next the matrix $\mathbf{S}$ will play an important role on the sensitivities calculation.

As indicated before, instead of the inverse, most of times the columns of the matrix $\mathbf{R}$, the terms $\mathbf{Sb}$, $\mathbf{Sc}$ and the columns of the matrix $\dot{\mathbf{R}}$ are obtained by means of appropriate velocity and acceleration analysis.

$$\underbrace{\mathbf{R}_j}_{n\text{x}1} = \dot{\mathbf{q}}\big|_{\boldsymbol{\Phi}_t=\mathbf{0},\dot{z}_j=1,\dot{z}_i=0\,(i\neq j)} = \begin{bmatrix}\boldsymbol{\Phi}_\mathbf{q}\\ \mathbf{B}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{0}\\ \dot{\mathbf{z}}\end{bmatrix}_{\dot{z}_j=1,\dot{z}_i=0\,(i\neq j)} \tag{17}$$

$$\mathbf{Sb} = \dot{\mathbf{q}}\big|_{\dot{\mathbf{z}}=\mathbf{0}} = \begin{bmatrix}\boldsymbol{\Phi}_\mathbf{q}\\ \mathbf{B}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{b}\\ \mathbf{0}\end{bmatrix} \tag{18}$$

$$\mathbf{Sc} = \ddot{\mathbf{q}}\big|_{\ddot{\mathbf{z}}=\mathbf{0}} = \begin{bmatrix}\boldsymbol{\Phi}_\mathbf{q}\\ \mathbf{B}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{c}\\ \mathbf{0}\end{bmatrix} \tag{19}$$

$$\begin{aligned}\underbrace{\dot{\mathbf{R}}_j}_{n\text{x}1} &= \ddot{\mathbf{q}}\big|_{\boldsymbol{\Phi}_t=\mathbf{0},\dot{\boldsymbol{\Phi}}_t=\mathbf{0},\ddot{\mathbf{z}}=\mathbf{0},\dot{z}_j=1,\dot{z}_i=0\,(i\neq j)} = \\ &\begin{bmatrix}\boldsymbol{\Phi}_\mathbf{q}\\ \mathbf{B}\end{bmatrix}^{-1}\begin{bmatrix}-\dot{\boldsymbol{\Phi}}_\mathbf{q}\dot{\mathbf{q}}\\ \mathbf{0}\end{bmatrix}_{\boldsymbol{\Phi}_t=\mathbf{0},\dot{\mathbf{q}}=\mathbf{R}_j}\end{aligned} \tag{20}$$

In practice, the velocity and acceleration analyses of equations (17), (19) and (20) can be implemented in a more efficient way than the inversion of the matrix proposed. A better option is a common velocity (or acceleration) analysis removing the columns of the Jacobian matrix corresponding to the degrees of freedom, or solving by least squares in case that redundant constraints are present. Moreover, the Jacobian matrix is always factorized from previous steps evaluating the kinematics of the system and only a forward and back substitution is needed for each analysis.

Finally, let's suppose that the multibody system is described by the equations of motion (10) and dependent on some design parameters $\boldsymbol{\rho}\in\mathbb{R}^p$ (typically masses, lengths, or other parameters related to forces chosen by the engineer). Then the equations of motion (10) become.

$$\bar{\mathbf{M}}(\mathbf{z},\boldsymbol{\rho})\,\ddot{\mathbf{z}} = \bar{\mathbf{Q}}(t,\mathbf{z},\dot{\mathbf{z}},\boldsymbol{\rho}) \tag{21}$$

Note that the mass matrix and generalized forces vector in (21) are now dependent of a parameter set $\boldsymbol{\rho}$ that contains the design variables of the system and therefore $\mathbf{z}=\mathbf{z}(t,\boldsymbol{\rho})$, $\dot{\mathbf{z}}=\dot{\mathbf{z}}(t,\boldsymbol{\rho})$, $\ddot{\mathbf{z}}=\ddot{\mathbf{z}}(t,\boldsymbol{\rho})$, being t the time variable.

DIRECT SENSITIVITY ANALYSIS

The direct sensitivity approach involves obtaining the sensitivity of a cost function defined in terms of some states and design parameters of the system. The objective functions considered here will have the following form.

$$\psi = \int_{t_0}^{t_F} g(\mathbf{q},\dot{\mathbf{q}},\boldsymbol{\rho})\,\mathrm{dt} \tag{22}$$

Note that the cost function (22) is supposed to depend explicitly on the dependent states $\mathbf{q}$ and their derivatives, instead of depending on the independent ones $\mathbf{z}$ and their derivatives.

The gradient of the cost function (22) can be obtained by the following expression.

$$\nabla_{\boldsymbol{\rho}}\psi^{\mathrm{T}} = \frac{\mathrm{d}\psi}{\mathrm{d}\boldsymbol{\rho}} = \int_{t_0}^{t_F}\left(\left(\frac{\partial g}{\partial\mathbf{q}}\frac{\partial\mathbf{q}}{\partial\mathbf{z}} + \frac{\partial g}{\partial\dot{\mathbf{q}}}\frac{\partial\dot{\mathbf{q}}}{\partial\mathbf{z}}\right)\frac{\partial\mathbf{z}}{\partial\boldsymbol{\rho}} + \frac{\partial g}{\partial\dot{\mathbf{q}}}\frac{\partial\dot{\mathbf{q}}}{\partial\dot{\mathbf{z}}}\frac{\partial\dot{\mathbf{z}}}{\partial\boldsymbol{\rho}} + \frac{\partial g}{\partial\boldsymbol{\rho}}\right)\mathrm{dt} \tag{23}$$

Using the common notation of a sub-index to express partial derivatives and commuting the temporal and parameter derivatives:

$$\frac{\partial\mathbf{z}}{\partial\boldsymbol{\rho}} = \mathbf{z}_{\boldsymbol{\rho}} \tag{24}$$

$$\frac{\partial\dot{\mathbf{z}}}{\partial\boldsymbol{\rho}} = \frac{d}{dt}\frac{\partial\mathbf{z}}{\partial\boldsymbol{\rho}} = \dot{\mathbf{z}}_{\boldsymbol{\rho}} \tag{25}$$

From (8) and (6) the following expressions can be derived:

$$\frac{\partial\mathbf{q}}{\partial\mathbf{z}} = \frac{\partial\dot{\mathbf{q}}}{\partial\dot{\mathbf{z}}} = \mathbf{R} \tag{26}$$

$$\frac{\partial\dot{\mathbf{q}}}{\partial\mathbf{z}} = \frac{\partial\mathbf{R}}{\partial\mathbf{z}}\dot{\mathbf{z}} + \frac{\partial\mathbf{Sb}}{\partial\mathbf{z}} \tag{27}$$

