Sensitivity Analysis and Optimization of Multibody Systems

Yitao Zhu

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Mechanical Engineering

Corina Sandu, Chair
Adrian Sandu, Co-chair
Daniel Dopico
Jan Helge Bøhn
Steve Southward

December 5, 2014
Blacksburg, Virginia

Keywords: Sensitivity Analysis, Optimization, Multibody Dynamics, Vehicle Dynamics

Sensitivity Analysis and Optimization of Multibody Systems

Yitao Zhu

Abstract

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 for multibody systems. Sensitivity analysis of multibody system dynamics, which is performed before optimization or in parallel, is essential for optimization.

Current sensitivity approaches have limitations in terms of efficiently performing sensitivity analysis for complex systems with respect to multiple design parameters. Thus, we bring new contributions to the state-of-the-art in analytical sensitivity approaches in this study. A direct differentiation method is developed for multibody dynamic models that employ Maggi's formulation. An adjoint variable method is developed for explicit and implicit first order Maggi's formulations, second order Maggi's formulation, and first and second order penalty formulations. The resulting sensitivities are employed to perform optimization of different multibody systems case studies. The collection of benchmark problems includes a five-bar mechanism, a full vehicle model, and a passive dynamic robot. The five-bar mechanism is used to test and validate the sensitivity approaches derived in this paper by comparing them with other sensitivity approaches. The full vehicle system is used to demonstrate the capability of the adjoint variable method based on the penalty formulation to perform sensitivity analysis and optimization for large and complex multibody systems with respect to multiple design parameters with high efficiency.

In addition, a new multibody dynamics software library MBSVT (Multibody Systems at Virginia Tech) is developed in Fortran 2003, with forward kinematics and dynamics,

sensitivity analysis, and optimization capabilities. Several different contact and friction models, which can be used to model point contact and surface contact, are developed and included in MBSVT.

Finally, this study employs reference point coordinates and the penalty formulation to perform dynamic analysis for the passive dynamic robot, simplifying the modeling stage and making the robotic system more stable. The passive dynamic robot is also used to test and validate all the point contact and surface contact models developed in MBSVT.

Acknowledgements

I would like to express my special gratitude to those people who have positively contributed to my experiences at Virginia Tech.

First, I'd like to thank my advisors, Drs. Corina and Adrian Sandu, thank you for your guidance, patience, fund, and encouragement; you helped me grow as a research scientist.

I thank my research partner and my best friend, Dr. Daniel Dopico. You helped me become a multibody dynamics engineer. You have taught me, among many other things, the art of debugging.

I thank my colleagues from the Advanced Vehicle Dynamics Lab (AVDL), and the Computational Science Lab (CSL) for the many laughs, discussion, and friendship.

I thank committee members, Dr. Daniel Dopico, Dr. Jan Helge Bøhn and Dr. Steve Southward, thank you for your contributions to my education and for your insight and feedback on my work.

I thank my parents, in-laws, and siblings, thank you for your love, prayers, and support.

I thank my wonderful wife Yinglian Yang, your sacrifice, patience, love, prayers, encouragement, confidence, and support were invaluable. I could not have done it without you.

Contents

List of Figures

List of Tables

Nomenclature

$\ddot{()}$ $= \dfrac{\mathrm{d}^2\,()}{\mathrm{d}t^2}$

$\dot{()}$ $= \dfrac{\mathrm{d}\,()}{\mathrm{d}t}$

$(\cdots)_0$ Means evaluation at the initial time $(\cdots)\,(t_0)$.

$(\cdots)_F$ Means evaluation at the final time $(\cdots)\,(t_F)$.

$()_{\mathbf{q}}$ $= \dfrac{\partial\,()}{\partial \mathbf{q}}$

$()_{\boldsymbol{\rho}}$ $= \dfrac{\partial\,()}{\partial \boldsymbol{\rho}}$

$()_t$ $= \dfrac{\partial\,()}{\partial t}$

T Time.

