Dynamic Response Optimization of Complex Multibody Systems in a Penalty Formulation using Adjoint Sensitivity

Yitao Zhu  
Daniel Dopico  
Advanced Vehicle Dynamics Laboratory and Computational 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

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

Adrian Sandu  
Computational Science Laboratory  
Department of Computer Science  
Virginia Tech  
Blacksburg, VA 24061, USA  
Email: sandu@cs.vt.edu

Multibody dynamics simulations are currently widely accepted as valuable means for dynamic performance analysis of mechanical systems. The evolution of theoretical and computational aspects of the multibody dynamics discipline make it conducive these days for other types of applications, in addition to pure simulations. One very important such application is design optimization. A very important first step towards design optimization is sensitivity analysis of multibody system dynamics. Dynamic sensitivities are often calculated by means of finite differences. Depending of the number of parameters involved, this procedure can be computationally expensive. Moreover, in many cases, the results suffer from low accuracy when real perturbations are used. The main contribution to the state-of-the-art brought by this study is the development of the adjoint sensitivity approach of multibody systems in the context of the penalty formulation. The theory developed is demonstrated on one academic case study, a five-bar mechanism, and on one real-life system, a 14-DOF vehicle model. The five-bar mechanism is used to illustrate the sensitivity approach derived in this paper. The full vehicle model is used to demonstrate the capability of the new approach developed to perform sensitivity analysis and gradient-based optimization for large and complex multibody systems with respect to multiple design parameters.

Nomenclature

DOF Degree or degrees of a freedom.  
DAE Differential algebraic equations.  
ODE Ordinary differential equations.  
CG Center of gravity.  
t Time.  
\( (\cdots) \)\(_0\) Means evaluation at the initial time \((\cdots)(t_0)\).  
\( (\cdots) \)\(_F\) Means evaluation at the final time \((\cdots)(t_F)\).  
\( \mathbf{q} \in \mathbb{R}^n \) Vector of coordinates of the system.  
\( \mathbf{\rho} \in \mathbb{R}^p \) Vector of parameters.  
\( \mathbf{M}(\mathbf{q},\mathbf{\rho}) \in \mathbb{R}^{n \times n} \) Generalized mass matrix of the system.  
\( \mathbf{Q}(\mathbf{q},\dot{\mathbf{q}},\mathbf{\rho}) \in \mathbb{R}^n \) Vector of generalized forces of the system.  
\( \Phi(\mathbf{q},\mathbf{\rho}) \in \mathbb{R}^m \) Vector of constraints that relate the dependent coordinates.  
\( \mathbf{A}_\times \in \mathbb{R}^{q \times r \times s} \) Third order tensor of derivatives of matrix \( \mathbf{A} \in \mathbb{R}^{q \times r} \) w.r.t. vector \( \mathbf{x} \in \mathbb{R}^s \).  
\( \mathbf{A}_\times^T = \left[ \frac{\partial \mathbf{A}}{\partial x_{i_1}} \ldots \frac{\partial \mathbf{A}}{\partial x_{i_l}} \ldots \frac{\partial \mathbf{A}}{\partial x_{i_s}} \right] \in \mathbb{R}^{r \times q \times s} \).  
\( \mathbf{A}_\times \mathbf{b} = \mathbf{A}_\times \otimes \mathbf{b} = \left[ \frac{\partial \mathbf{A}}{\partial x_{i_1}} \mathbf{b} \ldots \frac{\partial \mathbf{A}}{\partial x_{i_l}} \mathbf{b} \ldots \frac{\partial \mathbf{A}}{\partial x_{i_s}} \mathbf{b} \right] \in \mathbb{R}^{q \times s} \),  
where \( \mathbf{b} \in \mathbb{R}^r \) is a vector.

* Address all correspondence to this author.
\[ A \times B = A \otimes B = \begin{bmatrix} \frac{\partial A}{\partial x_1} B & \cdots & \frac{\partial A}{\partial x_i} B & \cdots & \frac{\partial A}{\partial x_s} B \end{bmatrix} \in \mathbb{R}^{q \times t \times s}, \]

where \( B \in \mathbb{R}^{q \times t} \) is a matrix.

\[ CA \times B = C \otimes A \times B = \begin{bmatrix} \frac{\partial A}{\partial x_1} B & \cdots & \frac{\partial A}{\partial x_i} B & \cdots & \frac{\partial A}{\partial x_s} B \end{bmatrix} \in \mathbb{R}^{r \times t \times s}, \]

where \( C \in \mathbb{R}^{r \times q} \) is a matrix.

1 Introduction

Multibody dynamics has become an essential tool for mechanical systems analysis. The progress made during the last decades lead to the development of advanced multibody dynamics techniques and complex models that can account for phenomena otherwise difficult to consider and not feasible to achieve with analytical models. One important opportunity to expand the state-of-the-art research in multibody models is the design optimization of multibody systems with respect to design parameters. Sensitivity analysis of the dynamic response of multibody systems is essential for gradient-based optimization.

Numerical sensitivities, when needed, are often calculated by means of finite differences. However, most of the time, the objective function is not only related to the design parameters but also related to the state variables of the equation of motion. Due to this reason, in order to obtain the numerical sensitivities, the equation of motion must be solved repeatedly. Thus, to calculate numerical sensitivities is computationally expensive. Moreover, in many cases, the results suffer from low accuracy due to computer round-off errors.

Due to the shortcomings of numerical sensitivities, development of analytical approaches to perform sensitivity analysis becomes essential. There are two well-known sensitivity approaches: the direct sensitivity approach and the adjoint sensitivity approach. Haug and Arora, 1978, first presented the adjoint sensitivity approach [1]. In a later study, the sensitivity analysis of dynamic mechanical systems was presented by Haug, Wehage, and Mani, 1984 [2]. The direct sensitivity approach was presented in the same year by Krishnaswami and Bhatti [3]. Methodologies based on these two sensitivity approaches, for various multibody formulations, have then been developed. For example, the direct sensitivity approaches using index-3 and index-1 differential algebraic equations (DAEs) formulations were developed by Haug in 1987 [4] and Chang in 1985 [5]: the direct sensitivity approaches using penalty and augmented Lagrangian formulations was developed by Pagalday in 1997 [6]. The adjoint sensitivity methods using index-3 and index-1 DAEs formulations were developed by Haug in 1981 [7], Haug in 1987 [4], and Bestle in 1992 [8]. For more sensitivity studies, the reader is referred to [9,10,11,12,13,14,15,16,17,18].