For (27):

$$\frac{\partial\mathbf{R}}{\partial\mathbf{z}}\dot{\mathbf{z}} = \left[\frac{\partial\mathbf{R}}{\partial z_1}\dot{\mathbf{z}}\ \ldots\ \frac{\partial\mathbf{R}}{\partial z_i}\dot{\mathbf{z}}\ \ldots\ \frac{\partial\mathbf{R}}{\partial z_{n-m}}\dot{\mathbf{z}}\right] \tag{28}$$

$$\underbrace{\frac{\partial\mathbf{R}}{\partial z_i}}_{n\text{x}(n-m)} = \sum_{j=1}^{n}\frac{\partial\mathbf{R}}{\partial q_j}\frac{\partial q_j}{\partial z_i} = \sum_{j=1}^{n}\frac{\partial\mathbf{R}}{\partial q_j}R_{ji} = -\mathbf{S}\sum_{j=1}^{n}\left(\frac{\partial\boldsymbol{\Phi}_\mathbf{q}}{\partial q_j}R_{ji}\right)\mathbf{R} \tag{29}$$

$$\frac{\partial\mathbf{Sb}}{\partial\mathbf{z}} = \frac{\partial\mathbf{Sb}}{\partial\mathbf{q}}\frac{\partial\mathbf{q}}{\partial\mathbf{z}} = \mathbf{S}\left(-\boldsymbol{\Phi}_{\mathbf{qq}}\mathbf{Sb} + \frac{\partial\mathbf{b}}{\partial\mathbf{q}}\right)\mathbf{R} \tag{30}$$

$$\boldsymbol{\Phi}_{\mathbf{qq}}\mathbf{Sb} = \left[\frac{\partial\boldsymbol{\Phi}_\mathbf{q}}{\partial q_1}\mathbf{Sb}\ \ldots\ \frac{\partial\boldsymbol{\Phi}_\mathbf{q}}{\partial q_i}\mathbf{Sb}\ \ldots\ \frac{\partial\boldsymbol{\Phi}_\mathbf{q}}{\partial q_n}\mathbf{Sb}\right] \tag{31}$$

In equations (29) and (30) the following result derived from equations (17) and (18) respectively, was used.

$$\underbrace{\frac{\partial \mathbf{R}}{\partial q_j}}_{n\mathrm{x}(n-m)} = \begin{bmatrix} \mathbf{\Phi_q} \\ \mathbf{B} \end{bmatrix}^{-1} \begin{bmatrix} -\frac{\partial \mathbf{\Phi_q}}{\partial q_j}\mathbf{R} \\ \mathbf{0} \end{bmatrix} = -\mathbf{S}\frac{\partial \mathbf{\Phi_q}}{\partial q_j}\mathbf{R} \tag{32}$$

$$\frac{\partial \mathbf{Sb}}{\partial q_j} = \begin{bmatrix} \mathbf{\Phi_q} \\ \mathbf{B} \end{bmatrix}^{-1} \begin{bmatrix} -\frac{\partial \mathbf{\Phi_q}}{\partial q_j}\mathbf{Sb} + \frac{\partial \mathbf{b}}{\partial q_j} \\ \mathbf{0} \end{bmatrix} = \mathbf{S}\left(-\frac{\partial \mathbf{\Phi_q}}{\partial q_j}\mathbf{Sb} + \frac{\partial \mathbf{b}}{\partial q_j}\right) \tag{33}$$

Then equation (23) becomes:

$$\nabla_{\boldsymbol{\rho}} \psi^{\mathrm{T}} = \frac{\mathrm{d}\psi}{\mathrm{d}\boldsymbol{\rho}} = \int_{t_0}^{t_F} \left(\left(\frac{\partial g}{\partial \mathbf{q}}\mathbf{R} + \frac{\partial g}{\partial \dot{\mathbf{q}}}\left(\frac{\partial \mathbf{R}}{\partial \mathbf{z}}\dot{\mathbf{z}} + \frac{\partial \mathbf{Sb}}{\partial \mathbf{z}}\right)\right)\mathbf{z}_{\boldsymbol{\rho}} + \frac{\partial g}{\partial \dot{\mathbf{q}}}\mathbf{R}\dot{\mathbf{z}}_{\boldsymbol{\rho}} + \frac{\partial g}{\partial \boldsymbol{\rho}} \right) \mathrm{dt} \tag{34}$$

Where the derivatives of function g are known, since the objective function has a known expression and the derivatives $\mathbf{z}_{\boldsymbol{\rho}}$ and $\dot{\mathbf{z}}_{\boldsymbol{\rho}}$ are the sensitivities of the solution of the dynamical equations (21), that need to be obtained differentiating them like follows.

$$\frac{\mathrm{d}\bar{\mathbf{M}}}{\mathrm{d}\boldsymbol{\rho}}\ddot{\mathbf{z}} + \bar{\mathbf{M}}\frac{\partial \ddot{\mathbf{z}}}{\partial \boldsymbol{\rho}} = \frac{\mathrm{d}\bar{\mathbf{Q}}}{\mathrm{d}\boldsymbol{\rho}} \tag{35}$$

Expanding the total derivatives.

$$\frac{\partial \bar{\mathbf{M}}}{\partial \boldsymbol{\rho}}\ddot{\mathbf{z}} + \frac{\partial \bar{\mathbf{M}}}{\partial \mathbf{z}}\ddot{\mathbf{z}}\frac{\partial \mathbf{z}}{\partial \boldsymbol{\rho}} + \bar{\mathbf{M}}\frac{\partial \ddot{\mathbf{z}}}{\partial \boldsymbol{\rho}} = \frac{\partial \bar{\mathbf{Q}}}{\partial \mathbf{z}}\frac{\partial \mathbf{z}}{\partial \boldsymbol{\rho}} + \frac{\partial \bar{\mathbf{Q}}}{\partial \dot{\mathbf{z}}}\frac{\partial \dot{\mathbf{z}}}{\partial \boldsymbol{\rho}} + \frac{\partial \bar{\mathbf{Q}}}{\partial \boldsymbol{\rho}} \tag{36}$$

Finally, defining $()_{\boldsymbol{\rho}} = \frac{\partial ()}{\partial \boldsymbol{\rho}}$, $()_{\mathbf{q}} = \frac{\partial ()}{\partial \mathbf{q}}$ and grouping terms, the following ODE, called Tangent Linear Model (TLM) is obtained:

$$\bar{\mathbf{M}}\ddot{\mathbf{z}}_{\boldsymbol{\rho}} + \bar{\mathbf{C}}\dot{\mathbf{z}}_{\boldsymbol{\rho}} + \left(\bar{\mathbf{K}} + \bar{\mathbf{M}}_{\mathbf{z}}\ddot{\mathbf{z}}\right)\mathbf{z}_{\boldsymbol{\rho}} = \bar{\mathbf{Q}}_{\boldsymbol{\rho}} - \bar{\mathbf{M}}_{\boldsymbol{\rho}}\ddot{\mathbf{z}} \tag{37}$$

$$\mathbf{z}_{\boldsymbol{\rho}}(t_0) = \mathbf{z}_{\boldsymbol{\rho}0} \tag{38}$$

$$\dot{\mathbf{z}}_{\boldsymbol{\rho}}(t_0) = \dot{\mathbf{z}}_{\boldsymbol{\rho}0} \tag{39}$$

In (37), $\bar{\mathbf{K}}$, $\bar{\mathbf{C}}$ and $\bar{\mathbf{Q}}_{\boldsymbol{\rho}}$ are given by expressions (40), (41) and (42) and the terms $\bar{\mathbf{M}}_{\mathbf{z}}\ddot{\mathbf{z}}$ and $\bar{\mathbf{M}}_{\boldsymbol{\rho}}\ddot{\mathbf{z}}$ are derivatives of matrices times vectors, which are matrices obtained by means of expressions (43) and (44). Since the initial conditions (38) and (39) are initial sensitivities of the degrees of freedom of the system, its value can to be decided based on the physical interpretation of them.