$\mathbf{A}_{\mathbf{x}}$ $= \left[\begin{array}{ccccc} \dfrac{\partial \mathbf{A}}{\partial x_1} & \cdots & \dfrac{\partial \mathbf{A}}{\partial x_i} & \cdots & \dfrac{\partial \mathbf{A}}{\partial x_s} \end{array}\right] \in \mathbb{R}^{q\times r\times s}$. Third order tensor of derivatives of matrix $\mathbf{A} \in \mathbb{R}^{q\times r}$ w.r.t. vector $\mathbf{x} \in \mathbb{R}^s$.

$\mathbf{A}_{\mathbf{x}}^{\mathrm{T}}$ $= \left[\begin{array}{ccccc} \dfrac{\partial \mathbf{A}^{\mathrm{T}}}{\partial x_1} & \cdots & \dfrac{\partial \mathbf{A}^{\mathrm{T}}}{\partial x_i} & \cdots & \dfrac{\partial \mathbf{A}^{\mathrm{T}}}{\partial x_s} \end{array}\right] \in \mathbb{R}^{r\times q\times s}$.

$\mathbf{A}_{\mathbf{x}}\mathbf{B} = \mathbf{A}_{\mathbf{x}} \otimes \mathbf{B} = \left[\begin{array}{ccccc} \dfrac{\partial \mathbf{A}}{\partial x_1}\mathbf{B} & \cdots & \dfrac{\partial \mathbf{A}}{\partial x_i}\mathbf{B} & \cdots & \dfrac{\partial \mathbf{A}}{\partial x_s}\mathbf{B} \end{array}\right] \in \mathbb{R}^{q\times t\times s}$, where $\mathbf{B} \in \mathbb{R}^{r\times t}$ is a matrix.

$\mathbf{A}_{\mathbf{x}}\mathbf{b} = \mathbf{A}_{\mathbf{x}} \otimes \mathbf{b} = \left[\begin{array}{ccccc} \dfrac{\partial \mathbf{A}}{\partial x_1}\mathbf{b} & \cdots & \dfrac{\partial \mathbf{A}}{\partial x_i}\mathbf{b} & \cdots & \dfrac{\partial \mathbf{A}}{\partial x_s}\mathbf{b} \end{array}\right] \in \mathbb{R}^{q\times s}$, where $\mathbf{b} \in \mathbb{R}^r$ is a vector.

$\mathbf{C}\mathbf{A}_{\mathbf{x}}\mathbf{B} = \mathbf{C} \otimes \mathbf{A}_{\mathbf{x}}\mathbf{B} = \left[\begin{array}{ccccc} \mathbf{C}\dfrac{\partial \mathbf{A}}{\partial x_1}\mathbf{B} & \cdots & \mathbf{C}\dfrac{\partial \mathbf{A}}{\partial x_i}\mathbf{B} & \cdots & \mathbf{C}\dfrac{\partial \mathbf{A}}{\partial x_s}\mathbf{B} \end{array}\right] \in \mathbb{R}^{r\times t\times s}$, where $\mathbf{C} \in \mathbb{R}^{r\times q}$ is a matrix.

$\mathbf{M}\,(\mathbf{q}, \boldsymbol{\rho}) \in \mathbb{R}^{n\times n}$ Generalized mass matrix of the system.

$\mathbf{q} \in \mathbb{R}^n$ Vector of coordinates of the system.

$\mathbf{Q}(\mathbf{q}, \dot{\mathbf{q}}, t, \boldsymbol{\rho}) \in \mathbb{R}^n$ Vector of generalized forces of the system.

$\boldsymbol{\Phi}(\mathbf{q}, t, \boldsymbol{\rho}) \in \mathbb{R}^m$ Vector of constraints that relate the dependent coordinates.

$\boldsymbol{\rho} \in \mathbb{R}^p$ Vector of parameters.

CG Center of gravity.