Thus, the main task of the study presented in this paper is to overcome these drawbacks and to create a new approach in order to efficiently perform sensitivity analysis and optimization for large and complex systems with respect to a large number of parameters.

Unlike the direct sensitivity approach, the adjoint sensitivity approach works well when the number of parameters is large. On the other hand, solving ordinary differential equations (ODEs) is computationally easier than solving DAEs. Thus, the adjoint sensitivity approaches using ODEs formulations become popular. Dopico and Zhu, 2014, first developed the adjoint sensitivity approach using ODEs formulation [21]. In that paper, direct and adjoint sensitivity approaches are developed for the state-space formulation based on the projection Matrix \( R \) [22], or Maggi’s equations. For Maggi’s formulation, since the dynamic equations are transformed from dependent to independent coordinates at each time step, the approach is not stable when the multibody system goes through a singular or bifurcation position [23].

In this paper, the penalty formulation is used to compute sensitivities. The penalty formulation is an ODEs formulation with dependent coordinates, which was introduced in [24,22]. Comparing this formulation with the DAEs formulations and Maggi’s formulation, the penalty formulation is more stable, and it doesn’t fail around kinematic singularity; it also allows redundant constraints. Moreover, it is more computationally efficient to be solved than DAEs formulations [23]. In addition, unlike Maggi’s formulation, the penalty formulation doesn’t need to restart the numerical integrator for each time step. The shortcoming of the penalty formulation is that it requires an arbitrary value for its penalty factor and for two other coefficients. There is no rigorous method of determining acceptable values for these terms. This penalty factor is typically chosen based on the researcher’s experience with this formulation. For example, \( 10^3 \) is chosen as the value of the penalty factor in this paper and it works perfectly well. The other two coefficients are usually chosen such as to have critical damping in the vibrations associated to the constraints.

The new theory developed is demonstrated on one academic case study, a five-bar mechanism, and on one real-life system, a 14-DOF vehicle model. The five-bar mechanism is used to illustrate the sensitivity approach derived in this paper. The full vehicle model is used to demonstrate the capability of the new approach developed to perform sensitivity analysis and gradient-based optimization for large and complex multibody systems with respect to multiple design parameters.

Finally, using the outputs of the sensitivity analysis, a gradient-based optimization package (L-BFGS-B) [25] is employed to perform the dynamical optimization of the full vehicle system with respect to 6 design parameters.
2 Design optimization of mechanical systems

The design optimization of a mechanical system usually concerns a set of design parameters \( \mathbf{y} \in \mathbb{R}^p \); these parameters are related to the geometry, materials, or other characteristics that must be specified by the design engineer. The optimization theory can considerably help the engineer make such decisions.

The objective of the optimization is to find the values of the selected design parameters that produce the best performance/behavior of the system, under the given constraints. The behavior of the system is represented mathematically by a cost or objective function \( \psi = \psi(\mathbf{y}) \), which is minimized by the optimal value of the parameters.

In cases where the optimization is based on the dynamical behavior of the system under given inputs and initial conditions, the objective function often depends directly on the states of the system in the form \( \psi = \psi(y) \). The system states depend on the parameters \( y = y(\mathbf{y}) \) through the dynamics of the system.

It is also quite usual that the vector of design variables cannot have any values while in the same time it is subjected to some design constraints. The design constraints should be equality or inequality relations, e.g., \( \Psi(\mathbf{y}) = 0 \).

Many advanced numerical optimization methods require the gradient of the objective function/constraints with respect to the parameters. In this paper, based on the outputs of the sensitivity analysis, a gradient-based optimization package (L-BFGS-B) is applied to perform the dynamical optimization. Subsequent sections present the approach developed to perform this sensitivity analysis.

3 Description of the multibody formulation

The equations of motion (EOM) in the penalty formulation \([24]\) have the following expression

\[
\mathbf{M}\ddot{\mathbf{q}} + \Phi^Tq\alpha (\ddot{\Phi} + 2\xi\omega\dot{\Phi} + \omega^2\Phi) = \mathbf{Q},
\]

where \( \alpha \) is the penalty factor, \( \xi \) and \( \omega \) are coefficients of the method, and the rest of the terms are described in the nomenclature.

Equations \([1]\) constitute an ordinary differential equation (ODE) that replaces the constraints of the original index-3 DAEs system by a penalty term that makes possible to estimate the Lagrange multipliers associated to the constraint reactions by the following formula:

\[
\lambda^* = \alpha (\ddot{\Phi} + 2\xi\omega\dot{\Phi} + \omega^2\Phi).
\]

Expanding equation \([1]\), one obtains the following second order ODE system with accelerations as unknowns:

\[
\begin{align*}
\ddot{\mathbf{M}}(\mathbf{q}, \mathbf{p})\ddot{\mathbf{q}} &= \ddot{\mathbf{Q}}(\mathbf{q}, \dot{\mathbf{q}}, t, \mathbf{p}) \quad \text{(3)} \\
\dddot{\mathbf{M}} &= \mathbf{M} + \Phi^Tq\alpha \dddot{\Phi} \quad \text{(4)} \\
\dddot{\mathbf{Q}} &= \mathbf{Q} - \Phi^Tq\alpha (\dddot{\Phi}q + \dddot{\Phi} + 2\xi\omega\dot{\Phi} + \omega^2\Phi) \quad \text{(5)}
\end{align*}
\]

where the following kinematic identities hold

\[
\begin{align*}
\Phi &= \Phi_q \dot{\mathbf{q}} + \Phi_r, \quad \text{(6)} \\
\dddot{\Phi} &= \Phi_q \dddot{\mathbf{q}} + \Phi_r \dddot{\mathbf{q}} + \dddot{\Phi}.
\end{align*}
\]

Note that the EOM \([3]\) depend on some design parameters \( \mathbf{p} \in \mathbb{R}^p \) (typically masses, lengths, or other parameters related to forces chosen by the engineer). Therefore \( \mathbf{q} = \mathbf{q}(t, \mathbf{p}), \dot{\mathbf{q}} = \dot{\mathbf{q}}(t, \mathbf{p}), \) and \( \dddot{\mathbf{q}} = \dddot{\mathbf{q}}(t, \mathbf{p}) \).

