Sensitivity analysis for multibody systems formulated on a Lie group

Valentin Sonneville and Olivier Brüls

Department of Aerospace and Mechanical Engineering (LTAS), University of Liège
Chemin des Chevreuils 1, 4000 Liège, Belgium

Abstract

A direct differentiation method and an adjoint variable method are proposed for the efficient semi-analytical evaluation of the sensitivities of multibody systems formulated in a matrix Lie group framework. These methods rely on the linearization of the equations of motion and/or of the time integration procedure. The simpler structure of the equations of motion in the Lie group formalism appears as an advantage for that purpose. Lie bracket contributions and the nonlinearity of the exponential map need to be taken into account in the sensitivity algorithms. Nevertheless, essential characteristics of formulations of the direct differentiation method and the adjoint variable method on linear spaces are recovered. Some implementation issues are discussed and two relevant examples illustrate the properties of these methods.

Keywords: multibody systems, Lie groups, sensitivity analysis, direct differentiation method, adjoint variable method

1 Introduction

In the context of the large scale optimization of multibody systems based on dynamic simulations, a critical step is the efficient evaluation of the sensitivities of an objective function or a design constraint with respect to a set of design parameters. Relevant applications involving many design variables and potentially many design criteria can be found in structural optimization [1] and optimal control problems [2]. Although finite difference approximations of the sensitivities can be easily implemented in general, they become rapidly inefficient when the number of design variables increases. In contrast, more reasonable performance can be expected from semi-analytical methods such as the direct differentiation or the adjoint variable methods. They need to be manually implemented within a simulation code, which allows to optimize the computational procedure for the problem at hand. They mainly require the availability of the complete linearized form of the equations of motion and/or of the time integration procedure in the code. Their development and implementation are thus closely related to the formulation exploited for the analysis of the multibody system and to the structure of the simulation code. The automatic differentiation method is also widely used [3, 4]. Based on an available computation code written in Fortran or C, it yields an additional code dedicated to the computation of the sensitivities. Relevant discussions and applications of this method can be found for instance in [5]. As it is shown, on the one hand it does a large amount of work for the user and provides accurate results, but on the other hand it cannot be used as a black box and the resulting code must considered with caution.

This paper shows that semi-analytical methods for sensitivity analysis can be naturally implemented in a simulation code for multibody systems based on a Lie group formalism. We consider mechanical systems described using an absolute coordinate formulation, i.e. using displacements and rotations with respect to an inertial reference frame [6]. At first, these two kinds of variables have a different nature: displacements are classical vectors while rotations are linear maps of vectors, namely rotation matrices. However, both can be handled at the same time if the configuration space is described as a k-dimensional manifold G that has a matrix Lie group structure. According to [7, 8], the Lie group approach can be exploited not only to describe the configuration, but also to formulate the equations of motion and to integrate them in the time domain. This approach offers a number of advantages since no parameterization of the motion is needed, parameterization singularities are avoided, the equations of motion take a simpler form, their non-linearities are reduced and the frame invariance property is naturally satisfied. Moreover, the linearization of the equations of motion is greatly simplified, which is an attractive property for the development of algorithms for sensitivity analysis.

Ref. [9] focused on the direct differentiation method for dynamic systems on the group of 3D rotations $SO(3)$. In the continuity of that work, the developments of a direct differentiation method and an adjoint variable method in a general Lie group setting are the main contributions of the present paper. The literature on classical sensitivity analysis methods for dynamic systems evolving on a linear space is quite rich, see e.g. [10, 11, 12, 13, 14, 15, 16, 17, 18]. However, the extension of those methods to systems evolving on non-linear spaces is not trivial. Firstly, Lie groups are characterized by the lack of commutativity of derivatives, which is measured by the so-called Lie bracket or commutator. As will be shown here, Lie bracket contributions, which vanish for systems on linear spaces, need to be properly taken into account when developing sensitivity algorithms on a Lie group. Secondly, whereas classical time integration schemes systematically rely on linear formulae, Lie group integrators rely on non-linear time integration formulae involving most often the exponential map. The non-linearity of the exponential map plays an important role in the present analysis. The paper shows that, based on a careful understanding of these issues, the Lie group setting allows an efficient development and implementation of semi-analytical methods for sensitivity analysis. The resulting algorithms open the door for the optimization of 3D flexible multibody systems, e.g. for optimal structural and control design.

2 Lie group description of mechanical systems

Following the finite element approach for flexible multibody systems, see [6], the modelling of a mechanical system relies on the definition of a representative set of nodes. Kinematic assumptions are made to describe elements, e.g. rigid bodies or beams, and rotation matrices are introduced. Therefore, the configuration of a mechanical system is described in terms of absolute translations of the nodes and rotation matrices, which are collected in a variable denoted q. The configuration space, denoted G, is a k-dimensional manifold with a matrix Lie group structure, and we write $q \in G$. The nodal translations and rotations are generally not independent and m kinematic constraints restrict the configuration space to a subspace $G_{\mathbf{\Phi}}$ of dimension $k - m$:

$$G_{\mathbf{\Phi}} = \{q \in G \ : \mathbf{\Phi}(q) = \mathbf{0}\}$$

Let us introduce some fundamentals about Lie groups. One may refer to [19] or [20] for a more detailed description. A group is a set of elements with a composition rule, that associates an element of the group to two elements of the group: if $q_1, q_2 \in G$, then $q_1 \circ q_2 = q_3 \in G$. In the present case, we consider matrix groups and the composition rule is the classical matrix product written as $q_1 q_2 = q_3$. This composition rule has several properties, e.g. the existence of a neutral element e ($qe = eq = q$), which is simply the identity matrix, and the existence of an inverse ($\forall q \in G,\ \exists!\ q^{-1}:\ \ qq^{-1} = q^{-1}q = e$). It follows that the elements of such a matrix group are square and invertible matrices. A matrix Lie group is a continuous matrix group for which the composition rule and the inverse are smooth. Therefore, a matrix Lie group is, geometrically speaking, a differentiable manifold, and differential geometry can be used to perform operations on the group.

The composition operation can be seen as a particular map called the left translation map:

$$L_y(q) : G \to G,\ q \to yq \tag{1}$$

The tangent space at q is a vector space and is denoted T_qG. The notation $\delta(\bullet)$ means an arbitrary infinitesimal variation of the argument, and thus δq is an arbitrary element belonging to T_qG. Similarly, $\delta(yq)$ belongs to $T_{yq}G$, the tangent space at yq. The definition (1) yields a relationship between the tangent spaces T_qG and $T_{yq}G$:

$$\delta(yq) = \delta(L_y(q)) = L_*(\delta q) = y\delta q \tag{2}$$

where $L_* :\ T_qG \to T_{yq}G$ is an invertible linear map. The particular case $y = q^{-1}$ leads to an interesting fact: the tangent space of any $q \in G$ is related to the tangent space at the identity element e through a linear map.

The tangent space at the identity of a Lie group is called the Lie algebra $\mathfrak{g}$. It is isomorphic to $\mathbb{R}^k$ through the invertible linear map:

$$\widetilde{(\bullet)} : \mathbb{R}^k \to \mathfrak{g},\ \mathbf{x} \to \widetilde{\mathbf{x}}$$

Using these notations in Eq. (2), we obtain

$$\delta q = q\widetilde{\delta\mathbf{q}} \tag{3}$$

Also, the derivative of q with respect to any parameter $s \in \mathbb{R}$ is written as:

$$\frac{dq}{ds} = q\widetilde{\mathbf{u}} \tag{4}$$

where $\widetilde{\mathbf{u}}$ is an element of the Lie algebra. Furthermore, combining the derivative of Eq. (3) and the variation of $\widetilde{\mathbf{u}}$ yields

$$\delta\widetilde{\mathbf{u}} = \frac{d(\delta\widetilde{\mathbf{q}})}{ds} + \widetilde{\mathbf{u}}\widetilde{\delta\mathbf{q}} - \widetilde{\delta\mathbf{q}}\widetilde{\mathbf{u}} = \frac{d(\delta\widetilde{\mathbf{q}})}{ds} + [\widetilde{\mathbf{u}}, \widetilde{\delta\mathbf{q}}] \tag{5}$$

where $[\bullet, \bullet]$ is called the Lie bracket. This shows that in general derivatives do not commute due to the non-linearity of the Lie group. Equation (5) can be written in terms of vectors in $\mathbb{R}^k$:

$$\delta\mathbf{u} = \frac{d(\delta\mathbf{q})}{ds} + \hat{\mathbf{u}}\delta\mathbf{q} \tag{6}$$

where $\hat{\bullet}$ is a linear operator which maps a vector in $\mathbb{R}^k$ into a $k \times k$ matrix.

Equation (4) can be seen as a differential equation on the Lie group. If the vector field $\widetilde{\mathbf{u}}$ does not depend on s, the solution is

$$q(s) = q_0 \exp(\widetilde{\mathbf{u}}s)$$

where $q_0 = q(0)$ and exp is the exponential operator, which maps an element of the Lie algebra to an element of the Lie group:

$$\exp : \mathfrak{g} \to G,\ \widetilde{\mathbf{x}} \to \exp(\widetilde{\mathbf{x}})$$

As the Lie algebra is isomorphic to $\mathbb{R}^k$, the exponential map introduces a local parametrization of the Lie group around any $q_0 \in G$. Indeed, any $q \in G$ may be represented as a function of $\widetilde{\mathbf{x}} \in \mathfrak{g}$ using the exponential operator and the composition with q_0 according to

$$q = q_0 \exp(\widetilde{\mathbf{x}}) \tag{7}$$

Thus, the derivative with respect to a parameter $s \in \mathbb{R}$ reads

$$\frac{dq}{ds} = q_0 \ \mathrm{Dexp}(\widetilde{\mathbf{x}}) \cdot \frac{d\widetilde{\mathbf{x}}}{ds} \tag{8}$$

where $\mathrm{Dexp}(\widetilde{\mathbf{x}}) \cdot \widetilde{\mathbf{y}}$ is the directional derivative of the exponential map in the direction of $\widetilde{\mathbf{y}}$. The comparison of Eq. (8) with Eq. (4) yields

$$\widetilde{\mathbf{u}} = (\exp(\widetilde{\mathbf{x}}))^{-1}\mathrm{Dexp}(\widetilde{\mathbf{x}}) \cdot \frac{d\widetilde{\mathbf{x}}}{ds}$$

This gives a relationship between the element $\widetilde{\mathbf{u}}$ involved in the left translation map and the element $\widetilde{\mathbf{x}}$ defined by the exponential map representation in Eq. (7). It can be expressed as a linear relationship from $\mathbb{R}^k$ to $\mathbb{R}^k$:

$$\mathbf{u} = \mathbf{T}(\mathbf{x})\frac{d\mathbf{x}}{ds} \tag{9}$$

Likewise, we can also write

$$\delta\mathbf{q} = \mathbf{T}(\mathbf{x})\delta\mathbf{x}$$

3 Equations of motion in a Lie group setting

3.1 General formulation

In this section, the equations of motion and their linearized form are derived from Hamilton's principle using the Lagrange multipliers method.