DOF Degree or degrees of a freedom

Chapter 1

Introduction

1.1 Motivation and Objectives

1.1.1 Motivation

Multibody dynamics has become an essential tool for mechanical systems analysis and design. Progress during the last decades lead to the development of complex multibody models that consider phenomena difficult to take into account in the past and impossible to achieve with analytical models.

One of the most interesting problems that brought attraction since even before of the earlier developments of the multibody systems techniques, is the optimization of the dynamic response of mechanical systems [6]. Nowadays, with the improvement of computer technology and computational methods, the possibility of performing the optimization of complex industrial problems becomes an interesting topic in the multibody community. Sensitivity analysis, which quantifies the effect of specific design parameters of interest on the dynamic response of a given multibody system, is often performed before the optimization or in parallel with it. The purpose of sensitivity analysis is to obtain the gradient of an objective function with respect to the design parameters. The outputs of sensitivity analysis are next used by gradient-based optimization packages to perform optimization.

Numerical sensitivities, when needed, are often calculated by means of finite differences. However, to calculate numerical sensitivities is computationally expensive. Moreover, in many cases the results suffer from low accuracy due to round-off error generated by the computer.

Due to the shortcomings of numerical sensitivities, the development of analytical approaches to perform sensitivity analysis becomes essential. There are two well-known sensitivity approaches: the direct differentiation method (DDM) and the adjoint variable method (AVM). Haug and Arora, 1978, first extended the adjoint variable method from control theory to multibody systems optimization[7]. Later on, the sensitivity analysis of dynamical mechanical systems was presented by Haug, Wehage, and Mani, 1984 [8]. The direct differentiation method was presented in the same year by Krishnaswami and Bhatti [9]. These two sensitivity approaches for different multibody formulations have already been developed, as summarized in section 1.3.1. However, these methods have some drawbacks that prevent them from easily computing sensitivities for large and complex multibody systems. For instance, the direct differentiation method works well when the number of parameters is small, but it becomes computationally expensive when the number of parameters is large. In contrast, the adjoint variable method works well when the number of parameters is large, but it doesn't work well when the number of adjoint variables becomes large. On the other hand, the direct differentiation methods and the adjoint sensitivity methods using index-3 and index-1 differential algebraic system of equations (DAE) formulations are not practical due to the numerical difficulties associated to the underlying formulations [10, 11]. To date, the sensitivity analysis and optimization of large and complex multibody systems with respect to a large number of design parameters is still an open topic.

Moreover, current commercial software packages and academic software packages for multibody dynamics have some drawbacks. First, most of these packages don't release the source code and the user is constrained to work within the package's given limitations. Second, some of these packages don't have the capability to perform sensitivity analysis and optimization for multibody system while others are able to perform sensitivity analysis and optimization only for very simple multibody systems. Thus, in order to perform sensitivity analysis and optimization for large and complex multibody systems with respect to a large number of parameters, a new multibody dynamics software that includes new sensitivity approaches should be developed and validated.

An interesting area of application for the multibody dynamics techniques is robotics. In general, relative coordinates and a Lagrange formulation are used for the analysis of robotic systems. There are several drawbacks associated to the use of this approach. First, it is difficult to assess if all the *DOF* associated with these coordinates will be valid during the entire simulation; the system may become unstable when it goes through a singular or a bifurcation position. Second, it is more difficult to model and to write the equations with relative than with reference point coordinates. Thus, a more general-purpose method is of interest.

1.1.2 Objectives

- The first objective of this study is to overcome the drawbacks of current sensitivity approaches and to create new algorithms and techniques in order to efficiently perform sensitivity analysis for large and complex systems with respect to a large number of parameters.

- The second objective is to validate and test these new analytical sensitivity approaches by applying them to an illustrative example. The case study selected here is a five-bar mechanism; the results are compared with other analytical and numerical approaches.