$$\bar{\mathbf{K}} = -\frac{\partial \bar{\mathbf{Q}}}{\partial \mathbf{z}} = -\left(\frac{\partial \bar{\mathbf{Q}}}{\partial \mathbf{q}}\frac{\partial \mathbf{q}}{\partial \mathbf{z}} + \frac{\partial \bar{\mathbf{Q}}}{\partial \dot{\mathbf{q}}}\frac{\partial \dot{\mathbf{q}}}{\partial \mathbf{z}}\right) = -\left(\frac{\partial \bar{\mathbf{Q}}}{\partial \mathbf{q}}\mathbf{R} + \frac{\partial \bar{\mathbf{Q}}}{\partial \dot{\mathbf{q}}}\left(\frac{\partial \mathbf{R}}{\partial \mathbf{z}}\dot{\mathbf{z}} + \frac{\partial \mathbf{Sb}}{\partial \mathbf{z}}\right)\right) \tag{40}$$

$$\bar{\mathbf{C}} = -\frac{\partial \bar{\mathbf{Q}}}{\partial \dot{\mathbf{z}}} = -\frac{\partial \bar{\mathbf{Q}}}{\partial \dot{\mathbf{q}}}\frac{\partial \dot{\mathbf{q}}}{\partial \dot{\mathbf{z}}} = -\frac{\partial \bar{\mathbf{Q}}}{\partial \dot{\mathbf{q}}}\mathbf{R} \tag{41}$$

$$\bar{\mathbf{Q}}_{\boldsymbol{\rho}} = \frac{\partial \bar{\mathbf{Q}}}{\partial \boldsymbol{\rho}} = \mathbf{R}^{\mathrm{T}}\left(\mathbf{Q}_{\boldsymbol{\rho}} - \mathbf{M}_{\boldsymbol{\rho}}\mathbf{Sc}\right) \tag{42}$$

$$\bar{\mathbf{M}}_{\mathbf{z}}\ddot{\mathbf{z}} = \frac{\partial \bar{\mathbf{M}}}{\partial \mathbf{z}}\ddot{\mathbf{z}} = \frac{\partial \mathbf{R}^{\mathrm{T}}}{\partial \mathbf{z}}\mathbf{M}\mathbf{R}\ddot{\mathbf{z}} + \mathbf{R}^{\mathrm{T}}\mathbf{M}_{\mathbf{z}}\mathbf{R}\ddot{\mathbf{z}} + \mathbf{R}^{\mathrm{T}}\mathbf{M}\frac{\partial \mathbf{R}}{\partial \mathbf{z}}\ddot{\mathbf{z}} \tag{43}$$

$$\bar{\mathbf{M}}_{\boldsymbol{\rho}}\ddot{\mathbf{z}} = \frac{\partial \bar{\mathbf{M}}}{\partial \boldsymbol{\rho}}\ddot{\mathbf{z}} = \mathbf{R}^{\mathrm{T}}\mathbf{M}_{\boldsymbol{\rho}}\mathbf{R}\ddot{\mathbf{z}} \tag{44}$$

The derivatives $\frac{\partial \bar{\mathbf{Q}}}{\partial \mathbf{q}}$ and $\frac{\partial \bar{\mathbf{Q}}}{\partial \dot{\mathbf{q}}}$ of equations (40) and (41) are calculated as follows.

$$\frac{\partial \bar{\mathbf{Q}}}{\partial \mathbf{q}} = \frac{\partial \mathbf{R}^{\mathrm{T}}}{\partial \mathbf{q}}(\mathbf{Q} - \mathbf{MSc}) + \mathbf{R}^{\mathrm{T}}\frac{\partial (\mathbf{Q} - \mathbf{MSc})}{\partial \mathbf{q}} = \frac{\partial \mathbf{R}^{\mathrm{T}}}{\partial \mathbf{q}}(\mathbf{Q} - \mathbf{MSc}) - \mathbf{R}^{\mathrm{T}}\left(\mathbf{K} + \mathbf{M_q Sc} + \mathbf{M}\frac{\partial \mathbf{Sc}}{\partial \mathbf{q}}\right) \tag{45}$$

$$\frac{\partial \bar{\mathbf{Q}}}{\partial \dot{\mathbf{q}}} = \mathbf{R}^{\mathrm{T}}\frac{\partial (\mathbf{Q} - \mathbf{MSc})}{\partial \dot{\mathbf{q}}} = -\mathbf{R}^{\mathrm{T}}\left(\mathbf{C} + \mathbf{M}\frac{\partial \mathbf{Sc}}{\partial \dot{\mathbf{q}}}\right) \tag{46}$$

Where $\mathbf{K} = -\frac{\partial \mathbf{Q}}{\partial \mathbf{q}}$, $\mathbf{C} = -\frac{\partial \mathbf{Q}}{\partial \dot{\mathbf{q}}}$ and the terms of $\frac{\partial \mathbf{R}^{\mathrm{T}}}{\partial \mathbf{q}}$ are the transpose of (32) and the following relations hold.

$$\mathbf{M_q Sc} = \left[\frac{\partial \mathbf{M}}{\partial q_1}\mathbf{Sc} \ldots \frac{\partial \mathbf{M}}{\partial q_i}\mathbf{Sc} \ldots \frac{\partial \mathbf{M}}{\partial q_n}\mathbf{Sc}\right] \tag{47}$$

$$\frac{\partial \mathbf{Sc}}{\partial \mathbf{q}} = \mathbf{S}\left(-\mathbf{\Phi_{qq}}\mathbf{Sc} + \frac{\partial \mathbf{c}}{\partial \mathbf{q}}\right) \tag{48}$$

$$\frac{\partial \mathbf{Sc}}{\partial \dot{\mathbf{q}}} = \mathbf{S}\frac{\partial \mathbf{c}}{\partial \dot{\mathbf{q}}} \tag{49}$$

$$\frac{\partial \mathbf{c}}{\partial \mathbf{q}} = -\dot{\mathbf{\Phi}}_{\mathbf{q}\dot{\mathbf{q}}}\dot{\mathbf{q}} - \dot{\mathbf{\Phi}}_{t\mathbf{q}} \tag{50}$$

$$\frac{\partial \mathbf{c}}{\partial \dot{\mathbf{q}}} = -\mathbf{\Phi_{qq}}\dot{\mathbf{q}} - \dot{\mathbf{\Phi}}_{\mathbf{q}} - \dot{\mathbf{\Phi}}_{t\dot{\mathbf{q}}} \tag{51}$$

In equations (48) and (51), the terms $\mathbf{\Phi_{qq}Sc}$ and $\mathbf{\Phi_{qq}\dot{q}}$ need to be calculated in the same way as $\mathbf{\Phi_{qq}Sb}$ in (31), only replacing the vector $\mathbf{Sb}$ by $\mathbf{Sc}$ and $\dot{\mathbf{q}}$ respectively. In equation (50) the following expression has to be used.

$$\dot{\mathbf{\Phi}}_{\mathbf{q}\dot{\mathbf{q}}}\dot{\mathbf{q}} = \left[\frac{\partial \dot{\mathbf{\Phi}}_{\mathbf{q}}}{\partial \dot{q}_1}\dot{\mathbf{q}} \ldots \frac{\partial \dot{\mathbf{\Phi}}_{\mathbf{q}}}{\partial \dot{q}_i}\dot{\mathbf{q}} \ldots \frac{\partial \dot{\mathbf{\Phi}}_{\mathbf{q}}}{\partial \dot{q}_n}\dot{\mathbf{q}} \right] \tag{52}$$