4 Adjoint sensitivity approach

The adjoint approach seeks to obtain the sensitivity of a cost function, \( \psi \), with respect to the set of parameters \( \mathbf{p} \). For practical applications, very general cost functions depend not only on positions and velocities, but also on accelerations and reaction forces:

\[
\psi = w(\mathbf{q}_F, \dot{\mathbf{q}}_F, \ddot{\mathbf{q}}_F, \mathbf{p}, \lambda^*_F) + \int_{t_0}^{t_F} g(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}}, \lambda^*_F) \, dt.
\]

The system \( [3] \) can be transformed into a first order system by simply defining a new set of variables by the relation \( \mathbf{q} = \mathbf{v} \),

\[
\begin{bmatrix} 1 & 0 \\ 0 & \mathbf{M} \end{bmatrix} \begin{bmatrix} \dot{\mathbf{q}} \\ \dot{\mathbf{p}} \end{bmatrix} = \begin{bmatrix} \mathbf{v} \\ \mathbf{Q} \end{bmatrix} \iff \mathbf{M}(\mathbf{y}, \mathbf{p}) \ddot{\mathbf{q}} = \dddot{\mathbf{Q}}(t, \mathbf{y}, \mathbf{p}).
\]

In \([7]\), the new state vector is \( \mathbf{y} = [\mathbf{q}^T \mathbf{v}^T]^T \). Taking the inverse of the leading matrix, the system \([9]\) can be expressed as a first order explicit ODE

\[
\ddot{\mathbf{y}} = \mathbf{M}^{-1}(\mathbf{y}, \mathbf{p}) \dddot{\mathbf{q}}(t, \mathbf{y}, \mathbf{p}) = \mathbf{f}(t, \mathbf{y}, \mathbf{p}).
\]

Similarly, the objective function \([3]\) can be expressed as a function of the first order states

\[
\psi = w(\mathbf{y}_F, \dot{\mathbf{y}}_F, \mathbf{p}_F, \lambda^*_F) + \int_{t_0}^{t_F} g(\mathbf{y}, \dot{\mathbf{y}}, \mathbf{p}, \lambda^*_F) \, dt.
\]

Following \([26]\), we consider the following Lagrangian, given by the cost function subject to the EOM constraints

\[
\mathcal{L}(\mathbf{p}) = \psi - \int_{t_0}^{t_F} \mathbf{\mu}^T(\ddot{\mathbf{y}} - \mathbf{f}(t, \mathbf{y}, \mathbf{p})) \, dt,
\]

where \( \mathbf{\mu} \) is the vector of Lagrange multipliers or adjoint variables. Applying variational calculus

\[
\delta \mathcal{L} = \delta \psi - \int_{t_0}^{t_F} \delta \mathbf{\mu}^T(\ddot{\mathbf{y}} - \mathbf{f}(t, \mathbf{y}, \mathbf{p})) \, dt
\]

\[
- \int_{t_0}^{t_F} \mathbf{\mu}^T(\delta \ddot{\mathbf{y}} - \delta \mathbf{f}_y - \delta \mathbf{f}_p \delta \mathbf{p}) \, dt
\]
The central term vanishes if the EOM are fulfilled at each time step.

The variation of the cost function is

\[
\delta \psi = (w_y \delta y + w_y \delta \dot{y} + w_p \delta \dot{p} + w_\lambda \lambda^* \dot{\lambda}^*) F + \int_{t_0}^{T} (g_y \delta y + g_y \delta \dot{y} + g_p \delta \dot{p} + g_\lambda \lambda^* \dot{\lambda}^*) \, dt. \tag{14}
\]

From Eqn. (13)

\[
\delta \lambda^* = \alpha (\delta \Phi + 2\xi_0 \delta \Phi + \omega^2 \delta \Phi), \tag{15a}
\]

where

\[
\begin{align*}
\delta \Phi &= \Phi_q \delta q + (\Phi_{qq} q + \Phi_q + \Phi_q q) \delta q \\
&+ \left( \Phi_{qq} q + (\Phi_q q + (\Phi_q) q) \right) \delta q \\
&+ \left( \Phi_{qp} q + (\Phi_q) p + (\Phi_p) q \right) \delta p \\
\delta \Phi &= \Phi_q \delta q + (\Phi_{qq} q + \Phi_q q + \Phi_p) \delta q \\
&+ \left( \Phi_q \delta q + (\Phi_{qq} q + \Phi_q q + \Phi_q) \delta q \right) \delta q \\
&+ \left( \Phi_{qp} q + (\Phi_q) p + (\Phi_p) q \right) \delta p
\end{align*}
\]

Grouping together the terms associated to \( \delta q \), \( \delta q \), \( \delta \Phi \), \( \delta p \) and taking into account that \( y = [q^T v^T]^T \), Eqn. (15a) becomes

\[
\delta \lambda^* = \lambda^*_y \delta y + \lambda^*_q \delta q + \lambda^*_p \delta p, \tag{16}
\]

Identifying the common terms in (15a) and (16) and using the identity \( v = q \) one obtains

\[
\begin{align*}
\lambda_y &= \left[ \lambda^*_y \lambda^*_q \right] \tag{17} \\
\lambda_q &= \left[ 0 \lambda^*_q \right] \tag{18} \\
\lambda^*_p &= \alpha \Phi_q \tag{19} \\
\lambda^*_q &= \alpha \left[ \Phi_{qq} v + \Phi_q + \Phi_q q + 2\xi_0 \Phi_q q \right] \tag{20} \\
\lambda^*_p &= \alpha \left[ \Phi_{qq} v + \Phi_q + \Phi_q q + 2\xi_0 \Phi_q q \right] + \alpha \left[ \Phi_{qq} v + \Phi_q + \Phi_q q + 2\xi_0 \Phi_q q \right] + \omega^2 \Phi_q \tag{21} \\
\lambda^*_p &= \left[ \Phi_{qq} v + \Phi_q + \Phi_q q + 2\xi_0 \Phi_q q \right] + \omega^2 \Phi_q \tag{22}
\end{align*}
\]