In order to evaluate the kinetic energy of the mechanical system, the velocity field is introduced with an element of the Lie algebra according to Eq. (4):

$$\dot{q} = q\widetilde{\mathbf{v}} \tag{10}$$

where the variable $\widetilde{\mathbf{v}} \in \mathfrak{g}$, or $\mathbf{v} \in \mathbb{R}^k$, thus represents the velocities of the system under study. $\dot{\bullet}$ refers to the time derivative. Then, the kinetic energy of the system can be written as a quadratic form of the vector $\mathbf{v}$:

$$\mathcal{K}(\mathbf{v}) = \frac{1}{2}\mathbf{v}^T\mathbf{M}\mathbf{v}$$

where $\mathbf{M}$ is the symmetric mass matrix, which turns out to be constant in most practical situations.

Using the Lagrange multipliers method to treat the constraints $\mathbf{\Phi}(q) = 0$, Hamilton's principle reads

$$\delta \int_{t_i}^{t_f} (\mathcal{K}(\mathbf{v}) - \mathcal{V}(q) - \boldsymbol{\lambda}^T\mathbf{\Phi}(q)) \ dt = 0$$

together with the condition that the variations of the initial and final configurations vanish. $\boldsymbol{\lambda}$ are the Lagrange multipliers and $\mathcal{V}(q)$ is the potential energy. Computing the variations yields

$$\int_{t_i}^{t_f} (\delta\mathbf{v}^T\mathbf{M}\mathbf{v} - \delta\mathbf{q}^T\mathbf{f}(q) - \delta\mathbf{q}^T\mathbf{B}^T(q)\boldsymbol{\lambda} - \delta\boldsymbol{\lambda}^T\mathbf{\Phi}(q)) \ dt = 0 \tag{11}$$

where the force vector $\mathbf{f}$ and the constraint gradient $\mathbf{B}$ are defined from the directional derivatives of $\mathcal{V}$ and $\mathbf{\Phi}$, respectively:

$$D\mathcal{V} \cdot \widetilde{\delta\mathbf{q}} = \delta\mathbf{q}^T\mathbf{f}(q)$$
$$D\mathbf{\Phi} \cdot \widetilde{\delta\mathbf{q}} = \mathbf{B}(q)\delta\mathbf{q}$$

where the directional derivative of a function $f(q)$ is the direction $\delta\widetilde{\mathbf{q}}$ is defined as

$$D_q f \cdot \widetilde{\delta\mathbf{q}} = \frac{d}{dt}\bigg|_{t=0} \Big(f\big(q\exp(t\widetilde{\delta\mathbf{q}})\big)\Big) \tag{12}$$

Equation (6) leads to

$$\delta\mathbf{v} = \delta\dot{\mathbf{q}} + \hat{\mathbf{v}}\delta\mathbf{q} \tag{13}$$

and inserting this expression into Eq. (11), an integration by parts yields

$$[\delta\mathbf{q}^T\mathbf{M}\mathbf{v}]_{t_i}^{t_f} - \int_{t_i}^{t_f} (\delta\mathbf{q}^T(\mathbf{M}\dot{\mathbf{v}} - \hat{\mathbf{v}}^T\mathbf{M}\mathbf{v} + \mathbf{f}(q) + \mathbf{B}^T(q)\boldsymbol{\lambda}) + \delta\boldsymbol{\lambda}^T\mathbf{\Phi}(q))\ dt = 0 \tag{14}$$

where the first term drops by hypothesis and the assumption of a constant mass matrix has been made. Accounting for the kinematic compatibility condition (10) and considering the arbitrary character of the variations $\delta\mathbf{q}$ and $\delta\boldsymbol{\lambda}$, we obtain a set of differential-algebraic equations (DAE):

$$\dot{q} = q\widetilde{\mathbf{v}} \tag{15}$$
$$\mathbf{M}\dot{\mathbf{v}} - \hat{\mathbf{v}}^T\mathbf{M}\mathbf{v} + \mathbf{f}(q) + \mathbf{B}^T(q)\boldsymbol{\lambda} = \mathbf{0} \tag{16}$$
$$\mathbf{\Phi}(q) = \mathbf{0} \tag{17}$$

Notice that the first equation is a matrix equation and that the other two are vector equations. In the following, the left hand side term of Eq. (16) is denoted as $\mathbf{r}(q, \mathbf{v}, \dot{\mathbf{v}}, \boldsymbol{\lambda})$. The initial conditions in position and velocity for the mechanical system satisfy

$$\boldsymbol{\zeta}(q(t_0)) = e \tag{18}$$
$$\boldsymbol{\chi}(q(t_0), \mathbf{v}(t_0)) = \mathbf{0} \tag{19}$$

In simple cases, one may have $\boldsymbol{\zeta}(q(t_0)) = q_0^{-1}q(t_0)$ and $\boldsymbol{\chi}(q(t_0), \mathbf{v}(t_0)) = \mathbf{v}(t_0) - \mathbf{v}_0$.

Let us develop the linearized form of the equations of motion, which is useful for implicit time integration as well as for the derivation of the sensitivity equations. For infinitesimal variations about a point $(q^*, \mathbf{v}^*, \dot{\mathbf{v}}^*, \boldsymbol{\lambda}^*)$, we have

$$\begin{aligned} \delta\mathbf{r} &= \mathbf{K}_t\delta\mathbf{q} + \mathbf{C}_t\delta\mathbf{v} + \mathbf{M}\delta\dot{\mathbf{v}} + \mathbf{B}^T\delta\boldsymbol{\lambda} \\ \delta\mathbf{\Phi} &= \mathbf{B}\delta\mathbf{q} \end{aligned}$$

where $\delta\mathbf{r} = \mathbf{r}(q^*(e+\widetilde{\delta\mathbf{q}}), \mathbf{v}^*+\delta\mathbf{v}, \dot{\mathbf{v}}^*+\delta\dot{\mathbf{v}}, \boldsymbol{\lambda}^*+\delta\boldsymbol{\lambda}) - \mathbf{r}(q^*, \mathbf{v}^*, \dot{\mathbf{v}}^*, \boldsymbol{\lambda}^*)$ and $\delta\mathbf{\Phi} = \mathbf{\Phi}(q^*(e+\widetilde{\delta\mathbf{q}})) - \mathbf{\Phi}(q^*)$. The tangent damping and stiffness matrices are defined from $\partial\mathbf{r}(q, \mathbf{v}, \dot{\mathbf{v}}, \boldsymbol{\lambda})/\partial\mathbf{v} = \mathbf{C}_t$ and the directional derivative $D_q\mathbf{r}(q, \mathbf{v}, \dot{\mathbf{v}}, \boldsymbol{\lambda}) \cdot \widetilde{\delta\mathbf{q}} = \mathbf{K}_t\delta\mathbf{q}$, respectively.

3.2 Example

As an example, the equations of motion for a spinning top formulated on $\mathbb{R}^3 \times SO(3)$ are given here (see e.g. [22] for more information and an alternative formulation on $SE(3)$). Let us consider a rigid body of mass m and whose inertia tensor is $\mathbf{J}$. The inertial frame is centered at the fixed point of the body. Following the same philosophy as the finite element method [6] the configuration variables are the position vector $\mathbf{x}$ of the center of mass and the rotation matrix $\mathbf{R}$ representing the orientation of the rigid body, i.e. $q = (\mathbf{R}, \mathbf{x})$, and the fixed point constraint is introduced in a second step. Explicitly, q can be represented by the 7×7 matrix

$$q = \begin{bmatrix} \mathbf{R} & \mathbf{0}_{3\times 3} & \mathbf{0}_{3\times 1} \\ \mathbf{0}_{3\times 3} & \mathbf{I}_{3\times 3} & \mathbf{x} \\ \mathbf{0}_{1\times 3} & \mathbf{0}_{1\times 3} & 1 \end{bmatrix} \tag{20}$$

The composition law for this matrix Lie group formally reads $(\mathbf{R}_1, \mathbf{x}_1)(\mathbf{R}_2, \mathbf{x}_2) = (\mathbf{R}_1\mathbf{R}_2, \mathbf{x}_1 + \mathbf{x}_2)$, and we have $e = (\mathbf{I}_{3\times 3}, \mathbf{0})$ and $(\mathbf{R}, \mathbf{x})^{-1} = (\mathbf{R}^T, -\mathbf{x})$. Thus, referring to (10), the velocity vector $\mathbf{v}$ and $\hat{\mathbf{v}}$ are

$$\mathbf{v} = \begin{bmatrix} \dot{\mathbf{x}} \\ \mathbf{\Omega} \end{bmatrix} \quad ; \quad \hat{\mathbf{v}} = \begin{bmatrix} \mathbf{0}_{3\times 3} & \mathbf{0}_{3\times 3} \\ \mathbf{0}_{3\times 3} & \widetilde{\mathbf{\Omega}} \end{bmatrix}$$

where $\mathbf{\Omega}$ is the 3×1 material rotational velocity vector containing the non-diagonal elements of the skew-symmetric matrix $\widetilde{\mathbf{\Omega}} = \mathbf{R}^T\dot{\mathbf{R}}$. Indeed, the tilde operator of $SO(3)$ defines a map between $\mathbb{R}^3$ and the Lie algebra $\mathfrak{so}(3)$ as

$$\mathbf{\Omega} = \begin{bmatrix} \Omega_1 \\ \Omega_2 \\ \Omega_3 \end{bmatrix} \in \mathbb{R}^3 \qquad \text{and} \qquad \widetilde{\mathbf{\Omega}} = \begin{bmatrix} 0 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 0 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 0 \end{bmatrix} \in \mathfrak{so}(3)$$