- The third objective is to demonstrate the capability of the new approaches to perform sensitivity analysis for large and complex multibody systems with respect to multiple design parameters. This is done by applying the newly developed sensitivity analysis techniques to a three-dimensional (3D) full vehicle system.

- Using the outputs of the sensitivity analysis, several gradient-based optimization packages can be used to perform dynamical optimization. Thus, the fourth objective is to perform dynamic response optimization of large and complex multibody systems with respect to multiple parameters. The case studies presented here are for vehicle ride optimization and vehicle handling optimization with respect to suspension parameters.

- The fifth objective is to develop a multibody package for research and education of multibody dynamics and vehicle dynamics, and to make it available to the scientific community. The new package, Multibody Systems at Virginia Tech (MBSVT), uses the new approaches developed in this study and is able to perform sensitivity analysis and optimization for large and complex multibody systems.

- The last objective is to develop a more systematic and general-purpose approach for the dynamic analysis of legged robotic systems. A passive dynamic two-legged robot is presented as a case study. The robot can also be used to test and validate all the point contact and surface contact models developed in MBSVT.

1.2 Multibody dynamics formulations

1.2.1 Coordinates system

Multibody systems can be described with different coordinates. There are three important types of coordinates: reference point coordinates, natural coordinates, and relative coordinates.

- Relative coordinates

  Relative coordinates define the relative position between the current element and the previous element in the kinematic chain. If two elements are linked by a revolute joint, then their relative position is the relative angle between these two elements. If they are linked by a translational joint, then their relative position is the distance between two points on these two elements. In the spatial multibody systems, there are more types of joints and unlike the planar systems, most of these joints allow more than one *DOF*. Denavit-Hartenberg Notation is mainly used to establish the relative coordinates in spatial systems, especially robotic systems. Relative coordinates are especially suitable for open-loop multibody systems attached to the ground because in

this case the number of coordinates is equal to the number of DOF and no constraints are needed.

- Reference point coordinates

  Reference point coordinates are also called Cartesian coordinates. For the planar systems, they determine the position of a point of a body (typically the CG) by using two Cartesian coordinates and they determine the orientation of the body by a angle. For spatial systems, they define the position of a body by using the absolute position of a point (typically the CG). They define the orientation of the body by three angles (typically Euler angles or Fixed angles) or four Euler parameters. Modeling multibody systems with reference point coordinates normally requires more variables than when using relative coordinates. However, since the terms in the equations of motion (EOM) are sparse, one may make the formulation efficient by some special techniques.

- Natural coordinates

  Natural coordinates were introduced by J. Garcia de Jalon and Serna in 1981, 1982, 1986, and 1987 for planar systems and spatial systems [12–16]. For the planar systems, these coordinates require the absolute positions of two points (typically at the joints) for one body. For spatial systems, these coordinates require a sufficient number of points and unit vectors to define both the position and orientation of bodies. For more details about natural coordinates, the reader is referred to [17].

1.2.2 Formulations of equations of motion

There is abundant literature about multibody dynamics formulations of EOM, only the most relevant ones are introduced here. Initially, Swiss mathematician Leonhard Euler and Italian mathematician Joseph-Louis Lagrange developed Lagrange's equations, which became one of the theoretical foundations of multibody dynamics. For a constrained system,

the Lagrange's equations become

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}}\right) - \frac{\partial L}{\partial \mathbf{q}} + \mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}}\boldsymbol{\lambda} = \mathbf{Q}_{ex} \tag{1.1}$$

where $L = T - V$ is the Lagrangian function, $\mathbf{\Phi}_{\mathbf{q}}$ is the Jacobian matrix of the vector of constraint equations, $\boldsymbol{\lambda}$ is the vector of Lagrange multipliers associated to the constraints forces, and $\mathbf{Q}_{ex}$ is the vector of external forces. This equation is only true if the constraint equations are satisfied. The kinetic energy of a multibody system can be written as follows:

$$T = \frac{1}{2}\dot{\mathbf{q}}^{\mathrm{T}}\mathbf{M}\dot{\mathbf{q}} \tag{1.2}$$