Replacing (16) in (14)

\[
\begin{align*}
\delta \psi &= \left[ (w_y + w_y \lambda^*_q + (w_y + w_y \lambda^*_q) \dot{\lambda}^*) \right] \delta y \\
&+ \left( w_p + w_p \lambda^*_p \right) \delta p \right] \int_{t_0}^{T} + \left[ (g_y + g_y \lambda^*_q + (g_y + g_y \lambda^*_q) \dot{\lambda}^*) \right] \delta y \\
&+ \left( g_p + g_p \lambda^*_p \right) \delta p \right] \delta p \right] \, dt \tag{23}
\end{align*}
\]

For convenience, \( \delta y \) in (23) can be expressed as a function of \( \delta y \). Differentiating Eqn. (10)

\[
\delta y = f_y \delta y + f_p \delta p \tag{24}
\]

and replacing Eqn. (24) in (23) leads to

\[
\begin{align*}
\delta \psi &= \left[ (w_y + w_y \lambda^*_q + (w_y + w_y \lambda^*_q) \dot{\lambda}^*) \right] \delta y \\
&+ \left( w_p + w_p \lambda^*_p + (w_y + w_y \lambda^*_q) \dot{\lambda}^*) \right) \delta p \right] \int_{t_0}^{T} + \left[ (g_y + g_y \lambda^*_q + (g_y + g_y \lambda^*_q) \dot{\lambda}^*) \right] \delta y \\
&+ \left( g_p + g_p \lambda^*_p + (g_y + g_y \lambda^*_q) \dot{\lambda}^*) \right) \delta p \right] \, dt \tag{25}
\end{align*}
\]

The variation of the full Lagrangian (13) can be obtained by replacing (25) in (13)

\[
\delta L = \left[ (w_y + w_y \lambda^*_q + (w_y + w_y \lambda^*_q) \dot{\lambda}^*) \right] \delta y \\
+ \left( w_p + w_p \lambda^*_p + (w_y + w_y \lambda^*_q) \dot{\lambda}^*) \right) \delta p \right] \int_{t_0}^{T} + \left[ (g_y + g_y \lambda^*_q + (g_y + g_y \lambda^*_q) \dot{\lambda}^*) \right] \delta y \\
+ \left( g_p + g_p \lambda^*_p + (g_y + g_y \lambda^*_q) \dot{\lambda}^*) \right) \delta p \right] \, dt \tag{26}
\]

In Eqn. (26), the variation of the parameters \( \delta p \) is known, and variations \( \delta y \) and \( \delta \dot{y} \) could be calculated by solving the linearized form of the EOM (19), but this is computationally expensive. Instead of calculating them, the idea is to cancel these variations. Integrating by parts the integral terms involving \( \delta \dot{y} \) the variation can be removed from the integral

\[
\int_{t_0}^{T} -\mu^T \delta y \, dt = -\mu^T \delta y \big|_{t_0}^{t_f} + \int_{t_0}^{T} \mu^T \delta y \, dt. \tag{27}
\]

Therefore

\[
\delta L = \left[ (w_y + w_y \lambda^*_q + (w_y + w_y \lambda^*_q) \dot{\lambda}^*) \right] \delta y \\
+ \left( w_p + w_p \lambda^*_p + (w_y + w_y \lambda^*_q) \dot{\lambda}^*) \right) \delta p \right] \int_{t_0}^{T} + \left[ (\mu^T + g_y + g_y \lambda^*_q) \right] \delta y \\
+ \left( g_p + g_p \lambda^*_p + (\mu^T + g_y + g_y \lambda^*_q) \dot{\lambda}^*) \right) \delta p \right] \, dt \tag{28}
\]

In Eqn. (28), it is possible to cancel \( \delta \dot{y} \) by choosing \( \mu \) to be the solution of following adjoint ODE system

\[
\mu = -f_y^T \left( \mu + g_y^T + \lambda_y^T \dot{\lambda}_y^* \right) - g_y^T \lambda_y^T \dot{\lambda}_y^*, \tag{29}
\]

\[
\mu_F = \left[ w_y^T + \lambda_y^T \dot{w}_y^* + f_y^T \left( w_y + w_y \lambda_y^* \right) \right]^T. \tag{30}
\]
The adjoint system (29) is a first order linear ODE in $\mu$. Since the initial conditions (30) are given at the final time $t_f$, it has to be integrated backward in time from $t_f$ to $t_0$ as an initial value problem.

Finally, from Eqns. (28) the gradient of the cost function with respect to parameters can be obtained as

$$
\nabla_{\rho} \mathcal{J} = \left[ w_p^T + \lambda_p^T w_{\lambda}^T + f_p^T (w_y + w_{\lambda}^T \lambda_y) \right]_{t_f}^T + \left[ \frac{\partial y_f^T}{\partial \rho} \mu \right]_{t_f}^T + \int_{t_f}^{t_0} \left[ \frac{\partial f_p^T}{\partial \rho} \mu + g_f^T (y_f + \lambda_y^T \lambda_f) + g_p^T + \lambda_p^T \lambda_f \right] dt,
$$

where the identity $\delta y = \delta L$ was used. This holds if the EOM are satisfied, as can be seen from Eqn. (12).