The body is subjected to constraint equations since it has a fixed point, which is at the origin of the inertial frame. The body undergoes gravity, denoted by the vector $\mathbf{g}$ acting on the center of mass. Finally, the equations of motion (16-17) read

$$\begin{aligned} m\ddot{\mathbf{x}} + \boldsymbol{\lambda} &= m\mathbf{g} && (21) \\ \mathbf{J}\dot{\mathbf{\Omega}} + \mathbf{\Omega} \times \mathbf{J}\mathbf{\Omega} - \widetilde{\mathbf{X}}\mathbf{R}^T\boldsymbol{\lambda} &= \mathbf{0} && (22) \\ \mathbf{x} - \mathbf{R}\mathbf{X} &= \mathbf{0} && (23) \end{aligned}$$

where Eq. (23) are the constraint equations and $\mathbf{X}$ is the position vector of the center of mass with respect to the inertial frame. $\mathbf{X}$ is expressed in the material frame. We obtain a system of nine equations including three constraints, which means, as expected, that the rigid body has three degrees of freedom. The tangent matrices involved in the linearized equations of motion are

$$\mathbf{M} = \begin{bmatrix} m\mathbf{I}_{3\times 3} & \mathbf{0}_{3\times 3} \\ \mathbf{0}_{3\times 3} & \mathbf{J} \end{bmatrix} \quad ; \quad \mathbf{C}_t = \begin{bmatrix} \mathbf{0}_{3\times 3} & \mathbf{0}_{3\times 3} \\ \mathbf{0}_{3\times 3} & \widetilde{\mathbf{\Omega}}\mathbf{J} - \widetilde{\mathbf{J}\mathbf{\Omega}} \end{bmatrix} \quad ; \quad \mathbf{K}_t = \begin{bmatrix} \mathbf{0}_{3\times 3} & \mathbf{0}_{3\times 3} \\ \mathbf{0}_{3\times 3} & -\widetilde{\mathbf{X}}\widetilde{\mathbf{R}^T\boldsymbol{\lambda}} \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} \mathbf{I}_{3\times 3} & \mathbf{R}\widetilde{\mathbf{X}} \end{bmatrix}$$

4 Time integration method

Equations (15-17) can be solved using the Lie group version of the generalized-α scheme proposed in Ref. [8]. This algorithm preserves the Lie group structure of the problem. It has a proven unconditional second-order convergence and some numerical damping can be used to lessen the high frequency content.

The integration method relies on the discretized equations of motion

$$\begin{aligned} q_{n+1} &= q_n \exp(\widetilde{\mathbf{x}}_{n+1}) && (24) \\ \mathbf{M}\dot{\mathbf{v}}_{n+1} - \hat{\mathbf{v}}_{n+1}^T\mathbf{M}\mathbf{v}_{n+1} + \mathbf{f}(q_{n+1}) + \mathbf{B}^T(q_{n+1})\boldsymbol{\lambda}_{n+1} &= \mathbf{0} && (25) \\ \mathbf{\Phi}(q_{n+1}) &= \mathbf{0} && (26) \end{aligned}$$

and the time integration formulae

$$
\begin{aligned}
\mathbf{x}_{n+1} &= h\mathbf{v}_n + (0.5-\beta)h^2\mathbf{a}_n + \beta h^2\mathbf{a}_{n+1} && (27)\\
\mathbf{v}_{n+1} &= \mathbf{v}_n + (1-\gamma)h\mathbf{a}_n + \gamma h\mathbf{a}_{n+1} && (28)\\
(1-\alpha_m)\mathbf{a}_{n+1} + \alpha_m\mathbf{a}_n &= (1-\alpha_f)\dot{\mathbf{v}}_{n+1} + \alpha_f\dot{\mathbf{v}}_n && (29)
\end{aligned}
$$

where n refers to the time step and h is the time step size. The numerical parameters of the integrator can be selected to achieve a desired spectral radius $\rho \in [0,1)$ at high frequency:

$$\alpha_m = \frac{2\rho-1}{\rho+1} \quad ; \quad \alpha_f = \frac{\rho}{\rho+1} \quad ; \quad \gamma = 0.5 + \alpha_f - \alpha_m \quad ; \quad \beta = \frac{(\gamma+0.5)^2}{4}$$

At each time step, this set of equations is solved for the unknown $\mathbf{x}_{n+1}$, $\mathbf{v}_{n+1}$, $\dot{\mathbf{v}}_{n+1}$ and $\mathbf{a}_{n+1}$ by a Newton iterative procedure. The linearization of Eqs. (24) gives $\Delta\mathbf{q}_{n+1} = \mathbf{T}(\mathbf{x}_{n+1})\Delta\mathbf{x}_{n+1}$ and the linearization of Eqs. (27-28) gives $\Delta\dot{\mathbf{v}}_{n+1} = (1-\alpha_m)/(\beta h^2(1-\alpha_f))\Delta\mathbf{x}_{n+1}$ and $\Delta\mathbf{v}_{n+1} = \gamma/(\beta h)\Delta\mathbf{x}_{n+1}$, where $\Delta(\bullet)$ denotes a small variation. Thus, the linearized form of Eqs. (25-26) is

$$\begin{bmatrix} \Delta\mathbf{r} \\ \Delta\boldsymbol{\Phi} \end{bmatrix} = \begin{bmatrix} ((1-\alpha_m)/(\beta h^2(1-\alpha_f)))\mathbf{M} + \gamma/(\beta h)\mathbf{C}_t + \mathbf{K}_t\mathbf{T}) & \mathbf{B}^T \\ \mathbf{BT} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \Delta\mathbf{x}_{n+1} \\ \Delta\boldsymbol{\lambda}_{n+1} \end{bmatrix} \tag{30}$$

5 Criterion function and design parameter

Let us consider a criterion function with the following form

$$\Psi = \Psi_G + \Psi_F = G(q(t_f), \mathbf{v}(t_f), p) + \int_{t_0}^{t_f} F(q, \mathbf{v}, \dot{\mathbf{v}}, \boldsymbol{\lambda}, p)\; dt \tag{31}$$

where t_0 and t_f are the initial time and the final time. In optimization problems, a criterion function can be an objective function to be minimized or a constraint to be satisfied. The function depends on the response of the mechanical system as well as on a design parameter p. In the following, the notations are simplified by referring to a single scalar criterion function and a single scalar design parameter. However, the development also applies for problems with several criterion functions or/and several parameters if suitable indices are introduced in the notation.

6 Direct differentiation method

In the direct differentiation method, the equations are directly differentiated with respect to the parameters to compute the sensitivities. This method has been treated in the literature for linear spaces, see e.g. [17]. Based on the reasoning of [9], we extend it to equations of motion formulated on a Lie group. The integration of the differentiated equations of motion and the differentiation of the time-discretized equations of motion lead to different algorithms. Both cases are considered below.

According to Eq. (4), the derivative of q with respect to the parameter p is represented by Lie algebra element $\widetilde{\mathbf{w}}$ such that

$$q' = q\widetilde{\mathbf{w}} \tag{32}$$

where $\bullet'$ denotes the derivative with respect to parameter p. Considering Eq. (6), a relation is obtained between the second derivatives $\dot{\mathbf{w}}$ and $\mathbf{v}'$, which involves the Lie bracket:

$$\dot{\mathbf{w}} = \mathbf{v}' - \hat{\mathbf{v}}\mathbf{w}$$

The derivative of any function $f(q, \mathbf{v}, \dot{\mathbf{v}}, \boldsymbol{\lambda}, p)$ with respect to a parameter p is thus

$$(f(q, \mathbf{v}, \dot{\mathbf{v}}, p))' = f_{\mathbf{q}}\mathbf{w} + f_{\mathbf{v}}\mathbf{v}' + f_{\dot{\mathbf{v}}}\dot{\mathbf{v}}' + f_{\boldsymbol{\lambda}}\boldsymbol{\lambda}' + f_p$$

where $f_{\mathbf{q}}\mathbf{w}$ is the directional derivative of f in the direction of $\widetilde{\mathbf{w}}$; $f_{\mathbf{v}}$, $f_{\dot{\mathbf{v}}}$, $f_{\boldsymbol{\lambda}}$ and f_p are the partial derivatives of f with respect to $\mathbf{v}$, $\dot{\mathbf{v}}$, $\boldsymbol{\lambda}$ and p, respectively. The derivative of the criterion in Eq. (31) with respect to the parameter p is then

$$\frac{d\Psi}{dp} = (G_{\mathbf{q}}\mathbf{w} + G_{\mathbf{v}}\mathbf{v}' + G_p)|_{t_f} + \int_{t_i}^{t_f} (F_{\mathbf{q}}\mathbf{w} + F_{\mathbf{v}}\mathbf{v}' + F_{\dot{\mathbf{v}}}\dot{\mathbf{v}}' + F_{\boldsymbol{\lambda}}\boldsymbol{\lambda}' + F_p)\, dt \tag{33}$$

In the direct differentiation method, the evaluation of $d\Psi/dp$ thus requires the knowledge of the sensitivities $\mathbf{w}, \mathbf{v}', \dot{\mathbf{v}}'$ and $\boldsymbol{\lambda}'$ over the trajectory.

6.1 Differentiation of the equations of motion

The differentiation of Eq. (15-17) leads to a coupled system of first-order linear differential-algebraic equations for the sensitivities $\mathbf{w}$ and $\mathbf{v}'$

$$\begin{aligned} \dot{\mathbf{w}} &= \mathbf{v}' - \hat{\mathbf{v}}\mathbf{w} \\ \mathbf{M}\dot{\mathbf{v}}' + \mathbf{C}_t\mathbf{v}' + \mathbf{K}_t\mathbf{w} + \mathbf{B}^T\boldsymbol{\lambda}' &= -\mathbf{r}_p \\ \mathbf{B}\mathbf{w} &= -\boldsymbol{\Phi}_p \end{aligned}$$

where $\mathbf{r}_p$ and $\boldsymbol{\Phi}_p$ act as pseudo-loads. The initial conditions for these equations are obtained by differentiating the initial conditions of the mechanical system:

$$\begin{aligned} \boldsymbol{\zeta}_{\mathbf{q}}\mathbf{w}_0 &= -\boldsymbol{\zeta}_p \\ \boldsymbol{\chi}_{\mathbf{v}}\mathbf{v}_0' + \boldsymbol{\chi}_{\mathbf{q}}\mathbf{w}_0 &= -\boldsymbol{\chi}_p \end{aligned}$$

An elimination of $\mathbf{v}'$ from these equations leads to a system of second-order linear differential-algebraic equations for the sensitivities $\mathbf{w}$

$$\begin{aligned} \mathbf{M}\ddot{\mathbf{w}} + (\mathbf{C}_t + \mathbf{M}\hat{\mathbf{v}})\dot{\mathbf{w}} + (\mathbf{K}_t + \mathbf{C}_t\hat{\mathbf{v}} + \mathbf{M}\hat{\dot{\mathbf{v}}})\mathbf{w} + \mathbf{B}^T\boldsymbol{\lambda}' &= -\mathbf{r}_p \\ \mathbf{B}\mathbf{w} &= -\boldsymbol{\Phi}_p \end{aligned}$$

with the initial conditions

$$\begin{aligned} \boldsymbol{\zeta}_{\mathbf{q}}\mathbf{w}_0 &= -\boldsymbol{\zeta}_p \\ \boldsymbol{\chi}_{\mathbf{v}}\dot{\mathbf{w}}_0 + (\boldsymbol{\chi}_{\mathbf{q}} + \boldsymbol{\chi}_{\mathbf{v}}\hat{\mathbf{v}}_0)\mathbf{w}_0 &= -\boldsymbol{\chi}_p \end{aligned}$$

This system of equations can be solved for the sensitivities using the generalized-α time integration method for linear spaces with the same numerical parameters as for the nominal problem. The sensitivity equations at time step $n+1$ can be formulated in terms of a specific $(k+m) \times (k+m)$ matrix but, unfortunately, this matrix is different from the iteration matrix defined in Eq. (30) used to solve the nominal problem.