Replacing (1.2) in (1.1), the well-know Lagrange multipler form of EOM is obtained as follow:

$$\mathbf{M}\ddot{\mathbf{q}} + \mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}}\boldsymbol{\lambda} = \mathbf{Q} \tag{1.3}$$

where $\mathbf{Q} = \mathbf{Q}_{ex} - \dot{\mathbf{M}}\dot{\mathbf{q}} + T_{\mathbf{q}} - V_{\mathbf{q}}$, which contains the external forces and velocity dependent inertia forces. Equation (1.3) can also be obtained by the method of virtual power. Since this is a constrained system, the position, velocity, and acceleration constraint equations must be fulfilled as follows:

$$\mathbf{\Phi} \equiv \mathbf{\Phi}(\mathbf{q}, t) = \mathbf{0} \tag{1.4}$$

$$\dot{\mathbf{\Phi}} \equiv \mathbf{\Phi}_{\mathbf{q}}\dot{\mathbf{q}} + \mathbf{\Phi}_t = \mathbf{0} \tag{1.5}$$

$$\ddot{\mathbf{\Phi}} \equiv \mathbf{\Phi}_{\mathbf{q}}\ddot{\mathbf{q}} + \dot{\mathbf{\Phi}}_{\mathbf{q}}\dot{\mathbf{q}} + \dot{\mathbf{\Phi}}_t = \mathbf{0} \tag{1.6}$$

Equations (1.3) with (1.4) constitute an index-3 DAE formulation. Equations (1.3) with (1.5) constitute an index-2 DAE formulation. To solve a high index DAE formulation is always a difficult problem, which needs index-reduction techniques. Thus, these formulations are not the most popular formulations in the multibody community. Equations (1.3) with (1.6) constitute an index-1 DAE formulation, which can also be written in matrix format as

follows:

$$\begin{bmatrix} \mathbf{M} & \mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}} \\ \mathbf{\Phi}_{\mathbf{q}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{q}} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{Q} \\ \mathbf{c} \end{bmatrix} \tag{1.7}$$

where $\mathbf{c} = -\dot{\mathbf{\Phi}}_{\mathbf{q}}\dot{\mathbf{q}} - \dot{\mathbf{\Phi}}_t$. This system matrix is also called augmented matrix, which is presented by Negrut, Serban and Potra in 1997 [18].

A direct numerical solution of (1.7) suffers from drift-off [19], meaning that any small perturbation in the acceleration constraints leads to an error in the position constraints that grows quadratically with time. One of the popular solutions is the Baumgarte stabilization method that was developed by Baumgarte in 1972 [20].

In general, the multibody dynamics equations, constitute a DAE that it is not usually directly solved because of the numerical difficulties involved [10, 11]. Due to the drawbacks of DAE formulations, some of the most advanced families of formulations presented in the eighties and nineties are based on ODE-like EOM in dependent coordinates or independent coordinates, such as Maggi's formulation [17] and the penalty formulation [21].

Maggi's formulation, which can be also called state-space formulation based on the projection Matrix R, is one of the well-known formulations in independent coordinates. It was described in [17]. For Maggi's formulation, the dynamic equations are transformed from dependent to independent coordinates at each time step by projecting the dependent coordinates vector on the rows of a constant matrix $\mathbf{B}$ of size $((n-m) \times n)$ where n is the number of dependent coordinates and m is the number of constraints. This formulation leads to a smaller size system of EOM. However, it's not stable when the multibody system goes through a singular or bifurcation position [22].

E. Bayo, J. Garca de Jalon, and M.A. Serna, 1988 [21], presented the penalty formulation, which is one of the ODE-like formulations in dependent coordinates. Compared with Maggi's formulation or other formulations in independent coordinates, the penalty formulation is more stable when the multibody system goes through a singular or bifurcation position. In addtion, it doesn't fail around kinematic singularity. Furthermore, the penalty

formulation allows redundant constraints, significantly simplifying the modeling process.