In Eqns. (31) and (29) the derivatives of function $g$ are known, since the objective function has a known expression. The derivatives of $f$ are obtained using (9) as

$$
\dot{M} \frac{\partial f}{\partial y} + \dot{M}_y f = \frac{\partial Q}{\partial y} \Rightarrow f_y = \dot{M}^{-1} (Q_y - \dot{M}_y f),
$$

$$
\dot{M} \frac{\partial f}{\partial \rho} + \dot{M}_\rho f = \frac{\partial Q}{\partial \rho} \Rightarrow f_\rho = \dot{M}^{-1} (Q_\rho - \dot{M}_\rho f).
$$

The derivatives $f_y$ and $f_\rho$ can be calculated in block form as

$$
f_y = \left[ \begin{array}{cc} 1 & 0 \\ 0 & M^{-1} \end{array} \right] \left[ \begin{array}{cc} 0 & 1 \\ -\dot{K} & -\dot{C} \end{array} \right] \left[ \begin{array}{cc} 0 & 0 \\ \dot{M}_y & 0 \end{array} \right] =
$$

$$
\left[ \begin{array}{cc} 0 & 1 \\ -M^{-1} (K + \dot{M}_y \dot{y}) & -M^{-1} \dot{C} \end{array} \right],
$$

$$
f_\rho = \left[ \begin{array}{cc} 1 & 0 \\ 0 & M \end{array} \right]^{-1} \left[ \begin{array}{cc} 0 & 0 \\ \dot{Q}_\rho & -\dot{M}_\rho \dot{\rho} \end{array} \right] =
$$

$$
\left[ \begin{array}{cc} 0 & 1 \\ \dot{M}^{-1} (\dot{Q}_\rho - \dot{M}_\rho \dot{\rho}) \end{array} \right].
$$

In Eqns. (33) and (34) the terms $\dot{K}$, $\dot{C}$, $\dot{Q}_p$, $\dot{M}_p \dot{q}$, and $\dot{M}_\rho \dot{q}$ are given by the following expressions:

$$
\dot{K} = \frac{\partial Q}{\partial q} = K + \Phi_{qq}^T \alpha (\Phi_y q + \Phi_\rho q + 2\dot{x} \omega \Phi + \omega^2 \Phi) +
$$

$$
\Phi_{qq}^T \alpha \left( (\Phi_y q)_q + (\Phi_\rho q)_q + 2\dot{x} \omega (\Phi_{qq} q + \Phi_q q) + \omega^2 \Phi_q \right),
$$

$$
\dot{C} = \frac{\partial Q}{\partial \rho} = C + \Phi_{\rho q}^T \alpha (\Phi_{qq} \dot{q} + \Phi_\rho \dot{q} + 2\dot{x} \omega \Phi_q),
$$

$$
\dot{Q}_p = \frac{\partial Q}{\partial \rho} = Q_p - \Phi_{\rho q}^T \alpha (\Phi_y q + \Phi_\rho q + 2\dot{x} \omega \Phi + \omega^2 \Phi) -
$$

$$
\Phi_{\rho q}^T \alpha \left( (\Phi_y q)_p + \Phi_\rho \dot{q} + 2\dot{x} \omega \Phi_p + \omega^2 \Phi_p \right),
$$

$$
\dot{M}_q = \dot{M}_p q + \Phi_{qq}^T (\alpha \Phi_y q + \Phi_{qq} \alpha (\Phi_q q)),
$$

$$
\dot{M}_\rho = \dot{M}_\rho q + \Phi_{\rho q}^T (\alpha \Phi_y q + \Phi_{\rho q} \alpha (\Phi_q q)).
$$

In Eqns. (35) and (36), $K = -Q_q$ and $C = -Q_\rho$ respectively. For Eqns. (38) and (39), the following magnitudes are tensor-vector products that have to be calculated as explained in the nomenclature:

$$
M_q q \equiv M_q \otimes \dot{q},
$$

$$
M_\rho q \equiv M_\rho \otimes \dot{q},
$$

$$
\Phi_{qq}^T (\alpha \Phi_y q) \equiv \Phi_{qq}^T \otimes (\alpha \Phi_y q),
$$

$$
\Phi_{\rho q}^T (\alpha \Phi_y q) \equiv \Phi_{\rho q}^T \otimes (\alpha \Phi_y q),
$$

$$
\Phi_{qq}^T \alpha (\Phi_q q) \equiv \Phi_{qq}^T \alpha (\Phi_q q),
$$

$$
\Phi_{\rho q}^T \alpha (\Phi_q q) \equiv \Phi_{\rho q}^T \alpha (\Phi_q q).
$$

To obtain expression (35), the kinematic relation (6) was employed, and for expression (36) the relations $(\Phi_y)_{q_k} = \Phi_{qq}, \ (\Phi_\rho)_{q_k} = \Phi_{qq}$, were used. The last two relations can be checked by considering the following differentials

$$
\delta \Phi_q = \Phi_{qq} \delta q \Rightarrow \delta \Phi_q = \Phi_{qq} \delta q + \Phi_{qq} \delta q =
$$

$$
\delta \Phi_q = \Phi_{qq} \delta q + \Phi_{qq} \delta q \Rightarrow \Phi_q = \Phi_{qq},
$$

$$
\delta \Phi_q = \Phi_{qq} \delta q \Rightarrow \delta \Phi_q = \Phi_{qq} \delta q + \Phi_{qq} \delta q =
$$

$$
\delta \Phi_q = \Phi_{qq} \delta q + \Phi_{qq} \delta q \Rightarrow \Phi_q = \Phi_{qq}.
$$

5 Sensitivity analysis of five-bar mechanism

The mechanism chosen as a case study on which to test the sensitivity approach developed is the five-bar mechanism with two degrees of freedom shown in Fig. 1. The five bars are constrained by five revolute joints located at 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 of 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 as $L_{01} = \sqrt{2^2 + 1^2} \, m$ and $L_{02} = \sqrt{2^2 + 0.5^2} \, m$, coincident with the initial configuration shown in Fig. 1.

The following objective function is proposed to obtain its sensitivity with respect to the parameters $p = [L_{01} \, L_{02}]$.

$$\psi = \int_{t_0}^{t_f} (r_2 - r_{20})^T (r_2 - r_{20}) \, dt,$$  \hspace{1cm} (48)

where $r_2$ is the global position of the point 2 and $r_{20}$ is the initial position of the same point.

In order to validate the sensitivity approach derived in this paper, the sensitivities obtained for the five-bar mechanism are calculated using different approaches and compared. The sensitivities, $\nabla_p \psi$, were obtained by the following approaches:

1. Direct sensitivity approach using the index-3 DAEs formulation
2. Direct sensitivity approach using the index-1 DAEs formulation
3. Direct sensitivity approach using the penalty formulation
4. Direct sensitivity approach using Maggi’s formulation
5. Adjoint sensitivity approach using the index-3 DAEs formulation
6. Adjoint sensitivity approach using the index-1 DAEs formulation
7. Adjoint sensitivity approach using the penalty formulation
8. Adjoint sensitivity approach using Maggi’s formulation
9. Finite difference method with perturbation $\delta = 10^{-7} \, m$.
10. Finite difference method with perturbation $\delta = 10^{-4} \, m$.