The Lie bracket operator, which is a mark of the non-linearity of the Lie group, appears through the contributions $\hat{\mathbf{v}}$ and $\hat{\dot{\mathbf{v}}}$. In the special case where G is a linear space, the Lie bracket vanishes, $\hat{\mathbf{v}} = \hat{\dot{\mathbf{v}}} = \mathbf{0}$, and the sensitivity equations reduce to their classical form for problems on $\mathbb{R}^k$.

6.2 Differentiation of the time integration method

Rather than differentiating the equations of motion on the Lie group, the time integration method introduced in section 4 can be differentiated. Relying on the definition in Eq. (32), Eq. (24) is differentiated as

$$\begin{aligned} q'_{n+1} &= q'_n \exp(\widetilde{\mathbf{x}}_{n+1}) + q_n \exp'(\widetilde{\mathbf{x}}_{n+1}) \\ \Leftrightarrow \widetilde{\mathbf{w}}_{n+1} &= (\exp(\widetilde{\mathbf{x}}_{n+1}))^{-1} \widetilde{\mathbf{w}}_n \exp(\widetilde{\mathbf{x}}_{n+1}) + (\exp(\widetilde{\mathbf{x}}_{n+1}))^{-1} \exp'(\widetilde{\mathbf{x}}_{n+1}) \end{aligned}$$

and finally, considering Eq. (9), we have the linear relationship

$$\mathbf{w}_{n+1} = \mathbf{D}(\exp(\widetilde{\mathbf{x}}_{n+1}))\mathbf{w}_n + \mathbf{T}(\mathbf{x}_{n+1})\mathbf{x}'_{n+1}$$

where $\mathbf{D}(q)$ is a matrix satisfying $\widetilde{\mathbf{D}(q)\mathbf{w}} = q^{-1}\widetilde{\mathbf{w}}q$. Hence, the differentiation of the discretized equations of motion (25-26) leads to

$$\begin{aligned} \mathbf{w}_{n+1} &= \mathbf{D}(\exp(\widetilde{\mathbf{x}}_{n+1}))\mathbf{w}_n + \mathbf{T}(\mathbf{x}_{n+1})\mathbf{x}'_{n+1} \\ \mathbf{M}\dot{\mathbf{v}}'_{n+1} + \mathbf{C}_t\mathbf{v}'_{n+1} + \mathbf{K}_t\mathbf{w}_{n+1} + \mathbf{B}^T(q_{n+1})\boldsymbol{\lambda}'_{n+1} &= -\mathbf{r}_p \\ \mathbf{B}(q_{n+1})\mathbf{w}_{n+1} &= -\boldsymbol{\Phi}_p \end{aligned}$$

The pseudo-loads and the initial conditions are the same as for the differentiation of the equations of motion in the previous section. The complete algorithm for sensitivity analysis is obtained by combining these equations with the differentiated form of the time integration formulae (27-29):

$$\begin{aligned} \mathbf{x}'_{n+1} &= h\mathbf{v}'_n + (0.5-\beta)h^2\mathbf{a}'_n + \beta h^2\mathbf{a}'_{n+1} \\ \mathbf{v}'_{n+1} &= \mathbf{v}'_n + (1-\gamma)h\mathbf{a}'_n + \gamma h\mathbf{a}'_{n+1} \\ (1-\alpha_m)\mathbf{a}'_{n+1} + \alpha_m\mathbf{a}'_n &= (1-\alpha_f)\dot{\mathbf{v}}'_{n+1} + \alpha_f\dot{\mathbf{v}}'_n \end{aligned}$$

where the time step size h is assumed to be constant. One may easily verify that the sensitivity equations at time step $n+1$ can be formulated based on exactly the same iteration matrix as for the nominal problem. We will come back to this point in the discussion below.

The exponential operator and its derivative appear explicitly in the sensitivity equations, which is a mark of the non-linearity of the Lie group. In the special case where G is a linear space, $\mathbf{D}$ and $\mathbf{T}$ are simply identity matrices and one obtains the classical sensitivity relations for problems in $\mathbb{R}^k$.

6.3 Discussion

In general, the algorithms obtained by time integration of the sensitivity equations or by differentiation of the time integrator are not the same due to the non-linear structure of the Lie group. The difference between these two approaches vanishes in the special case where G is a linear space.

The direct differentiation method leads to as many additional systems of linear DAE to be solved at each time step as design parameters, regardless the number of criterion functions. These systems of linear DAE are closely related to the linearized equations of the nominal problem. Moreover, even though the solution of the nominal problem may require several Newton iterations, only one iteration is necessary to solve the linear sensitivity problem at each time step.

In the direct differentiation of the equations of motion, the system of linear DAE for the sensitivities can be solved using any classical time integrator for linear spaces, but the involved matrices are affected by the Lie group structure. The formulation is independent of the particular Lie group time integrator used to solve the nominal problem.

When differentiating the Lie group time integration scheme, a discrete system of linear DAE is obtained and the iteration matrix for the sensitivity problem is the same as the iteration matrix for the nominal problem. This property tends to simplify the practical implementation of the algorithm in an existing simulation code, which is advantageous compared to the formulation based on the time integration of the sensitivity equations. Furthermore, using the same numerical parameters in the time integration formulae, the differentiation of the integration scheme allows to compute the sensitivity of the numerical solution exactly.

For both methods, the sensitivities may be computed on the fly, simultaneously with the mechanical system, using the following procedure at each time step. Firstly the nominal nonlinear problem is solved using the usual Newton procedure. Then after convergence the iteration matrix for the sensitivity problem, which is the same for each design parameter, is evaluated once. For each design parameter, the pseudo-loads need to be evaluated using either analytical developments or a finite difference approach. The sensitivities are then evaluated by solving one system of linear DAE for each pseudo-load. The factorization of the iteration matrix for the sensitivities should thus be performed only once at the current time step, whatever the number of design parameters and the number of criterion functions.

7 Adjoint variable method

In the adjoint variable method, the equations of motion are used as constraints on the criterion thanks to Lagrange multiplier variables, called adjoint variables. After some manipulations, a system of equations is obtained for the adjoint variables and the computation of the derivatives of the generalized coordinates with respect to the parameters is avoided. This method can be found in the literature, see e.g. [16]. However, to our knowledge, it has not yet been developed for Lie group formulation of multibody systems.

7.1 Preliminary developments

The variation of any function $f(q, \mathbf{v}, \dot{\mathbf{v}}, \boldsymbol{\lambda}, p)$ can be expressed as

$$\delta f(q, \mathbf{v}, \dot{\mathbf{v}}, \boldsymbol{\lambda}, p) = f_{\mathbf{q}}\delta\mathbf{q} + f_{\mathbf{v}}\delta\mathbf{v} + f_{\dot{\mathbf{v}}}\delta\dot{\mathbf{v}} + f_{\boldsymbol{\lambda}}\delta\boldsymbol{\lambda} + f_p\delta p$$

Using Eq. (13) and its derivative $\delta\dot{\mathbf{v}} = \delta\ddot{\mathbf{q}} + \hat{\mathbf{v}}\delta\dot{\mathbf{q}} + \hat{\dot{\mathbf{v}}}\delta\mathbf{q}$, the variations $\delta\mathbf{v}$ and $\delta\dot{\mathbf{v}}$ can be eliminated, so that

$$\delta f(q, \mathbf{v}, \dot{\mathbf{v}}, \boldsymbol{\lambda}, p) = f_{\mathbf{K}}\delta\mathbf{q} + f_{\mathbf{C}}\delta\dot{\mathbf{q}} + f_{\mathbf{M}}\delta\ddot{\mathbf{q}} + f_{\boldsymbol{\lambda}}\delta\boldsymbol{\lambda} + f_p\delta p \tag{34}$$

with the compact notations $f_{\mathbf{M}} = f_{\dot{\mathbf{v}}}$, $f_{\mathbf{C}} = f_{\mathbf{v}} + f_{\dot{\mathbf{v}}}\hat{\mathbf{v}}$ and $f_{\mathbf{K}} = f_{\mathbf{q}} + f_{\mathbf{v}}\hat{\mathbf{v}} + f_{\dot{\mathbf{v}}}\hat{\dot{\mathbf{v}}}$.

Using these definitions, the variation of the equations of motion reads

$$\begin{aligned}
\delta\mathbf{r}(q, \mathbf{v}, \dot{\mathbf{v}}, \boldsymbol{\lambda}, p) &= \mathbf{M}\delta\ddot{\mathbf{q}} + \mathbf{r}_{\mathbf{C}}\delta\dot{\mathbf{q}} + \mathbf{r}_{\mathbf{K}}\delta\mathbf{q} + \mathbf{B}^T\delta\boldsymbol{\lambda} + \mathbf{r}_p\delta p && (35)\\
\delta\boldsymbol{\Phi}(q, p) &= \mathbf{B}\delta\mathbf{q} + \boldsymbol{\Phi}_p\delta p && (36)\\
\delta\boldsymbol{\zeta}(q(t_0), p) &= \boldsymbol{\zeta}_{\mathbf{q}}\delta\mathbf{q}(t_0) + \boldsymbol{\zeta}_p\delta p && (37)\\
\delta\boldsymbol{\chi}(q(t_0), \mathbf{v}(t_0), p) &= \boldsymbol{\chi}_{\mathbf{v}}\delta\dot{\mathbf{q}}(t_0) + \boldsymbol{\chi}_{\mathbf{K}}\delta\mathbf{q}(t_0) + \boldsymbol{\chi}_p\delta p && (38)
\end{aligned}$$

with $\mathbf{r}_{\mathbf{C}} = \mathbf{C}_t + \mathbf{M}\hat{\mathbf{v}}$, $\mathbf{r}_{\mathbf{K}} = \mathbf{K}_t + \mathbf{C}_t\hat{\mathbf{v}} + \mathbf{M}\hat{\dot{\mathbf{v}}}$ and $\boldsymbol{\chi}_{\mathbf{K}} = \boldsymbol{\chi}_{\mathbf{q}} + \boldsymbol{\chi}_{\mathbf{v}}\hat{\mathbf{v}}$.