Finally, for equations (43) and (44):

$$\mathbf{M_z R\ddot{z}} = \left[\frac{\partial \mathbf{M}}{\partial z_1}\mathbf{R\ddot{z}} \ldots \frac{\partial \mathbf{M}}{\partial z_i}\mathbf{R\ddot{z}} \ldots \frac{\partial \mathbf{M}}{\partial z_{n-m}}\mathbf{R\ddot{z}} \right] \tag{53}$$

$$\underbrace{\frac{\partial \mathbf{M}}{\partial z_i}}_{nxn} = \sum_{j=1}^{n} \frac{\partial \mathbf{M}}{\partial q_j}\frac{\partial q_j}{\partial z_i} = \sum_{j=1}^{n} \frac{\partial \mathbf{M}}{\partial q_j} R_{ji} \tag{54}$$

$$\mathbf{M}_{\boldsymbol{\rho}}\mathbf{R\ddot{z}} = \left[\frac{\partial \mathbf{M}}{\partial \rho_1}\mathbf{R\ddot{z}} \ldots \frac{\partial \mathbf{M}}{\partial \rho_i}\mathbf{R\ddot{z}} \ldots \frac{\partial \mathbf{M}}{\partial \rho_p}\mathbf{R\ddot{z}} \right] \tag{55}$$

ADJOINT SENSITIVITY ANALYSIS

Two approaches to the problem have been developed, depending on how the equations of motion are considered:

1. First approach: equations of motion are written as a first order explicit ODE system.
2. Second approach: equations of motion are written as a first order semi-explicit ODE system.

Equations of motion written as a first order explicit ODE system

The system (21) can be transformed into a first order semi-explicit one, by simply defining a new set of variables by the relation $\dot{\mathbf{z}} = \mathbf{v}$,

$$\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{M}} \end{bmatrix} \begin{bmatrix} \dot{\mathbf{z}} \\ \dot{\mathbf{v}} \end{bmatrix} = \begin{bmatrix} \mathbf{v} \\ \bar{\mathbf{Q}} \end{bmatrix} \tag{56}$$

$$\hat{\mathbf{M}}(\mathbf{y},\boldsymbol{\rho})\dot{\mathbf{y}} = \hat{\mathbf{Q}}(t,\mathbf{y},\boldsymbol{\rho}) \tag{57}$$

In (57) the new vector $\mathbf{y} = \left[\mathbf{z}^{\mathrm{T}} \; \mathbf{v}^{\mathrm{T}} \right]^{\mathrm{T}}$ was defined in order to lead the system from second to first order. Taking the inverse of the leading matrix in (57) the system can be expressed as a first order explicit one.

$$\dot{\mathbf{y}} = \hat{\mathbf{M}}^{-1}(\mathbf{y},\boldsymbol{\rho})\hat{\mathbf{Q}}(t,\mathbf{y},\boldsymbol{\rho}) = \mathbf{f}(t,\mathbf{y},\boldsymbol{\rho}) \tag{58}$$

The cost function (22) becomes in terms of the new states:

$$\psi = \int_{t_0}^{t_F} g(\mathbf{y},\boldsymbol{\rho})\, dt \tag{59}$$

Following the work of [8], let's consider the following Lagrangian, given by the cost function subjected to the equations of motion.

$$L(\boldsymbol{\rho}) = \int_{t_0}^{t_F} g(\mathbf{y},\boldsymbol{\rho})\, dt - \int_{t_0}^{t_F} \boldsymbol{\mu}^{\mathrm{T}}(\dot{\mathbf{y}} - \mathbf{f}(t,\mathbf{y},\boldsymbol{\rho}))\, dt \tag{60}$$

Where $\boldsymbol{\mu}$ is the vector of Lagrange multipliers. Applying variational calculus.

$$\begin{aligned} \delta L = &\int_{t_0}^{t_F} \left(\frac{\partial g}{\partial \mathbf{y}}\delta\mathbf{y} + \frac{\partial g}{\partial \boldsymbol{\rho}}\delta\boldsymbol{\rho} \right) dt \\ &- \int_{t_0}^{t_F} \delta\boldsymbol{\mu}^{\mathrm{T}}(\dot{\mathbf{y}} - \mathbf{f}(t,\mathbf{y},\boldsymbol{\rho}))\, dt \\ &- \int_{t_0}^{t_F} \boldsymbol{\mu}^{\mathrm{T}} \left(\delta\dot{\mathbf{y}} - \frac{\partial \mathbf{f}}{\partial \mathbf{y}}\delta\mathbf{y} - \frac{\partial \mathbf{f}}{\partial \boldsymbol{\rho}}\delta\boldsymbol{\rho} \right) dt \end{aligned} \tag{61}$$

The parenthesis in the central term are the equations of motion, therefore if they are fulfilled in each time step, the term vanishes. For the last term, integration by parts can be applied.

$$\int_{t_0}^{t_F} \boldsymbol{\mu}^{\mathrm{T}}\delta\dot{\mathbf{y}}dt = \boldsymbol{\mu}^{\mathrm{T}}\delta\mathbf{y}\Big|_{t_0}^{t_F} - \int_{t_0}^{t_F} \dot{\boldsymbol{\mu}}^{\mathrm{T}}\delta\mathbf{y}dt \tag{62}$$

Therefore.

$$\begin{aligned} \delta L = &\int_{t_0}^{t_F} \left(\frac{\partial g}{\partial \mathbf{y}} + \boldsymbol{\mu}^{\mathrm{T}}\frac{\partial \mathbf{f}}{\partial \mathbf{y}} + \dot{\boldsymbol{\mu}}^{\mathrm{T}} \right) \delta\mathbf{y}dt \\ &+ \int_{t_0}^{t_F} \left(\frac{\partial g}{\partial \boldsymbol{\rho}} + \boldsymbol{\mu}^{\mathrm{T}}\frac{\partial \mathbf{f}}{\partial \boldsymbol{\rho}} \right) \delta\boldsymbol{\rho}dt - \\ &\boldsymbol{\mu}^{\mathrm{T}}(t_F)\,\delta\mathbf{y}(t_F) + \boldsymbol{\mu}^{\mathrm{T}}(t_0)\,\delta\mathbf{y}(t_0) \end{aligned} \tag{63}$$

In equation (63), $\delta\mathbf{y}(t_0)$ in last term is known and the previous term can be cancelled choosing $\boldsymbol{\mu}(t_F) = \mathbf{0}$. Moreover, to avoid calculating $\delta\mathbf{y}$, the first integral can be canceled by choosing $\boldsymbol{\mu}$ to be the solution of following adjoint ODE system.

$$\dot{\boldsymbol{\mu}} = -\frac{\partial \mathbf{f}}{\partial \mathbf{y}}^{\mathrm{T}}\boldsymbol{\mu} - \frac{\partial g}{\partial \mathbf{y}}^{\mathrm{T}} \tag{64}$$

$$\boldsymbol{\mu}(t_F) = \mathbf{0} \tag{65}$$

Therefore, from equation (63) the gradient of the cost function with respect to parameters can be obtained as

$$\nabla_{\boldsymbol{\rho}} \psi = \frac{\partial \psi}{\partial \boldsymbol{\rho}}^{\mathrm{T}} = \frac{\partial \mathbf{y_0}}{\partial \boldsymbol{\rho}}^{\mathrm{T}} \boldsymbol{\mu}(t_0) + \int_{t_0}^{t_F} \left(\frac{\partial \mathbf{f}}{\partial \boldsymbol{\rho}}^{\mathrm{T}} \boldsymbol{\mu} + \frac{\partial g}{\partial \boldsymbol{\rho}}^{\mathrm{T}} \right) dt \quad (66)$$

In the previous result the identity $\delta \psi = \delta L$ was used, which holds if the equations of motion are satisfied, as can be derived from (60).