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

As it can be seen in Table 1, all the approaches employed yield similar results, as expected, except the finite difference method with perturbations $\delta = 10^{-7} \, m$, thus perfectly validating the adjoint sensitivity approach using penalty formulation developed in this study. It is important to remark that, when the perturbation is too small, the sensitivities generated by finite difference method suffer from low accuracy due to round-off error generated by the computer. On the other hand, if the perturbation is too big, the finite difference method is not accurate anymore. Thus, under such circumstances, it is concluded that the finite difference method is not reliable.

6 Sensitivity analysis and optimization of the dynamic response of a full vehicle

In addition to the five-bar mechanism, the approach derived in this paper has been used to perform the sensitivity analysis and the optimization of vehicle ride response for the Iltis vehicle shown in Fig. 2. The Iltis vehicle was proposed as a benchmark problem by the European automobile industry to check multibody dynamic codes. The vehicle model is extensively described in [27], therefore only a summary of the most important parameters of the model is given here. Because the tire model employed in this study is different than the one described in the reference, it will be fully described here.

A speed bump test has been implemented. The objective function for this scenario is the fourth power of the chassis’ CG vertical acceleration. It is minimized while the vehicle goes straight over a road with a small step located at some distance ahead from the initial point.

The sensitivities obtained for the vehicle model are not presented separately, but they are applied to perform the design optimization using L-BFGS-B. The results for the optimization are presented.

6.1 Iltis vehicle model

6.1.1 Vehicle topology

The vehicle is represented in Fig. 2 and a topology diagram of the model is given in Fig. 3, showing that the model

<table>
<thead>
<tr>
<th>Approach</th>
<th>Parameters</th>
<th>$d\psi/dL_{01}$</th>
<th>$d\psi/dL_{02}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Direct index-3</td>
<td>$h = 10^{-2} , s$</td>
<td>-4.2381</td>
<td>3.2170</td>
</tr>
<tr>
<td>2: Direct index-1</td>
<td>$h = 10^{-2} , s$</td>
<td>-4.2383</td>
<td>3.2169</td>
</tr>
<tr>
<td>3: Direct penalty</td>
<td>$h = 10^{-2} , s$</td>
<td>-4.2305</td>
<td>3.2154</td>
</tr>
<tr>
<td>4: Direct Maggi’s</td>
<td>$h = 10^{-2} , s$</td>
<td>-4.2300</td>
<td>3.2112</td>
</tr>
<tr>
<td>5: Adjoint index-3</td>
<td>$h = 10^{-2} , s$</td>
<td>-4.2287</td>
<td>3.2090</td>
</tr>
<tr>
<td>6: Adjoint index-1</td>
<td>$h = 10^{-2} , s$</td>
<td>-4.2294</td>
<td>3.2094</td>
</tr>
<tr>
<td>7: Adjoint penalty</td>
<td>$h = 10^{-2} , s$</td>
<td>-4.2293</td>
<td>3.2137</td>
</tr>
<tr>
<td>8: Ajoint Maggi’s</td>
<td>$h = 10^{-2} , s$</td>
<td>-4.2294</td>
<td>3.2093</td>
</tr>
<tr>
<td>9: Num. diff.</td>
<td>$\delta = 10^{-7} , m$</td>
<td>-9.7390</td>
<td>-4.0344</td>
</tr>
<tr>
<td>10: Num. diff.</td>
<td>$\delta = 10^{-4} , m$</td>
<td>-4.2194</td>
<td>3.2055</td>
</tr>
</tbody>
</table>
is composed of 20 bodies: the chassis, 4 bodies per suspension, 1 tie rod per each one of the front suspensions, and the steering rod. The bodies of the model are joined by 25 kinematic joints plus 3 extra primitive constraints: 16 revolute joints, 8 spherical joints, 1 translational joint, 2 constraints to avoid the rotation of the tie rods, and a rheonomic constraint to control the steering rod.

The total number of coordinates is 140 and the total number of constraints is 132 (6 of them redundant) giving a total count of 14 DOF: 6 DOF for the chassis, 4 DOF for the suspensions and 4 DOF for the wheels rotation. The steering is controlled by means of the mentioned rheonomic constraint and therefore it is not a true DOF since it is kinematically determined.

Masses and moments of inertia are given in Table 2. As indicated in [27], the masses of bodies not included in the table are neglected, and all the moments of inertia are principal, therefore they are given in their centroidal reference frames and all products of inertia are considered to be zero.

Centers of mass locations are given in Table 3, expressed in the reference frame $C$, indicated in Fig. 2.

The geometry of the left front suspensions is shown in Fig. 4. The rear suspensions have a similar topology, but without the steering system. Note that the leaf spring is modeled as a link and a linear spring.

### Table 2. Mass and principal moments of inertia

<table>
<thead>
<tr>
<th>Body</th>
<th>Mass [kg]</th>
<th>$I_{xx}$ [kg m²]</th>
<th>$I_{yy}$ [kg m²]</th>
<th>$I_{zz}$ [kg m²]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chassis</td>
<td>1260</td>
<td>130</td>
<td>1620</td>
<td>1670</td>
</tr>
<tr>
<td>Wheel/hub/brake assembly</td>
<td>57.35</td>
<td>1.2402</td>
<td>1.908</td>
<td>1.2402</td>
</tr>
<tr>
<td>A-arm</td>
<td>6.0</td>
<td>0.052099</td>
<td>0.023235</td>
<td>0.068864</td>
</tr>
</tbody>
</table>