However, the shortcoming of the penalty formulation is that it requires to choose the value for its penalty factor. It is not always clear how to determine reasonable values for this factor. In addition, although a large penalty factor ensures convergence to the constraints within a tight tolerance, it usually generates round-off errors and leads to numerical conditioning problems. To improve the numerical conditioning of the penalty formulation, an augmented Lagrangian formulation is described in [21]. All the formulations introduced above are in dependent coordinates.

Flexible multibody dynamics established itself as a new field in early 1970s. The number of coordinates required for flexible multibody systems is much larger than the number of coordinates required for rigid multibody systems. Shabana, 1997 [23], reviewed some of the most famous formulations for flexible multibody dynamics, such as the floating frame of reference formulation, the finite element incremental methods, large rotation vector formulations, the finite segment method, the linear theory of elastodynamics, and one of the most recent formulations, the absolute nodal coordinate formulation (ANCF). Since this study focuses only on the sensitivity analysis and optimization of rigid body dynamics, these methods are not going to be introduced in detail. For more details about these methods, the reader is referred to [24–37].

1.3 Optimization of multibody systems by using sensitivity analysis

The design optimization of a mechanical system usually concerns a set of design parameters $\boldsymbol{\rho} \in \mathbb{R}^p$. These parameters are related to the geometry, materials, or other characteristics that need to be specified by the engineer. The optimization theory can considerably help the engineer to make such decisions.

The objective of the optimization is to find a design that makes the behavior of the system optimal. The behavior of the system is represented mathematically by a cost or objective function $\psi = \psi(\boldsymbol{\rho})$, 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 $\psi = \psi(\mathbf{y})$. The system states depend on the parameters $\mathbf{y} = \mathbf{y}(\boldsymbol{\rho})$ through the dynamics of the system. Thus, the objective function often depends indirectly on the parameters in the form $\psi = \psi(\mathbf{y}(\boldsymbol{\rho}), \boldsymbol{\rho})$.

It is also quite usual that the vector of design variables cannot have any value and it is subjected to some design constraints. The design constraints should be equality or inequality relations, e.g., $\boldsymbol{\Psi}(\boldsymbol{\rho}) = 0$.

Many advanced numerical optimization methods require the gradient of the objective function with respect to the parameters. In this study, with the outputs of the sensitivity analysis, a gradient-based optimization package (L-BFGS-B) is used to perform the dynamical optimization. More details about how to perform this sensitivity analysis will be discussed in section 3.

1.3.1 Sensitivity analysis

State of the art

Sensitivity analysis is defined as analyzing and quantifying the effects of the system parameters on the outputs. It is the basis of gradient-based optimization, and usually, performed before the optimization procedures or in parallel.

For a large and complex multibody system, there is a large number of system parameters. To perform optimization with respect to all the parameters is extremely difficult. For this reason, the first objective of sensitivity analysis is to find those parameters that are most

relevant to the design objective such that the other unimportant parameters don't have to be considered. By doing this, the problem can be simplified and the optimization with respect to these important parameters becomes easier and more efficient. The second objective is to obtain the gradient of the objective function with respect to the system parameters. The gradient can be used by different optimization packages to perform optimization.

Numerical sensitivities, when needed, are often calculated by means of finite differences. However, most of the time, the objective function depends directly on the states of the system in the form $\psi = \psi(\mathbf{y})$. The system states depend on the parameters $\mathbf{y} = \mathbf{y}(\boldsymbol{\rho})$ through the dynamics of the system. Because of that, in order to obtain the numerical sensitivities, the EOM has to be solved many times. Thus, to calculate numerical sensitivities is computationally expensive. Moreover, in many cases ,the results suffer from low accuracy due to round-off error generated by the computer. For more details about the drawbacks of numerical sensitivities, the reader is referred to [38], which introduces the numerical sensitivity analysis of DAE formulations and the problems of numerical sensitivities.