In general, the elimination of $\delta\mathbf{v}$ and $\delta\dot{\mathbf{v}}$ brings additional terms due to $\hat{\mathbf{v}}$ and $\hat{\dot{\mathbf{v}}}$. On a linear space, $\hat{\mathbf{v}}$ and $\hat{\dot{\mathbf{v}}}$ vanish.

7.2 Equations for the adjoint variables and sensitivity of a criterion

Using Eq. (34), the variation of the criterion in Eq. (31) reads

$$\begin{aligned}\delta\Psi &= \delta\Psi_G + \delta\Psi_F \\ &= G_{\mathbf{K}}\delta\mathbf{q}(t_f) + G_{\mathbf{C}}\delta\dot{\mathbf{q}}(t_f) + G_p\delta p \\ &\quad + \int_{t_0}^{t_f} (F_{\mathbf{K}}\delta\mathbf{q} + F_{\mathbf{C}}\delta\dot{\mathbf{q}} + F_{\mathbf{M}}\delta\ddot{\mathbf{q}} + F_{\boldsymbol{\lambda}}\delta\boldsymbol{\lambda} + F_p\delta p)\ dt\end{aligned}$$

Considering Eq. (16-19) as constraints for the problem, a set of Lagrange multipliers is introduced in order to eliminate the variations of $\delta\mathbf{q}$, its time derivatives and $\delta\boldsymbol{\lambda}$. The modified, nevertheless equivalent, criterion is

$$\delta\Psi = \delta\Psi + \boldsymbol{\pi}^T\delta\boldsymbol{\zeta} + \boldsymbol{\rho}^T\delta\boldsymbol{\chi} + \int_{t_0}^{t_f} (\boldsymbol{\mu}^T\delta\mathbf{r}(q, \mathbf{v}, \dot{\mathbf{v}}, \boldsymbol{\lambda}, t) + \boldsymbol{\nu}^T\delta\boldsymbol{\Phi}(q))\ dt$$

where $\boldsymbol{\pi}$, $\boldsymbol{\rho}$, $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ are the adjoint variables. Inserting the variations in Eq. (35-38) yields

$$\begin{aligned}\delta\Psi &= \delta\Psi_G + (\boldsymbol{\rho}^T\boldsymbol{\chi}_p + \boldsymbol{\pi}^T\boldsymbol{\zeta}_p)\delta p + \boldsymbol{\rho}^T\boldsymbol{\chi}_{\mathbf{v}}\delta\dot{\mathbf{q}}(t_0) + (\boldsymbol{\rho}^T\boldsymbol{\chi}_{\mathbf{K}} + \boldsymbol{\pi}^T\boldsymbol{\zeta}_{\mathbf{q}})\delta\mathbf{q}(t_0) \\ &\quad + \int_{t_0}^{t_f} \big[(F_{\mathbf{M}} + \boldsymbol{\mu}^T\mathbf{M})\delta\ddot{\mathbf{q}} + (F_{\mathbf{C}} + \boldsymbol{\mu}^T\mathbf{r}_{\mathbf{C}})\delta\dot{\mathbf{q}} + (F_{\mathbf{K}} + \boldsymbol{\mu}^T\mathbf{r}_{\mathbf{K}} + \boldsymbol{\nu}^T\mathbf{B})\delta\mathbf{q} \\ &\quad + (F_{\boldsymbol{\lambda}} + \boldsymbol{\mu}^T\mathbf{B}^T)\delta\boldsymbol{\lambda} + (F_p + \boldsymbol{\mu}^T\mathbf{r}_p)\delta p\big]\ dt\end{aligned}$$

Integrating by parts and gathering the terms with same variation leads to

$$\begin{aligned}\delta\Psi &= (G_p + \boldsymbol{\rho}^T\boldsymbol{\chi}_p + \boldsymbol{\pi}^T\boldsymbol{\zeta}_p)\delta p + \int_{t_0}^{t_f} (F_p + \boldsymbol{\mu}^T\mathbf{r}_p + \boldsymbol{\nu}^T\boldsymbol{\Phi}_p)\delta p\ dt \\ &\quad + (F_{\mathbf{M}} + \boldsymbol{\mu}^T\mathbf{M} + G_{\mathbf{C}})|_{t_f}\delta\dot{\mathbf{q}}(t_f) + [F_{\mathbf{C}} + \boldsymbol{\mu}^T\mathbf{r}_{\mathbf{C}} - \frac{d}{dt}(F_{\mathbf{M}} + \boldsymbol{\mu}^T\mathbf{M}) + G_{\mathbf{K}}]|_{t_f}\delta\mathbf{q}(t_f) \\ &\quad + [\boldsymbol{\rho}^T\boldsymbol{\chi}_{\mathbf{v}} - (F_{\mathbf{M}} + \boldsymbol{\mu}^T\mathbf{M})]|_{t_0}\delta\dot{\mathbf{q}}(t_0) + [\boldsymbol{\rho}^T\boldsymbol{\chi}_{\mathbf{K}} + \boldsymbol{\pi}^T\boldsymbol{\zeta}_{\mathbf{q}} - (F_{\mathbf{C}} + \boldsymbol{\mu}^T\mathbf{r}_{\mathbf{C}}) \\ &\quad + \frac{d}{dt}(F_{\mathbf{M}} + \boldsymbol{\mu}^T\mathbf{M})]|_{t_0}\delta\mathbf{q}(t_0) + \int_{t_0}^{t_f} (F_{\boldsymbol{\lambda}} + \boldsymbol{\mu}^T\mathbf{B}^T)\delta\boldsymbol{\lambda}\ dt \\ &\quad + \int_{t_0}^{t_f} [\frac{d^2}{dt^2}(F_{\mathbf{M}} + \boldsymbol{\mu}^T\mathbf{M}) - \frac{d}{dt}(F_{\mathbf{C}} + \boldsymbol{\mu}^T\mathbf{r}_{\mathbf{C}}) + F_{\mathbf{K}} + \boldsymbol{\mu}^T\mathbf{r}_{\mathbf{K}} + \boldsymbol{\nu}^T\mathbf{B}]\delta\mathbf{q}\ dt\end{aligned}$$

The adjoint variables are then defined such that only the variation of the parameter remains. Three sets of equations are thus obtained.

- An adjoint system of second-order linear DAE is obtained for $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$:

$$\mathbf{M}\ddot{\boldsymbol{\mu}} - \mathbf{r}_{\mathbf{C}}^T\dot{\boldsymbol{\mu}} + (\mathbf{r}_{\mathbf{K}} - \frac{d}{dt}\mathbf{r}_{\mathbf{C}})^T\boldsymbol{\mu} + \mathbf{B}^T\boldsymbol{\nu} = (-\frac{d^2}{dt^2}F_{\mathbf{M}} + \frac{d}{dt}F_{\mathbf{C}} - F_{\mathbf{K}})^T \tag{39}$$

$$\mathbf{B}\boldsymbol{\mu} = -F_{\boldsymbol{\lambda}}^T \tag{40}$$

  Developing more leads to

$$\begin{aligned}\mathbf{M}\ddot{\boldsymbol{\mu}} - (\mathbf{C}_t + \mathbf{M}\hat{\mathbf{v}})^T\dot{\boldsymbol{\mu}} + (\mathbf{K}_t + \mathbf{C}_t\hat{\mathbf{v}} - \dot{\mathbf{C}}_t)^T\boldsymbol{\mu} + \mathbf{B}^T\boldsymbol{\nu} &= (-\frac{d^2}{dt^2}F_{\mathbf{M}} + \frac{d}{dt}F_{\mathbf{C}} - F_{\mathbf{K}})^T \\ \mathbf{B}\boldsymbol{\mu} &= -F_{\boldsymbol{\lambda}}^T\end{aligned}$$

  where the mass matrix $\mathbf{M}$ has been assumed to be constant. Considering the final time conditions discussed next, this system can be solved backwards in time for $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$. The adjoint equations at time step $n+1$ can be formulated in terms of a specific $(k+m)\times(k+m)$ adjoint matrix but, unfortunately, this matrix is not equal to the transpose of the iteration matrix defined in Eq. (30) and used to solve the nominal problem. The Lie bracket operator appears in these equations through the contribution $\hat{\mathbf{v}}$. In the special case where G is a linear space, these contributions vanish but the time derivative of the damping matrix is still involved.

If the criterion is a non-linear function of the accelerations $\dot{\mathbf{v}}$, the first and second derivatives of the accelerations are needed since $F_{\mathbf{C}}$ and $F_{\mathbf{M}}$ depend on $\dot{\mathbf{v}}$. However, the third and fourth derivatives of the motion are usually not computed by time integrators. Several options can be investigated. Firstly additional differential equations $\mathbf{z} = \dot{\mathbf{y}}$ and $\mathbf{y} = \dot{\mathbf{v}}$ can be used to computed the further derivatives. The computation of $\mathbf{z}(t_0)$ and $\mathbf{y}(t_0)$ can be done by finite differentiation of the known accelerations at t_0. It can be easily implemented, and is not expensive if only few accelerations are involved in the criterion. However, higher derivatives can have a high frequency content and large amplitudes, which may lead to some numerical difficulties and deteriorate the expected properties. Secondly another approach could be to express the accelerations and its variations as functions of the state variables, i.e. the positions and the velocities, by solving

$$\begin{bmatrix} \dot{\mathbf{v}} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{M} & \mathbf{B}^T \\ \mathbf{B} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} -\mathbf{f}(q, \mathbf{v}, p) \\ -\dot{\mathbf{B}}\mathbf{v} \end{bmatrix} \tag{41}$$

The criterion can then be expressed as a function of the state variables only. Unfortunately, Eq. (41) involve the whole system and the time derivative of the constraint gradient $\mathbf{B}$. This approach may thus lead in general to complicated and computationally expensive expressions. Thirdly the adjoint variable method presented in [16] could be extended to the Lie group formalism. In this case, no time derivative of the criterion is needed at the cost of a larger number adjoint variables, and it leads to a system of first-order linear DAE, whose time integration must be carried out carefully. A detailed comparison of these three strategies require more investigation and is a subject of future research.