In (66) and (64) the derivatives of function g are known, since the objective function has a known expression. To obtain the derivatives of $\mathbf{f}$, expression (57) can be used.

$$\hat{\mathbf{M}} \frac{\partial \mathbf{f}}{\partial \mathbf{y}} + \frac{\partial \hat{\mathbf{M}}}{\partial \mathbf{y}} \mathbf{f} = \frac{\partial \hat{\mathbf{Q}}}{\partial \mathbf{y}} \Rightarrow \frac{\partial \mathbf{f}}{\partial \mathbf{y}} = \hat{\mathbf{M}}^{-1} \left(\frac{\partial \hat{\mathbf{Q}}}{\partial \mathbf{y}} - \frac{\partial \hat{\mathbf{M}}}{\partial \mathbf{y}} \mathbf{f} \right) \quad (67)$$

$$\hat{\mathbf{M}} \frac{\partial \mathbf{f}}{\partial \boldsymbol{\rho}} + \frac{\partial \hat{\mathbf{M}}}{\partial \boldsymbol{\rho}} \mathbf{f} = \frac{\partial \hat{\mathbf{Q}}}{\partial \boldsymbol{\rho}} \Rightarrow \frac{\partial \mathbf{f}}{\partial \boldsymbol{\rho}} = \hat{\mathbf{M}}^{-1} \left(\frac{\partial \hat{\mathbf{Q}}}{\partial \boldsymbol{\rho}} - \frac{\partial \hat{\mathbf{M}}}{\partial \boldsymbol{\rho}} \mathbf{f} \right) \quad (68)$$

The derivatives $\frac{\partial \mathbf{f}}{\partial \mathbf{y}}$ and $\frac{\partial \mathbf{f}}{\partial \boldsymbol{\rho}}$ can be calculated by blocks.

$$\frac{\partial \mathbf{f}}{\partial \mathbf{y}} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{M}}^{-1} \end{bmatrix} \left(\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\bar{\mathbf{K}} & -\bar{\mathbf{C}} \end{bmatrix} - \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \frac{\partial \bar{\mathbf{M}}}{\partial \mathbf{z}} \dot{\mathbf{v}} & \mathbf{0} \end{bmatrix} \right) = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\bar{\mathbf{M}}^{-1} \left(\bar{\mathbf{K}} + \bar{\mathbf{M}}_{\mathbf{z}} \dot{\mathbf{v}} \right) & -\bar{\mathbf{M}}^{-1} \bar{\mathbf{C}} \end{bmatrix} \quad (69)$$

$$\frac{\partial \mathbf{f}}{\partial \boldsymbol{\rho}} = \begin{bmatrix} \mathbf{0} \\ \bar{\mathbf{M}}^{-1} \left(\frac{\partial \bar{\mathbf{Q}}}{\partial \boldsymbol{\rho}} + \frac{\partial \bar{\mathbf{M}}}{\partial \boldsymbol{\rho}} \dot{\mathbf{v}} \right) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \bar{\mathbf{M}}^{-1} \left(\bar{\mathbf{Q}}_{\boldsymbol{\rho}} + \bar{\mathbf{M}}_{\boldsymbol{\rho}} \dot{\mathbf{v}} \right) \end{bmatrix} \quad (70)$$

Taking into account the identity $\mathbf{v} = \dot{\mathbf{z}}$ in (69), the terms $\bar{\mathbf{K}}$, $\bar{\mathbf{C}}$ and $\bar{\mathbf{M}}_{\mathbf{z}} \dot{\mathbf{v}}$ are given by equations (40), (41) and (43) respectively. Identically in (70) the terms $\bar{\mathbf{Q}}_{\boldsymbol{\rho}}$ and $\bar{\mathbf{M}}_{\boldsymbol{\rho}} \dot{\mathbf{v}}$ are given by (42) and (44) respectively.

Equations of motion written as a first order semi-explicit ODE system

Another way to write the adjoint ODE system (64) and the gradient of the cost function (66) can be obtained using the equations of motion (57) instead of writing the Lagrangian in terms of the equations of motion (58). Then the new Lagrangian is like follows.

$$L(\boldsymbol{\rho}) = \int_{t_0}^{t_F} g(\mathbf{y}, \boldsymbol{\rho}) \, dt - \int_{t_0}^{t_F} \boldsymbol{\mu}^{\mathrm{T}} \left(\hat{\mathbf{M}}(\mathbf{y}, \boldsymbol{\rho}) \dot{\mathbf{y}} - \hat{\mathbf{Q}}(t, \mathbf{y}, \boldsymbol{\rho}) \right) dt \quad (71)$$

Applying variational calculus.

$$\delta L = \int_{t_0}^{t_F} \left(\frac{\partial g}{\partial \mathbf{y}} \delta \mathbf{y} + \frac{\partial g}{\partial \boldsymbol{\rho}} \delta \boldsymbol{\rho} \right) dt - \int_{t_0}^{t_F} \delta \boldsymbol{\mu}^{\mathrm{T}} \left(\hat{\mathbf{M}}(\mathbf{y}, \boldsymbol{\rho}) \dot{\mathbf{y}} - \hat{\mathbf{Q}}(t, \mathbf{y}, \boldsymbol{\rho}) \right) dt - \int_{t_0}^{t_F} \boldsymbol{\mu}^{\mathrm{T}} \left(\hat{\mathbf{M}} \delta \dot{\mathbf{y}} + \frac{\partial \hat{\mathbf{M}}}{\partial \mathbf{y}} \dot{\mathbf{y}} \delta \mathbf{y} + \frac{\partial \hat{\mathbf{M}}}{\partial \boldsymbol{\rho}} \dot{\mathbf{y}} \delta \boldsymbol{\rho} - \frac{\partial \hat{\mathbf{Q}}}{\partial \mathbf{y}} \delta \mathbf{y} - \frac{\partial \hat{\mathbf{Q}}}{\partial \boldsymbol{\rho}} \delta \boldsymbol{\rho} \right) dt \quad (72)$$

Again the central term vanishes if the equations of motion are fulfilled at every instant. For the last term, integration by parts can be applied.

$$\int_{t_0}^{t_F} \boldsymbol{\mu}^{\mathrm{T}} \hat{\mathbf{M}} \delta \dot{\mathbf{y}} dt = \boldsymbol{\mu}^{\mathrm{T}} \hat{\mathbf{M}} \delta \mathbf{y} \Big|_{t_0}^{t_F} - \int_{t_0}^{t_F} \left(\dot{\boldsymbol{\mu}}^{\mathrm{T}} \hat{\mathbf{M}} + \boldsymbol{\mu}^{\mathrm{T}} \dot{\hat{\mathbf{M}}} \right) \delta \mathbf{y} dt \quad (73)$$