### Table 3. Positions of centers of mass (origin C. Fig. 2)

<table>
<thead>
<tr>
<th>Body</th>
<th>x [m]</th>
<th>y [m]</th>
<th>z [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>chassis</td>
<td>0</td>
<td>0</td>
<td>0.57</td>
</tr>
<tr>
<td>right front wheel with hub and brake assembly</td>
<td>0.97</td>
<td>-0.615</td>
<td>0.356</td>
</tr>
<tr>
<td>left rear wheel with hub and brake assembly</td>
<td>-1.047</td>
<td>0.615</td>
<td>0.356</td>
</tr>
<tr>
<td>right front A-arm</td>
<td>0.97</td>
<td>-0.4155</td>
<td>0.2655</td>
</tr>
<tr>
<td>left rear A-arm</td>
<td>-1.047</td>
<td>0.4155</td>
<td>0.2655</td>
</tr>
</tbody>
</table>

### Table 4. Positions of joints (left front suspension, origin C. Fig. 2)

<table>
<thead>
<tr>
<th>Point description</th>
<th>x [m]</th>
<th>y [m]</th>
<th>z [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>wheel center</td>
<td>0.97</td>
<td>0.615</td>
<td>0.356</td>
</tr>
<tr>
<td>A-arm to hub carrier</td>
<td>0.97</td>
<td>0.572</td>
<td>0.229</td>
</tr>
<tr>
<td>A-arm to chassis</td>
<td>0.97</td>
<td>0.259</td>
<td>0.302</td>
</tr>
<tr>
<td>leaf spring to hub carrier</td>
<td>0.97</td>
<td>0.488</td>
<td>0.531</td>
</tr>
<tr>
<td>leaf spring to chassis</td>
<td>0.97</td>
<td>0.1585</td>
<td>0.600</td>
</tr>
<tr>
<td>damper to A-arm</td>
<td>1.045</td>
<td>0.500</td>
<td>0.241</td>
</tr>
<tr>
<td>damper to chassis</td>
<td>1.045</td>
<td>0.297</td>
<td>0.632</td>
</tr>
<tr>
<td>tie rod to hub carrier</td>
<td>0.83</td>
<td>0.448</td>
<td>0.531</td>
</tr>
<tr>
<td>tie rod to chassis</td>
<td>0.83</td>
<td>0.07</td>
<td>0.600</td>
</tr>
<tr>
<td>steering rack to chassis</td>
<td>0.83</td>
<td>0.00</td>
<td>0.600</td>
</tr>
</tbody>
</table>
Each one of the four suspensions has three force elements: a linear leaf-spring that represents the stiffness of the leaf spring, a bump stop, and a non-linear spring-damper element. The suspension forces in the nominal configuration are given in Table 5.

Table 5. Suspension forces in the nominal configuration

<table>
<thead>
<tr>
<th>Force Type</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leaf spring force</td>
<td>2728.9 N</td>
</tr>
<tr>
<td>Non-linear Spring-Damper force</td>
<td>128.0 N</td>
</tr>
<tr>
<td>Bump stop force</td>
<td>0.0 N</td>
</tr>
</tbody>
</table>

The force of the leaf spring can be represented by the following equation,

\[ F_L = -k_L (L - (1 + 2728.9/35906 N/m)). \] (49)

where \( L \) is distance between the spring extreme points, the stiffness is originally \( k_L = 35906 N/m \), and in the nominal (initial) configuration \( L = 1 m \) and the leaf spring force is equal to \( F_L = 2728.9 N \).

The force of the bump stop is given by

\[ F_B = -10^7 (s - 0.93); s < 0.93 m, \] (50)

\[ F_B = 0; s \geq 0.93 m. \] (51)

The elastic and the damping forces of the nonlinear spring and damper system are given by the following expressions

\[ F_s = -4.0092 \cdot 10^6 + k_S \cdot 10^7 s \]

\[-6.7061 \cdot 10^5 \cdot s^2 + 5.2796 \cdot 10^7 s^3, \] (52)

\[ F_d = c_S \cdot v + 33955.72v^2 - 59832.25v^3 \]

\[-395651.0v^4; -0.2 < v < 0.21 m/s, \] (53)

\[ F_d = -416.4200 + 1844.3v; v < -0.2 m/s, \] (54)

\[ F_d = 1919.1638 + 1634.727v; v > 0.21 m/s. \] (55)

where \( s \) is the distance between the extreme points of the nonlinear spring-damper, \( v \) is the derivative of \( s \), \( c_S \) is the dominant damping coefficient, \( k_S \) is the dominant stiffness coefficient, and originally \( c_S = 9945.627 N s/m, k_S = 2.8397 \cdot 10^7 N/m \).

6.1.3 The tire model

The tire model consists of normal, longitudinal, and lateral components. The normal tire model component is a linear spring-damper element, and the longitudinal and lateral models are linearized tire models with saturation. The normal tire model is

\[ F_n = -k_n (r - R) n; r < R. \] (56)

where \( r \) is the distance from the center of the wheel to the ground, \( R \) is the tire radius, and \( n \) is the normal vector to the ground in the center of the contact region. The normal tire forces in the nominal configuration are given in Table 6.

Table 6. Tire forces in the nominal configuration

<table>
<thead>
<tr>
<th>Force Type</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>front tyre load</td>
<td>3829.6 N</td>
</tr>
<tr>
<td>rear tyre load</td>
<td>3593.6 N</td>
</tr>
</tbody>
</table>

The longitudinal and the lateral models implemented in this work are described in [28].

\[ F_l = F_x b + F_y (n \times b), \] (57)

\[ F_x = \begin{cases} 
\frac{\mu_x |F_{rad}|}{\kappa} & \kappa \leq \kappa_c, \\
\frac{\mu_x |F_{rad}|}{\kappa} & \kappa > \kappa_c, 
\end{cases} \] (58)

\[ F_y = \begin{cases} 
\frac{\mu_y |F_{rad}|}{\kappa} & \alpha \leq \alpha_c, \\
\frac{\mu_y |F_{rad}|}{\kappa} & \alpha > \alpha_c.
\end{cases} \] (59)

where \( b \) is the unit vector coincident with the wheel rotation axis, \( b = (u \times n) / |u \times n| \) is the longitudinal vector, \( \kappa \) is the longitudinal slip, \( \alpha \) is the slip angle, and \( \kappa_c, \alpha_c \) are the critical slip factors for the longitudinal and lateral models, which are parameters of the tire model.