- Final time conditions are obtained for $\boldsymbol{\mu}$ and $\dot{\boldsymbol{\mu}}$:

$$\mathbf{M}\boldsymbol{\mu}(t_f) = -(F_{\mathbf{M}} + G_{\mathbf{C}})|_{t_f}^T \tag{42}$$

$$\mathbf{M}\dot{\boldsymbol{\mu}}(t_f) = (F_{\mathbf{C}} + \boldsymbol{\mu}^T \mathbf{r}_{\mathbf{C}} - \frac{d}{dt} F_{\mathbf{M}} + G_{\mathbf{K}})|_{t_f}^T \tag{43}$$

This allows to eliminate the variations at the final time. In many cases, the mass matrix is not invertible and the constraint equations must be considered. Indeed, $\boldsymbol{\mu}(t_f)$ and $\dot{\boldsymbol{\mu}}(t_f)$ must satisfy the constraint equations (40) and their time derivatives. Therefore, based on [24], we have

$$\boldsymbol{\mu}(t_f) = \mathbf{A} \begin{bmatrix} -(F_{\mathbf{M}} + G_{\mathbf{C}})^T \\ -F_{\boldsymbol{\lambda}}^T \end{bmatrix}_{t_f} \quad ; \quad \dot{\boldsymbol{\mu}}(t_f) = \mathbf{A} \begin{bmatrix} (F_{\mathbf{C}} + \boldsymbol{\mu}^T \mathbf{r}_{\mathbf{C}} - \frac{d}{dt} F_{\mathbf{M}} + G_{\mathbf{K}})^T \\ -\dot{\mathbf{B}}\boldsymbol{\mu} - \dot{F}_{\boldsymbol{\lambda}}^T \end{bmatrix}_{t_f}$$

in which, defining $\bullet^+$ as the pseudo-inverse,

$$\mathbf{A} = \begin{bmatrix} (\mathbf{I} - \mathbf{B}^+\mathbf{B})\mathbf{M} \\ \mathbf{B} \end{bmatrix}_{t_f}^+$$

We observe that in order to compute $\dot{\boldsymbol{\mu}}(t_f)$ the time derivative of the constraint gradient at the final time step is required, as well as the time derivative of $F_{\boldsymbol{\lambda}}$.

- The variables $\boldsymbol{\rho}$ and $\boldsymbol{\pi}$ allow to deal with the initial conditions of the equations of motion. Once the adjoint system is solved, $\boldsymbol{\rho}$ and $\boldsymbol{\pi}$ are computed such that the variation of the initial conditions drops:

$$\begin{aligned}\boldsymbol{\chi}_{\mathbf{v}}^T\boldsymbol{\rho} &= (F_{\mathbf{M}} + \boldsymbol{\mu}^T\mathbf{M})|_{t_0}^T \\ \boldsymbol{\zeta}_{\mathbf{q}}^T\boldsymbol{\pi} &= (F_{\mathbf{C}} + \boldsymbol{\mu}^T\mathbf{r}_{\mathbf{C}} - \frac{d}{dt}F_{\mathbf{M}} - \dot{\boldsymbol{\mu}}^T\mathbf{M} - \boldsymbol{\rho}^T\boldsymbol{\chi}_{\mathbf{K}})|_{t_0}^T\end{aligned}$$

  If the initial conditions do not depend explicitly on the parameters, that is $\boldsymbol{\zeta}_p$ and $\boldsymbol{\chi}_p$ vanish, there is no need to compute $\boldsymbol{\rho}$ and $\boldsymbol{\pi}$.

Eventually, the gradient of the criterion is computed using

$$\frac{d\Psi}{dp} = G_p + \boldsymbol{\rho}^T\boldsymbol{\chi}_p + \boldsymbol{\pi}^T\boldsymbol{\zeta}_p + \int_{t_0}^{t_f} (F_p + \boldsymbol{\mu}^T\mathbf{r}_p + \boldsymbol{\nu}^T\boldsymbol{\Phi}_p)\ dt$$

Besides the ajoint variables, it only involves partial derivatives with respect to the parameters.

7.3 Discussion

The adjoint variable method leads to as many additional systems of linear DAE to be integrated backward in time as criterion functions, regardless the number of parameters. These systems of linear DAE are closely related to the linearized equations of the nominal problem and can be solved using any classical time integrator for linear spaces. Notice that the terms involved are affected by the Lie group structure and require some additional information compared to the nominal problem. Even though the solution of the nominal problem may require several Newton iterations, only one iteration is necessary to solve the linear adjoint problem at each time step.

Since the time integration needs to be performed backward in time, it is not possible to compute the sensitivities on the fly when the adjoint variable method is used. One needs to perform the nominal simulation from the initial to the final time in a first step, then to compute the sensitivities by backward integration in a second step. At each time step of the backward time integration, the procedure is as follows. Firstly the adjoint iteration matrix, which is the same for all criterion functions, is evaluated. Secondly the derivatives of each criterion function with respect to the generalized coordinates are evaluated using either an analytical formula or a finite difference approach. Thirdly the adjoint variables are computed by solving one system of linear DAE for each criterion function. The factorization of the adjoint iteration matrix should thus be performed only once at the current time step, whatever the number of criterion functions or the number of design parameters.

The presented method yields index-3 adjoint DAE and according to [25] the stability is guaranteed on a linear space. The adjoint DAE has the same second-order structure as the nominal problem and can be solved using the generalized-α integration formulae for linear spaces. There are alternative ways to introduce the adjoint variables and they could also be extended to the Lie group formalism. For instance in Ref. [15], the method requires to introduce some additional adjoint variables to deal with the kinematic compatibility equation and yields a system of first-order linear DAE. In this case, no time derivative of the functional F is involved in the pseudo-load terms but a suitable time integration for first-order systems must be selected carefully [26].

8 Example: Spinning top

The methods are now illustrated for the academic example of a spinning top, as shown in Fig. 1, formulated on the Lie group $\mathbb{R}^3 \times SO(3)$. The equations of motion were given in section 3.2.

In this example, we consider that the initial position vector $\mathbf{X}$ depends on a design parameter α_0 as follows

$$\mathbf{X}(\alpha_0) = l\mathbf{e}(\alpha_0) = l[0\ \cos(\alpha_0)\ \sin(\alpha_0)]^T$$

where l represents the distance between the origin and the center of mass C, $\mathbf{e}$ is the unit vector pointing towards the center of mass and α_0 represents an angle in the yz-plane. The spinning top has an initial rotational velocity ω_0 about $\mathbf{e}$. The initial conditions satisfy

$$\boldsymbol{\zeta}(\mathbf{x}_0, \mathbf{R}_0) : (\mathbf{I}_{3\times 3}, \mathbf{X}(\alpha_0))^{-1}(\mathbf{R}_0, \mathbf{x}_0) = e \Leftrightarrow \left\{ \begin{array}{ccc} \mathbf{x}_0 - \mathbf{X}(\alpha_0) & = & \mathbf{0}_{3\times 1} \\ \mathbf{R}_0 & = & \mathbf{I}_{3\times 3} \end{array} \right.$$

$$\boldsymbol{\chi}(\dot{\mathbf{x}}_0, \boldsymbol{\Omega}_0) : \left\{ \begin{array}{ccc} \dot{\mathbf{x}}_0 + \mathbf{R}_0\widetilde{\mathbf{X}}(\alpha_0)\boldsymbol{\Omega}_0 & = & \mathbf{0}_{3\times 1} \\ \boldsymbol{\Omega}_0 - \omega_0\mathbf{e}(\alpha_0) & = & \mathbf{0}_{3\times 1} \end{array} \right.$$

The criterion is the integral of the square of the displacement along the x-axis

$$\Psi = \int_{t_0}^{t_f} x^2(t)\ dt$$

The proposed methods are applied to evaluate the sensitivity of this criterion with respect to the angle α_0.

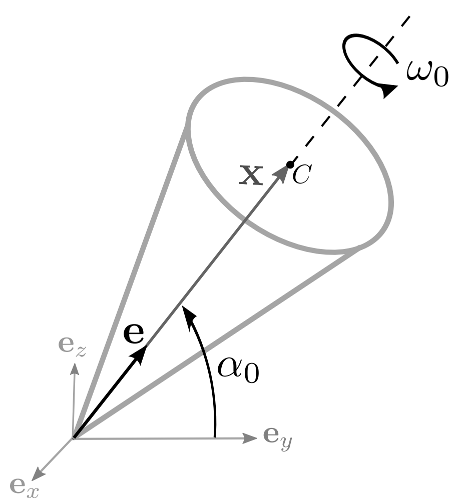

Figure 1: Spinning top example. The initial conditions are described using $\mathbf{e}$, α_0 and ω_0.

8.1 Pseudo-loads and common derivatives

The partial derivatives with respect to the parameters and the derivatives with respect to the generalized coordinates are involved in both the direct differentiation method and the adjoint variable method. Compared to (31), the criterion involves only $F(q)$. All derivatives vanish but

$$F_{\mathbf{q}} = \begin{bmatrix} 2x(t) & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and we have $F_{\mathbf{K}} = F_{\mathbf{q}}$ and $F_{\mathbf{C}} = F_{\mathbf{M}} = \mathbf{0}$.