Therefore,

$$\delta L = \int_{t_0}^{t_F} \left(\frac{\partial g}{\partial \mathbf{y}} - \boldsymbol{\mu}^{\mathrm{T}} \left(\frac{\partial \hat{\mathbf{M}}}{\partial \mathbf{y}} \dot{\mathbf{y}} - \frac{\partial \hat{\mathbf{Q}}}{\partial \mathbf{y}} - \dot{\hat{\mathbf{M}}} \right) + \dot{\boldsymbol{\mu}}^{\mathrm{T}} \hat{\mathbf{M}} \right) \delta \mathbf{y} dt + \int_{t_0}^{t_F} \left(\frac{\partial g}{\partial \boldsymbol{\rho}} - \boldsymbol{\mu}^{\mathrm{T}} \left(\frac{\partial \hat{\mathbf{M}}}{\partial \boldsymbol{\rho}} \dot{\mathbf{y}} - \frac{\partial \hat{\mathbf{Q}}}{\partial \boldsymbol{\rho}} \right) \right) \delta \boldsymbol{\rho} dt - \boldsymbol{\mu}^{\mathrm{T}} \hat{\mathbf{M}} \delta \mathbf{y} \Big|_{t_0}^{t_F} \quad (74)$$

Again, the first integral and the last term at the final time can be canceled if $\boldsymbol{\mu}$ fulfills the following adjoint ODE.

$$\hat{\mathbf{M}}^{\mathrm{T}} \dot{\boldsymbol{\mu}} = \left(\frac{\partial \hat{\mathbf{M}}}{\partial \mathbf{y}} \dot{\mathbf{y}} - \frac{\partial \hat{\mathbf{Q}}}{\partial \mathbf{y}} - \dot{\hat{\mathbf{M}}} \right)^{\mathrm{T}} \boldsymbol{\mu} - \frac{\partial g}{\partial \mathbf{y}}^{\mathrm{T}} \quad (75)$$

$$\boldsymbol{\mu}(t_F) = \mathbf{0} \quad (76)$$

where $\frac{\partial g}{\partial \mathbf{y}}$ is known and

$$\frac{\partial \hat{\mathbf{M}}}{\partial \mathbf{y}} \dot{\mathbf{y}} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{M}}_{\mathbf{z}} \dot{\mathbf{v}} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{M}}_{\mathbf{z}} \ddot{\mathbf{z}} & \mathbf{0} \end{bmatrix} \quad (77)$$

$$\frac{\partial \hat{\mathbf{Q}}}{\partial \mathbf{y}} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\bar{\mathbf{K}} & -\bar{\mathbf{C}} \end{bmatrix} \tag{78}$$

Taking into account that $\bar{\mathbf{M}} = \bar{\mathbf{M}}(\mathbf{z}, \boldsymbol{\rho})$ and the parameters do not vary with time,

$$\dot{\hat{\mathbf{M}}} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \dot{\bar{\mathbf{M}}} \end{bmatrix} \tag{79}$$

$$\dot{\bar{\mathbf{M}}} = \sum_i \frac{\partial \bar{\mathbf{M}}}{\partial z_i} \dot{z}_i \tag{80}$$

From equation (74), after removing the terms made zero in (75) and (76), the gradient of the cost function with respect to parameters can be obtained as

$$\nabla_{\boldsymbol{\rho}} \psi = \frac{\partial \mathbf{y_0}}{\partial \boldsymbol{\rho}}^{\mathrm{T}} \hat{\mathbf{M}}^{\mathrm{T}}(t_0)\, \boldsymbol{\mu}(t_0) + \int_{t_0}^{t_F} \left(\left(\frac{\partial \hat{\mathbf{Q}}}{\partial \boldsymbol{\rho}} - \frac{\partial \hat{\mathbf{M}}}{\partial \boldsymbol{\rho}} \dot{\mathbf{y}} \right)^{\mathrm{T}} \boldsymbol{\mu} + \frac{\partial g}{\partial \boldsymbol{\rho}}^{\mathrm{T}} \right) dt \tag{81}$$

Where.

$$\frac{\partial \hat{\mathbf{Q}}}{\partial \boldsymbol{\rho}} = \begin{bmatrix} \mathbf{0} \\ \bar{\mathbf{Q}}_{\boldsymbol{\rho}} \end{bmatrix} \tag{82}$$

$$\frac{\partial \hat{\mathbf{M}}}{\partial \boldsymbol{\rho}} \dot{\mathbf{y}} = \begin{bmatrix} \mathbf{0} \\ \bar{\mathbf{M}}_{\boldsymbol{\rho}} \dot{\mathbf{v}} \end{bmatrix} \tag{83}$$

being $\bar{\mathbf{Q}}_{\boldsymbol{\rho}}$ and $\bar{\mathbf{M}}_{\boldsymbol{\rho}} \dot{\mathbf{v}}$ given by (42) and (44) respectively.

VALIDATION OF THE COMPUTED SENSITIVITIES

Several approaches were used to make sure that the formulations proposed compute the sensitivities correctly and that all the derivatives proposed are correct. It is important to remark, that any mistake, even small, in the derivatives involved in the direct or adjoint approaches can lead to completely different results in the sensitivities computed.

The validation proposed and implemented here included several strategies:

1. Compare the results of direct and adjoint sensitivity approaches. They should be equal within the truncation error.
2. Compare the results of different formulations of the equations of motion. They should be almost equal if the formulations represent accurately the motion behavior. The alternative formulation of the equations of motion, employed in this work to doble-check the results was the penalty formulation [5,15].
3. Compute the sensitivities using a third party code: FATODE [14].
4. Use *real* finite differences to approximate whole sensitivities or individual derivatives. This approach can be very inaccurate or even completely useless.
5. Use *complex* finite differences to approximate whole sensitivities or individual derivatives. This approach is much more reliable than the previous one, but more complex to implement.

Compute the sensitivities using FATODE

The code computes the direct dynamics and sensitivities using adjoint techniques. Since the derivatives are provided by the user, the comparison can only detect errors in the algorithms but not in the derivatives.

The forward, adjoint, and tangent linear integration of ODEs (FATODE) is a library which provides explicit/implicit Runge-Kutta and Rosenbrock integrators for nonstiff and stiff ODEs. The forward model can solve ODE systems. The tangent linear model and the discrete adjoint model are used by the integrators in FATODE to perform sensitivity analysis. To use the integrators in FATODE for the forward simulations, two basic functions, $\mathbf{f}(t, \mathbf{y}, \boldsymbol{\rho})$ and $\frac{\partial \mathbf{f}}{\partial \mathbf{y}}$, are required. Besides, the objective function ψ, which is defined in (84), and several additional functions are also required for sensitivity analysis. In addition, for sensitivity analysis, an additional function is required by the integrators to initialize the adjoint variable $\boldsymbol{\lambda}_s$ and $\boldsymbol{\mu}_s$ before the backward simulation. The functions to define and their connections with the equations of this work, are the following.

ψ: the objective function, which is defined as follows:

$$\psi = r(\mathbf{y}(t_F), \boldsymbol{\rho}) + \int_{t_0}^{t_F} g(\mathbf{y}, \boldsymbol{\rho})\, \mathrm{dt} \tag{84}$$

$\mathbf{f}(t, \mathbf{y}, \boldsymbol{\rho})$: the right-hand side function of the ODE, which is defined in (58).
$\frac{\partial \mathbf{f}}{\partial \mathbf{y}}$: the Jacobian of the right-hand side function with respect to the state vector, which is defined in (69).
$\frac{\partial \mathbf{f}}{\partial \boldsymbol{\rho}}$: the Jacobian of the right-hand side function with respect to the parameters, which is defined in (70).
$g(\mathbf{y}, \boldsymbol{\rho})$: the function which is defined in (84).
$\frac{\partial g}{\partial \boldsymbol{\rho}}$: the partial derivative of $g(\mathbf{y}, \boldsymbol{\rho})$ with respect to the parameters $\boldsymbol{\rho}$.
$\frac{\partial g}{\partial \mathbf{y}}$: the partial derivative of $g(\mathbf{y}, \boldsymbol{\rho})$ with respect to the state vector $\mathbf{y}$.