The longitudinal slip and the slip angle can be defined according to the following expressions, respectively

\[ \kappa = -\frac{-b^T v_{slip}}{b^T v_e} = -\frac{b^T (v_e - v_c)}{b^T v_e}, \] (60)

\[ \alpha = -\arcsin \left( n^T \left( b \times \frac{v_e - (n^T v_c) n}{|v_e - (n^T v_c) n|} \right) \right). \] (61)

where \( v_e \) is the velocity of the center of the wheel, \( \omega \) is the angular velocity of the wheel, and \( r \) the effective radius defined before.

The saturation ellipse between the longitudinal and the lateral forces if given by the following expression

\[ \left( \frac{F_{sat}}{\mu_x} \right)^2 + \left( \frac{F_{sat}}{\mu_y} \right)^2 \leq |F_{rad}|^2. \] (62)

where \( \mu_x \) and \( \mu_y \) stand for the longitudinal and lateral friction coefficients and are parameters of the tire model.
If the components evaluated from Eqn. (58) and Eqn. (59) are not inside the ellipse Eqn. (62), the saturation of the forces take place and the previously calculated forces Eqn. (58) and Eqn. (59) don’t hold. In this case they have to be replaced by the following

\[
F_{x}^{\text{sat}} = \frac{|F_n| F_x}{f_{roz}} F_x, \quad F_{y}^{\text{sat}} = \frac{|F_n| F_y}{f_{roz}} F_y.
\]

\[\text{(63)}\]
\[\text{(64)}\]

### 6.2 Dynamic response optimization

A speed bump test is implemented as a case study on which to perform design optimization. Typically, speed bumps are modeled as cylindrical shapes with the axis below the ground. However, the computation of the contacts between the bumps and the tires is complex. In order to simplify the speed bumps test, a modified speed bumps test is employed in this study which a step is used to replace those bumps as shown in Figure 5. At the beginning, the vehicle is released from equilibrium with an initial velocity of 5 m/s in the longitudinal direction. The steering is not actuated and the vehicle goes straight. At a distance of 6 m ahead from the initial position in the longitudinal direction, a step of 10 cm is placed. After 1 s the vehicle drops down the step and oscillates until the static equilibrium in the vertical direction is reached.

The objective function is the integral in time of the fourth power of the chassis CG vertical acceleration

\[
\psi = \int_{0}^{t} \omega_{\text{chassis}}^4 \, dt.
\]

\[\text{(65)}\]

Six design parameters are chosen to perform design optimization. They are the stiffnesses of the rear and of the front leaf springs \([k_{L1}, k_{L2}]\) from Eqn. (49), the dominant damping coefficients of the rear and of the front suspension \([c_{S1}, c_{S2}]\) from Eqn. (52), and the dominant stiffnesses of the rear and of the front suspensions \([k_{S1}, k_{S2}]\) from Eqn. (52). The following constraints are imposed on the design parameters for the optimization problem:

\[
0 \leq k_{L1}, k_{L2}, k_{S1}, k_{S2} \leq \infty \, \text{N/m} \quad \text{(66)}
\]
\[
0 \leq c_{S1}, c_{S2} \leq \infty \, \text{N s/m} \quad \text{(67)}
\]

The evolutions of the parameters are given in Fig. 6, where it is shown that each parameter successfully converges after 40 iterations. The initial conditions of these parameters are the default values in [27]. The values of the non-optimized and of the optimized parameters and objective function are shown in Table 7.

In Fig. 7, the dynamic responses of original and optimized systems are shown. It can be noted that the response is significantly improved.
From the results above, it can be seen that the vehicle ride response is significantly improved by optimizing 6 parameters simultaneously; it is also shown that each parameter successfully converges after 40 iterations. From the gradient-based optimization point of view, the time of each iteration for optimization is highly related to the time needed to compute the sensitivities. If the sensitivity approach is computationally expensive, it takes a long time to finish each iteration, making gradient-based optimization very difficult. For the optimization implemented in this section, it takes 406 seconds to finish the whole process in 43 iterations, which is very fast. Also, since 6 design parameters were chosen to perform the optimization, one can see that the adjoint sensitivity approach using the penalty formulation developed in this study is able to efficiently compute sensitivities for complex multibody systems with respect to multiple parameters.

7 Conclusions
This paper develops the theoretical adjoint sensitivity approach for multibody system dynamics based on penalty formulations, bringing new contributions to the state-of-the-art in analytical approaches for sensitivity analysis of such systems. A five-bar mechanism and a 14-DOF vehicle model are implemented as case studies to test and validate this sensitivity approach. For the five-bar mechanism, the sensitivity approach is validated by comparing the sensitivities generated using various sensitivity approaches. For the vehicle model, the optimization results presented clearly show that the vehicle ride response is significantly improved by optimizing 6 parameters; this demonstrates the capability of the new approach developed to perform sensitivity analysis for large and complex multibody systems with respect to multiple design parameters with accuracy and efficiency.

Acknowledgements
This work has been partially supported by NSF Award no. 1130667, by the Computational Science Laboratory, and by the Advanced Vehicle Dynamics Laboratory at Virginia Tech.

Table 7. Optimized parameters and objective function

<table>
<thead>
<tr>
<th>Optimizations</th>
<th>Non-optimized value</th>
<th>Optimized value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{L1}$ [N/M]</td>
<td>35906</td>
<td>1497</td>
</tr>
<tr>
<td>$k_{L2}$ [N/M]</td>
<td>35906</td>
<td>3052</td>
</tr>
<tr>
<td>$c_{S1}$ [N s/M]</td>
<td>9946</td>
<td>14531</td>
</tr>
<tr>
<td>$k_{L2}$ [N s/M]</td>
<td>9946</td>
<td>8418</td>
</tr>
<tr>
<td>$k_{S1}$ [N/M]</td>
<td>28397000</td>
<td>28378728</td>
</tr>
<tr>
<td>$k_{S2}$ [N/M]</td>
<td>28397000</td>
<td>28382080</td>
</tr>
<tr>
<td>$\psi$ [N/s^2]</td>
<td>678</td>
<td>48</td>
</tr>
</tbody>
</table>

References


[28] Pablo Luque Rodríguez, Daniel Ivarez Mntaras, C. V.,