The pseudo-loads are

$$\mathbf{r}_p = \begin{bmatrix} \mathbf{0}_{3\times 1} \\ -l\widetilde{\mathbf{e}'}\mathbf{R}^T\boldsymbol{\lambda} \end{bmatrix} \quad ; \quad \boldsymbol{\Phi}_p = -l\mathbf{R}\mathbf{e}'$$

with $\mathbf{e}' = [0 \;\; -\sin(\alpha_0) \;\; \cos(\alpha_0)]^T$. Regarding the initial conditions, we have

$$\boldsymbol{\chi}_{\mathbf{v}} = \boldsymbol{\zeta}_{\mathbf{q}} = \begin{bmatrix} \mathbf{I}_{3\times 3} & \mathbf{R}_0\widetilde{\mathbf{X}} \\ \mathbf{0}_{3\times 3} & \mathbf{I}_{3\times 3} \end{bmatrix} \quad ; \quad \boldsymbol{\zeta}_p = \begin{bmatrix} \mathbf{0}_{3\times 1} \\ -l\mathbf{R}_0\mathbf{e}' \end{bmatrix}$$

$$\boldsymbol{\chi}_{\mathbf{q}} = \begin{bmatrix} \mathbf{0}_{3\times 3} & -\mathbf{R}_0\widetilde{\widetilde{\mathbf{X}}\boldsymbol{\Omega}_0} \\ \mathbf{0}_{3\times 3} & \mathbf{0}_{3\times 3} \end{bmatrix} \quad ; \quad \boldsymbol{\chi}_p = \begin{bmatrix} l\mathbf{R}_0\widetilde{\mathbf{e}'}\boldsymbol{\Omega}_0 \\ -\omega_0\mathbf{e}' \end{bmatrix}$$

8.2 Direct differentiation method

The direct differentiation of the time integration method is presented. The matrix $\mathbf{D}$ introduced in section 6.2 is given here by

$$\mathbf{D} = \begin{bmatrix} \mathbf{I}_{3\times 3} & \mathbf{0}_{3\times 3} \\ \mathbf{0}_{3\times 3} & \exp_{S0(3)}(\widetilde{\mathbf{x}}_{n+1}) \end{bmatrix}^T$$

where $\exp_{S0(3)}$ is the exponential map of $SO(3)$, which is simply given by the Rodrigues formula. For readability, we write $\mathbf{w} = [\mathbf{x}' \;\; \mathbf{w}_\Omega]^T$, and the system of equations for the sensitivities is obtained as

$$\begin{aligned} m\ddot{\mathbf{x}}' + \boldsymbol{\lambda}' &= \mathbf{0} \\ \mathbf{J}\dot{\boldsymbol{\Omega}}' + (\widetilde{\boldsymbol{\Omega}}\mathbf{J} - \widetilde{\mathbf{J}\boldsymbol{\Omega}})\boldsymbol{\Omega}' - \widetilde{\mathbf{X}}\widetilde{\mathbf{R}^T\boldsymbol{\lambda}}\mathbf{w}_\Omega - \widetilde{\mathbf{X}}\mathbf{R}^T\boldsymbol{\lambda}' &= l\widetilde{\mathbf{e}'}\mathbf{R}^T\boldsymbol{\lambda} \\ \mathbf{x}' + \mathbf{R}\widetilde{\mathbf{X}}\mathbf{w}_\Omega &= l\mathbf{R}\mathbf{e}' \end{aligned}$$

where all terms are evaluated at a time t_{n+1}. The initial conditions are computed as $\mathbf{w}_0 = -\boldsymbol{\zeta}_{\mathbf{q}}^{-1}\boldsymbol{\zeta}_p$ and $\mathbf{v}'_0 = -\boldsymbol{\chi}_{\mathbf{v}}^{-1}(\boldsymbol{\chi}_{\mathbf{q}}\mathbf{w}_0 + \boldsymbol{\chi}_p)$. Finally, the sensitivity of the criterion is

$$\frac{d\Psi}{d\alpha_0} = \int_{t_0}^{t_f} 2x(t)x'(t)\; dt$$

8.3 Adjoint variable method

The final conditions are trivially satisfied:

$$\begin{aligned} \boldsymbol{\mu}(t_t) &= \mathbf{0} \\ \dot{\boldsymbol{\mu}}(t_f) &= \mathbf{0} \end{aligned}$$

For readability, we write $\boldsymbol{\mu}^T = [\boldsymbol{\mu}_q^T \;\; \boldsymbol{\mu}_\Omega^T]$ and $\boldsymbol{\nu}^T = [\boldsymbol{\nu}_q^T \;\; \boldsymbol{\nu}_\Omega^T]$, and we obtain the system of equations for the adjoint variables

$$m\ddot{\boldsymbol{\mu}}_q + \boldsymbol{\nu}_q = -\begin{bmatrix} 2x(t) & 0 & 0 \end{bmatrix}^T \tag{44}$$

$$\mathbf{J}\ddot{\boldsymbol{\mu}}_\Omega - \mathbf{C}_{\boldsymbol{\mu}_\Omega}\dot{\boldsymbol{\mu}}_\Omega + \mathbf{K}_{\boldsymbol{\mu}_\Omega}\boldsymbol{\mu}_\Omega - \widetilde{\mathbf{X}}\mathbf{R}^T\boldsymbol{\nu}_\Omega = \mathbf{0} \tag{45}$$

$$\boldsymbol{\mu}_q + \mathbf{R}\widetilde{\mathbf{X}}\boldsymbol{\mu}_\Omega = \mathbf{0} \tag{46}$$

where

$$\begin{aligned} \mathbf{C}_{\boldsymbol{\mu}_\Omega} &= [\mathbf{J}\widetilde{\boldsymbol{\Omega}} + (\widetilde{\boldsymbol{\Omega}}\mathbf{J} - \widetilde{\mathbf{J}\boldsymbol{\Omega}})]^T \\ \mathbf{K}_{\boldsymbol{\mu}_\Omega} &= [(\widetilde{\mathbf{J}\dot{\boldsymbol{\Omega}}} - \dot{\widetilde{\boldsymbol{\Omega}}}\mathbf{J}) + (\widetilde{\boldsymbol{\Omega}}\mathbf{J} - \widetilde{\mathbf{J}\boldsymbol{\Omega}})\widetilde{\boldsymbol{\Omega}} - \widetilde{\mathbf{X}}\widetilde{\mathbf{R}^T\boldsymbol{\lambda}}]^T \end{aligned}$$

The adjoint variables related to the variations of the initial conditions are computed as $\boldsymbol{\rho} = \boldsymbol{\chi}_{\mathbf{v}}^{-T}\mathbf{M}\boldsymbol{\mu}$ and $\boldsymbol{\pi} = \boldsymbol{\zeta}_{\mathbf{q}}^{-T}(\mathbf{r}_c^T\boldsymbol{\mu} - \mathbf{M}\dot{\boldsymbol{\mu}} - (\boldsymbol{\chi}_{\mathbf{v}}\widehat{\mathbf{v}} + \boldsymbol{\chi}_{\mathbf{q}})^T\boldsymbol{\rho})$, where all terms are evaluated at t_0. Finally, the sensitivity of the criterion is

$$\frac{d\Psi}{d\alpha_0} = \boldsymbol{\rho}^T\boldsymbol{\chi}_p + \boldsymbol{\pi}^T\boldsymbol{\zeta}_p + \int_{t_0}^{t_f} (\boldsymbol{\mu}^T\mathbf{r}_p + \boldsymbol{\nu}^T\boldsymbol{\Phi}_p)\; dt$$

8.4 Numerical results

The properties of the spinning top are $m = 15$ kg, $J = \text{diag}(\begin{bmatrix}0.2 & 0.4 & 0.2\end{bmatrix})$ kgm^2, $l = 0.5$ m, $\alpha_0 = \pi/4$ rad and $\omega_0 = 10$ rad/s. The gravity field is $\mathbf{g} = \begin{bmatrix}0 & 0 & -9.81\end{bmatrix}$ m/s^2. The simulation runs for 1 second with the spectral radius $\rho = 0.9$. The reference value of the sensitivity is -0.002131825283595, which was computed using a fourth-order finite difference method with the aforementioned setting and a time step size of 10^{-5} s. We observe for both the direct differentiation method and the adjoint variable method the second-order of convergence of the sensitivities, i.e. the same order as for the nominal problem, and a high accuracy. The results are shown in Fig. 2. Although there is no formal proof that the convergence analysis applied to the nominal system also applies to the sensitivity problem and to the adjoint problem, the same order of convergence is observed in this example where the same time integration algorithm was used.

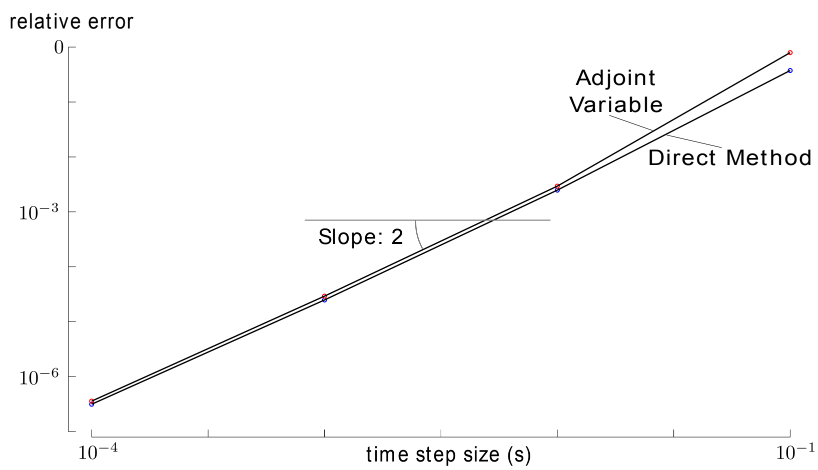

Figure 2: Second-order convergence of the relative error in the sensitivities with respect to the time step size.

9 Application: Car suspension

Here the methods are presented in the context of the parameter optimization of a quarter-car suspension. The model, shown in Fig. 3, relies on usual elements in multibody system modelling. Indeed, the chassis, the five bars, the knuckle and the wheel are considered as rigid bodies, while the tire and the suspension are modelled as springs. Regarding kinematic joints, the wheel is attached to the knuckle by a horizontal hinge joint and universal joints link the bars to the chassis, as well as the bars to the knuckle. The boundary conditions on the chassis allow its vertical displacement only. The system is described in a finite element approach (see [6]) and it has 30 nodes, 105 free nodal coordinates and 103 constraints. The two degrees of freedom are the vertical displacement of the chassis and the motion of the wheel with respect to the chassis.

In the case study, the car is passing over a speed bump, which is modelled as an imposed vertical motion of the wheel ground contact point. The objective of the optimization is to increase the comfort of the passengers by reducing the vertical accelerations of the chassis:

$$\Psi_0 = \int_{t_0}^{t_f} \dot{v}^2_{z,chassis}(t)\ dt \tag{47}$$

In order to guarantee road handling performances, the normal force F_{t-r}, exerted by the tire on the road, should always be above a lower threshold F_{min}:

$$\Psi_1 = F_{min} - F_{t-r}(t) \leq 0 \tag{48}$$

Finally, for feasibility reasons, the suspension length l must remain in-between two bounds l_{min} and l_{max}:

$$\Psi_2 = l_{min} - l(t) \leq 0 \tag{49}$$

$$\Psi_3 = l(t) - l_{max} \leq 0 \tag{50}$$

Notice that, after time discretization, Ψ_1, Ψ_2 and Ψ_3 actually induce as many criterion functions as time steps since they must be checked at each time of the simulation. However, in practice, these constraints are not all active at the same time.