$\boldsymbol{\lambda}_s$: the sensitivities of the objective function ψ with respect to the initial conditions, which is $\boldsymbol{\mu}$ from (38). $\boldsymbol{\lambda}_s$ should be initialized to be $\mathbf{0}$ from (43).

$\boldsymbol{\mu}_s$: the sensitivities of the objective function ψ with respect to the parameters. In this paper, $\boldsymbol{\mu}_s$ is the output of the sensitivities. It should be initialized to be $\mathbf{0}$

These functions are provided to the adjoint fully implicit Runge-Kutta solver to compute the forward solution and the sensitivities.

Real and complex differences approximation

Although impractical from the computational point of view, the finite differences approximation can be very useful to detect errors in the derivatives. The first order approximation for the derivatives with real perturbations, read as follows.

$$\frac{\mathrm{d}\psi}{\mathrm{d}\boldsymbol{\rho}_k} = \frac{\psi(\boldsymbol{\rho}+\delta\mathbf{e}_k)-\psi(\boldsymbol{\rho})}{\delta} \tag{85}$$

The truncation error in this case is $\vartheta(h)$, where h is the time-step, so it can be controlled decreasing it. Nevertheless, small h results in loss-of-significance (cancellation) errors due to the substraction. This fact can make this derivatives completely useless or untrustworthy.

The first order approximation for the derivatives with complex perturbations is the following.

$$\frac{\mathrm{d}\psi}{\mathrm{d}\boldsymbol{\rho}_k} = \frac{\Im\left(\psi(\boldsymbol{\rho}+i\delta\mathbf{e}_k)\right)}{\delta} \tag{86}$$

Where i is the imaginary unit and $\Im$ is the imaginary part of a complex number. The approach is much more trustworthy than the previous one, since there are no loss-of-significance errors involved in the calculation of the approximation, because there are not substractions in the imaginary parts and therefore the increments can be chosen arbitrarily small. The practical difficulty to apply complex finite differences is that not all codes can be changed easily to accommodate complex arithmetic. Special attention should be paid to the third party functions involved in the code (*transpose* functions, *norm* functions, numerical integrator chosen, etc).

This approach was used in this study to validate all the derivatives and results presented.

NUMERICAL EXPERIMENT

The mechanism chosen to test the formulations proposed in the paper is the five bar mechanism with 2 degrees of freedom shown in Fig.1. The five bars are constrained by five revolute joints located in points A, 1, 2, 3 and B. The five bars are constrained by five revolute joints located in points A, 1, 2, 3 and B. The masses of the bars are $m_1 = 1\ kg$, $m_2 = 1.5\ kg$, $m_3 = 1.5\ kg$, $m_4 = 1\ kg$ and the polar moments of inertia are calculated under the assumption of a uniform distribution of mass. The mechanism is subjected to the action of gravity and two elastic forces coming from the springs. The stiffness coefficients of the springs are $k_1 = k_2 = 100\ N/m$ and their natural lengths are initially chosen $L_{01} = \sqrt{2^2+1^2}\ m$ and $L_{02} = \sqrt{2^2+0.5^2}\ m$, coincident with the initial configuration shown in Fig.1.

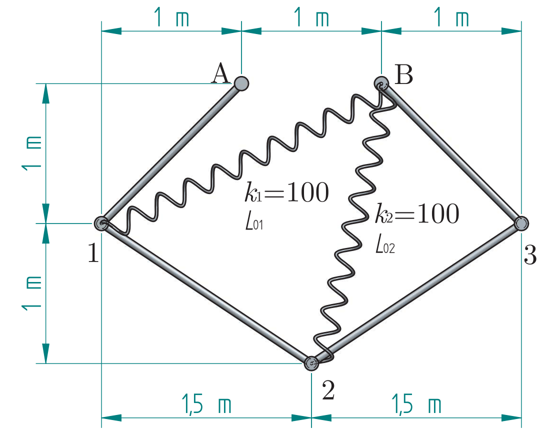

FIGURE 1. FIVE BAR MECHANISM

The mechanism can be balanced by properly selecting the two parameters $\boldsymbol{\rho}^{\mathrm{T}} = [L_{01}, L_{02}]$. Of course the problem can be solved by means of the static equations but the aim here is doing so by dynamical optimization.

The objective is to keep the mechanism still in the initial position which can be represented mathematically by the following objective function.

$$\psi = \int_{t_0}^{t_F} (\mathbf{r}_2-\mathbf{r}_{20})^{\mathrm{T}} (\mathbf{r}_2-\mathbf{r}_{20})\,\mathrm{d}t \tag{87}$$

Where $\mathbf{r}_2$ is the global position of the point 2 and $\mathbf{r}_{20}$ is the initial position of the same point.

The condition to obtain the minimum is the following.

$$\nabla_{\boldsymbol{\rho}}\psi = \mathbf{0} \tag{88}$$

The gradient (88) was obtained by the following approaches:

1. Direct sensitivity: using equation (34).

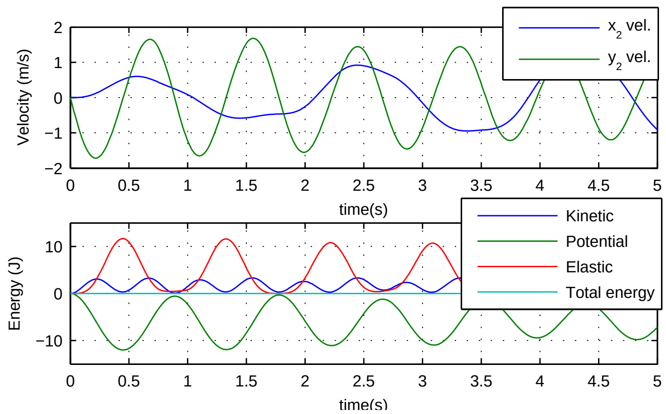

FIGURE 2. MECHANISM RESPONSE.

2. Adjoint sensitivity, explicit ODE: using equation (66).
3. Adjoint sensitivity, semi-explicit ODE: using equation (81).
4. Adjoint sensitivity with FATODE.
5. Numerical sensitivity with real perturbations.
6. Numerical sensitivity with complex perturbations.

The response of the system is shown in Fig.2 for a 5 seconds simulation. The upper plot represents the horizontal and vertical velocities of the point 2 while the lower one represents the energy taking as reference for the potential energy the initial configuration of the system.

The results for the sensitivities with the mentioned methods are presented in Table 1.

TABLE 1. RESULTS FOR THE FIVE BAR MECHANISM.

Approach	Parameters	$d\psi/dL_{01}$	$d\psi/dL_{02}$
1: Direct	$h = 10^{-2}s$	-4.2300	3.2112
2: Adjoint-1	$h = 10^{-2}s$	-4.2294	3.2092
3: Adjoint-2	$h = 10^{-2}s$	-42294	3.2087
3: FATODE	$Tol = 10^{-3}$	-4.2254	3.2083
4: Num. diff. real	$\delta = 10^{-7}m$	-4.2288	3.2116
5: Num. diff. complex	$\delta/i = 10^{-7}m$	-4.2288	3.2116

As can be seen in Table 1, all the approaches, except the numerical sensitivities with real perturbations, offer similar results which guarantees that the schemes proposed are correct. The numerical sensitivities with real perturbations are not reliable if

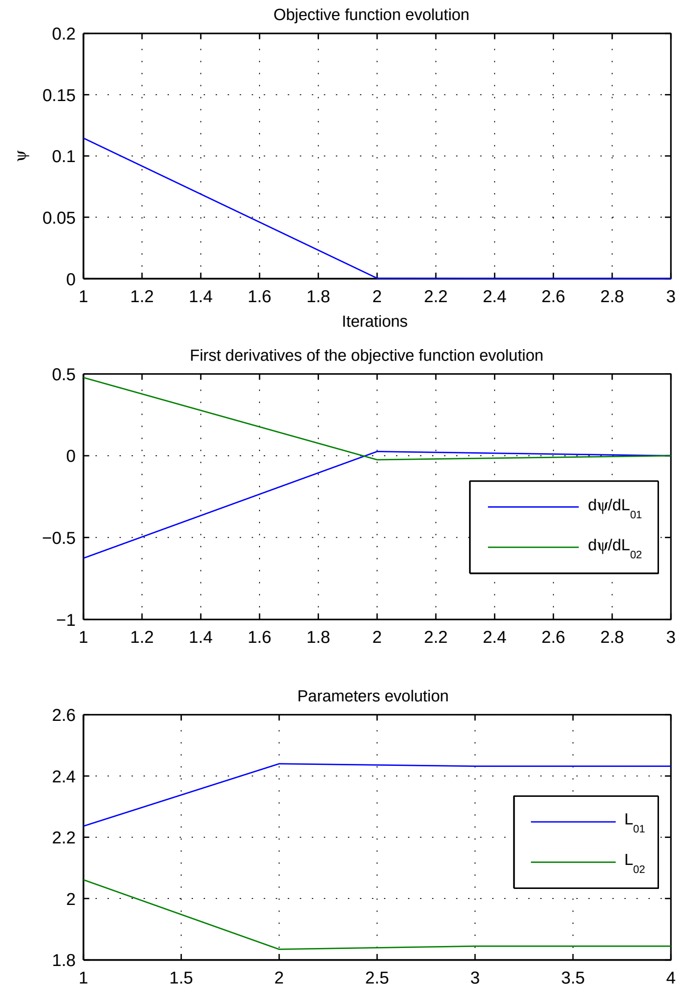

FIGURE 3. OBJECTIVE FUNCTION, GRADIENT AND PARAMETERS EVOLUTION.

accurate results for the sensitivities are important for the application to tackle. Given the simplicity of the system proposed, definitive conclusions in terms of efficiency cannot be stated.

The computed sensitivities can be employed for the optimization proposed. All the methods perform similar to solve the optimization problem. In this case the simulation time was reduced to $1s$ and the results for the objective function, derivatives and parameters are presented in Fig.3 for the *adjoint-1* approach. The plots for the direct approach coincide with the ones presented and they are not presented for clarity.

The optimization converges in three iterations, but in one is almost done. It is important to remark that approximate derivatives can be used to calculate the gradient and the optimization would converge at a lower pace.

Another important remark is that the tolerances in the solution of the forward dynamics are very important in order to obtain stable solutions for the TLM and adjoint ODEs, both of them strongly depends on the solution of the dynamics.

CONCLUSIONS

In this paper, two different approaches for the sensitivity analysis of multibody systems based on Maggi's formulations were developed: the direct sensitivity approach and the adjoint sensitivity approach. For the adjoint sensitivity approach, two different sets of equations were presented, based on the way to consider the equations of motion. Moreover, a strategy to validate the computed sensitivities was proposed.

The general expressions obtained can be employed to obtain the sensitivity equations of any ODE-like formulation, even penalty formulations.

A full strategy to validate the results obtained was proposed, all the results were tested and validated comparing: the two approaches between them; the results of another formulation of the equations of motion; a third party library and numerical results obtained by means of real and complex perturbations.

Finally, the results obtained were used for the dynamical optimization of a multibody system.

ACKNOWLEDGMENT

This work has been partially supported by NSF Award no. 1130667.

REFERENCES

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

[2] Ascher, U., and Petzold, L., 1998. *Computer methods for ordinary differential equations and differential-algebraic equations*. Philadelphia Society for Industrial and Applied Mathematics.

[3] Serna, M., Aviles, R., and Garca de Jaln, J., 1982. "Dynamic analysis of plane mechanisms with lower pairs in basic coordinates". *Mechanism and Machine Theory,* ***17***(6), pp. 397–403.

[4] Wehage, R., and Haug, E., 1982. "Generalized coordinate partitioning for dimension reduction in analysis of constrained mechanical systems". *Journal of Mechanical Design,* ***104***, pp. 247–255.

[5] Garcia de Jalon, J., and Bayo, E., 1994. *Kinematic and dynamic simulation of multibody systems: The real-time challenge*. Springer-Verlag, New York (USA).

[6] Haug, E. J., and Arora, J. S., 1979. *Applied Optimal Design: Mechanical and Structura Systems*. John Wiley & Sons Ltd.

[7] Cao, Y., Li, S., and Petzold, L., 2002. "Adjoint sensitivity analysis for differential-algebraic equations: algorithms and software". *Journal of Computational and Applied Mathematics,* ***149***(1), pp. 171 – 191. ¡ce:title¿Scientific and Engineering Computations for the 21st Century - Me thodologies and Applications Proceedings of the 15th Toyota Conference¡/ce:title¿.

[8] Cao, Y., Li, S., Petzold, L., and Serban, R., 2002. "Adjoint sensitivity analysis or differential-algebraic equations: The adjoint DAE system and its numerical solution". *SIAM Journal on Scientific Computing,* ***24***(3), pp. 1076–1089.

[9] Pagalday, J., and Avello, A., 1997. "Optimization of multibody dynamics using object oriented programming and a mixed numerical-symbolic penalty formulation". *Mechanism and Machine Theory,* ***32***(2), Feb, pp. 161–174.

[10] Dias, J., and Pereira, M., 1997. "Sensitivity analysis of rigid-flexible multibody systems". *Multibody System Dynamics,* ***1***, pp. 303–322. 10.1023/A:1009790202712.

[11] Anderson, K. S., and Hsu, Y., 2002. "Analytical fully-recursive sensitivity analysis for multibody dynamic chain systems". *Multibody System Dynamics,* ***8***, pp. 1–27. 10.1023/A:1015867515213.

[12] Ding, J.-Y., Pan, Z.-K., and Chen, L.-Q., 2007. "Second order adjoint sensitivity analysis of multibody systems described by differential-algebraic equations". *Multibody System Dynamics,* ***18***, pp. 599–617. 10.1007/s11044-007-9080-4.

[13] Schaffer, A., 2006. "Stabilized index-1 differential-algebraic formulations for sensitivity analysis of multibody dynamics". *Proceedings of the Institution of Mechanical Engineers Part K- Journal of Multi-Body Dynamics,* ***220***(3), SEP, pp. 141–156.

[14] Zhang, H., and Sandu, A., 2012. Fatode: a library for forward, adjoint, and tangent linear integration of odes, November.

[15] Bayo, E., Garca de Jaln, J., and Serna, M., 1988. "A modified lagrangian formulation for the dynamic analysis of constrained mechanical systems". *Computer Methods in Applied Mechanics and Engineering,* ***71***(2), 11, pp. 183–195.