The design parameters are the coefficients of the spring used to model the suspension, namely the stiffness k and the damping c.

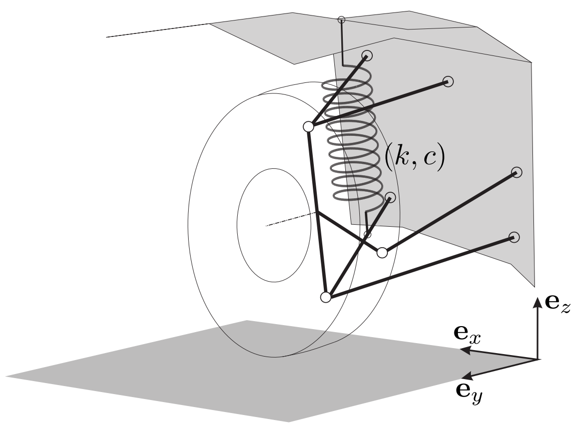

Figure 3: Quarter-car model. The parameters of the suspension are k and c, respectively the stiffness coefficient and the damping coefficient.

The choice of the method for the sensitivity analysis depends on the number of active constraints. If none of the constraints are active, the adjoint variable method is more attractive since only one additional system has to be solved to compute the sensitivities with respect to both parameters. However, when some constraints are activated, the direct differentiation method is more suitable because it leads to two additional systems of equations, regardless the number of criteria.

Some numerical results are shown in Fig. 4. In Fig. 4(a), we observe second-order convergence for both methods, and similar accuracy. In Fig. 4(b), we observe that the sensitivity computed with the adjoint variable method for the objective function has a slightly lower accuracy. This seems to be due to the non-linear dependence of the objective function on the acceleration, that has be treated with additional differential equations (see discussion in section 7.2).

10 Conclusions

This paper studies semi-analytical methods for efficient sensitivity analysis in large scale optimization problems. The main contribution is the development of a direct differentiation method

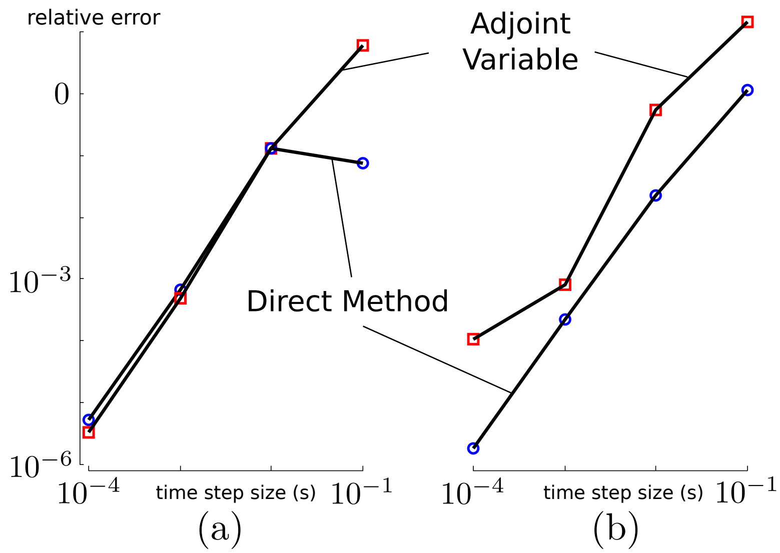

Figure 4: Convergence results for the sensitivity of (a) Constraint function Ψ_3 with respect to the damping coefficient c and (b) Objective function Ψ_0 with respect to the stiffness coefficient k

and an adjoint variable method in the context of a Lie group formulation in multibody dynamics.

The development and implementation of those methods require the complete linearization of the equations of motion and/or of the time integration procedure. The simpler structure of the equations of motion in the Lie group formalism turns out to be an advantage for that purpose. Lie bracket contributions and the non-linearity of the exponential map have to be explicitly taken into account in the sensitivity algorithms. In the special case where the Lie group is a vector space, these contributions vanish and classical algorithms are obtained.

Two direct differentiation methods are developed based on the time integration of the differentiated equations of motion and on the differentiation of the time integration algorithm. Due to the non-linearity of the Lie group, these two approaches lead to two different algorithms. In the second approach, the iteration matrix of the sensitivity problem and of the nominal problem are rigorously the same, which is an advantage. In both cases, the sensitivity problem leads to linear differential-algebraic equations and is solved by forward time integration with as many additional load cases as design parameters. In contrast, the adjoint variable method leads to linear differential-algebraic equations which must be solved backward in time with as many additional load cases as criterion functions. For each method, the essential features are similar to those obtained for linear spaces. Therefore, the existing work done on the implementation of these methods can largely be applied.

The properties of the numerical algorithms are illustrated with two examples: a spinning top and a vehicle suspension. The convergence analysis with respect to the time step size shows that the sensitivities converge with the same order of accuracy as the solution of the nominal problem, although no formal proof of this result is given. These methods thus have a high potential for the efficient and accurate evaluation of sensitivities in full-scale applications of multibody dynamics.

Acknowledgements

Valentin Sonneville would like to acknowledge the Belgian National Fund for Scientific research (FRIA) for its financial support. The authors also thank Martin Arnold for useful suggestions about the treatment of an acceleration-dependent functional using the adjoint variable method.

References

[1] O. Brüls, E. Lemaire, P. Duysinx, and P. Eberhard. Optimization of multibody systems and their structural components. In W. Blajer, J. Arczewski, K. Fraczek, and M. Wojtyra, editors, *Multibody Dynamics: Computational Methods and Applications, Computational Methods in Applied Sciences*, volume 23, pages 49–68. Springer, 2011.

[2] C.L. Bottasso, A. Croce, L. Ghezzi, and P. Faure. On the solution of inverse dynamics and trajectory optimization problems for multibody systems. *Multibody System Dynamics*, 11:1–22, 2004.

[3] P. Eberhard and C. Bischof. Automatic differentiation of numerical integration algorithms. *Mathematics of Computation*, 68:717–731, 1999.

[4] M.A. Neto, J.A.C. Ambrósio, and R.P. Leal. Sensitivity analysis of flexible multibody systems using composite materials components. *International Journal for Numerical Methods in Engineering*, 77:386–413, 2009.

[5] J.A.C. Ambrósio, M.A. Neto, and R.P. Leal. Optimization of a complex flexible multibody systems with composite materials. *Multibody System Dynamics*, 18:117–144, 2007.

[6] M. Géradin and A. Cardona. *Flexible Multibody Dynamics: A Finite Element Approach*. John Wiley & Sons, Chichester, 2001.

[7] O. Brüls and A. Cardona. On the use of Lie group time integrators in multibody dynamics. *ASME Journal of Computational and Nonlinear Dynamics*, 5(3):031002, 2010.

[8] O. Brüls, A. Cardona, and M. Arnold. Lie group generalized-α time integration of constrained flexible multibody systems. *Mechanism and Machine Theory*, 48:121–137, 2012.

[9] O. Brüls and P. Eberhard. Sensitivity analysis for dynamic mechanical systems with finite rotations. *International Journal for Numerical Methods in Engineering*, 74(13):1897–1927, 2008.

[10] R.T. Haftka and H.M. Adelman. Recent developments in structural sensitivity analysis. *Structural Optimization*, 1:137–151, 1989.

[11] E.J. Haug, K.K. Choi, and V. Komkov. *Design Sensitivity Analysis of Structural Systems*. Academic Press, Orlando, 1986.

[12] P. Kowalczyk. Design sensitivity analysis in large deformation elasto-plastic and elasto-viscoplastic problems. *International Journal for Numerical Methods in Engineering*, 66:1234–1270, 2006.

[13] P. Michaleris, D.A. Tortorelli, and C.A Vidal. Tangent operators and design sensitivity formulations for transient non-linear coupled problems with applications to elastoplasticity. *International Journal for Numerical Methods in Engineering*, 37:2471–2499, 1994.

[14] D.A. Tortorelli. Sensitivity analysis for non-linear constrained elastoplastic systems. *International Journal for Numerical Methods in Engineering*, 33:1643–1660, 1992.

[15] D. Bestle and P. Eberhard. Analyzing and optimizing multibody systems. *Mechanics of Structures and Machines*, 20:67–92, 1992.

[16] D. Bestle and J. Seybold. Sensitivity analysis of constrained multibody systems. *Archive of Applied Mechanics*, 62:181–190, 1992.

[17] J.M.P. Dias and M.S. Pereira. Sensitivity analysis of rigid-flexible multibody systems. *Multibody System Dynamics*, 1(3):303–322, 1997.

[18] T.M. Wasfy and K. Noor. Modeling and sensitivity analysis of multibody systems using new solid, shell and beam elements. *Computer Methods in Applied Mechanics and Engineering*, 138:187–211, 1996.

[19] D.D. Holm, T. Schmah, C. Stoica, and D.C.P. Ellis. *Geometric Mechanics and Symmetry: from Finite to Infinite Dimensions.* Oxford texts in applied and engineering mathematics. Oxford University Press, 2009.

[20] W.M. Boothby. *An Introduction to Differentiable Manifolds and Riemannian Geometry.* Academic Press, 2nd edition, 2003.

[21] E. Paraskevopoulos and S. Natsiavas. A new look into the kinematics and dynamics of finite rigid body rotations using lie group theory. *International Journal of Solids and Structures*, 50:57–72, 2013.

[22] O. Brüls, M. Arnold, and A. Cardona. Two Lie group formulations for dynamic multibody systems with large rotations. In *Proceedings of the IDETC/MSNDC Conference*, Washington D.C., U.S., August 2011.

[23] M.B. Nielsen and S. Krenk. Conservative integration of rigid body motion by quaternion parameters with implicit constraints. *International Journal for Numerical Methods in Engineering*, 92:734–752, 2012.

[24] F.E. Udwadia and P. Phohomsiri. Explicit equations of motion for constrained mechanical systems with singular mass matrices and applications to multi-body dynamics. *Proceedings of the Royal Society of London, Series A*, 462:2097–2117, 2006.

[25] A.S. Schaffer. *On the Adjoint Formulation of Design Sensitivity Analysis of Multibody Dynamics.* PhD thesis, The University of Iowa, 2005.

[26] Y. Cao, S. Li, L. Petzold, and R. Serban. Adjoint sensitivity analysis for differential-algebraic equations: The adjoint dae system and its numerical solution. *SIAM Journal on Scientific Computing*, 24(3):1076–1089, 2003.