Due to the shortcomings of numerical sensitivities, analytical approaches to perform sensitivity analysis becomes essential. As shown in section 1.1, there are two well-known sensitivity approaches: the direct differentiation method and the adjoint variable method. Haug and Arora, 1978, first applied the adjoint variable method to perform optimization of multibody systems [7]. Later on, the sensitivity analysis of kinematic and dynamic mechanical systems was presented by Haug, Sohoni, Wehage, and Mani, in 1982 [39] and 1984 [8] respectively. The direct differentiation method was presented in the same year by Krishnaswami and Bhatti [9]. These two sensitivity approaches based on different multibody formulations have already been developed in the last 30 years. The direct differentiation method using index-3 and index-1 DAE formulations were developed by Chang in 1985 [40] and Haug in 1987 [41] respectively. The adjoint variable methods using index-3 and index-1 DAE formulations were developed by Haug in 1981 [42], Haug in 1987 [41], and Bestle in 1992 [43]. The direct differentiation method using formulations based on velocity transformations was developed by Ashrafiuon and Mani in 1990[44]. Liu, 1996, applied the adjoint variable

method to constrained flexible multi-body systems [45]. The direct differentiation method using the penalty and the augmented Lagrangian formulations was developed by Pagalday in 1997 [46]. Dias and Pereira, 1997, applied the direct differentiation method to rigid-flexible multibody systems [47]. Wang, Haug, and Pan employed the direct differentiation method and implicit Runge-Kutta numerical integration algorithm based on generalized coordinate partitioning to solve the DAE and to perform sensitivity analysis for rigid multibody system in 2005 [48]. Schaffer, 2005, overcame some drawbacks of the adjoint variable method by combining the advantages of the adjoint variable method and the direct differentiation method [49]. Ding, Pan, and Chen applied the adjoint variable method to the second-order sensitivity analysis of DAE formulations in 2007 [50]. Sonneville and Brüls developed the direct differentiation method and the adjoint variable method based on a Lie group formulation in 2013 [51]. For more sensitivity approaches, the reader is referred to [52–78].

Contributions

Although there are so many sensitivity approaches developed in the past 30 years, they all have some drawbacks that prevent them from easily computing sensitivities for large and complex multibody systems with respect to a large number of design parameters. For instance, the direct differentiation method works well when the number of parameters is small, but it becomes computationally expensive when the number of parameters is large. In contrast, the adjoint variable method works well when the number of parameters is large, but it doesn't work well when the number of adjoint variables becomes large. On the other hand, the direct differentiation methods and the adjoint sensitivity methods using index-3 and index-1 DAE formulations are not practical due to the numerical difficulties associated to the underlying formulations.

Thus, the most important task of this study is to overcome those drawbacks and create new algorithms and approaches in order to efficiently perform sensitivity analysis for large and complex systems with respect to a large number of parameters.

Unlike the direct differentiation method, the adjoint variable method works well when the number of parameters is large. On the other hand, solving ordinary differential equations (ODE) is computationally easier than solving DAE. Due to these reasons, the adjoint variable methods using ODE-like formulations become popular.

In this study, we developed a direct differentiation method and an adjoint variable method based on two ODE formulations: Maggi's formulation and the penalty formulation. More specifically, the direct differentiation method is developed with Maggi's formulation, the adjoint variable method is developed with explicit and implicit first order Maggi's formulations, second order Maggi's formulation, and first and second order penalty formulations. These sensitivity approaches are tested and validated by applying them to calculate the sensitivities for a five-bar mechanism and comparing the results with the results generated from other analytical and numerical approaches.

The objective of this study is not only to develop new sensitivity approaches, but also to develop a new approach for large and complex multibody systems. As mentioned before, the adjoint variable methods using ODE-like formulations have several advantages and the adjoint variable methods based on the penalty formulation and Maggi's formulation are developed in this study. Therefore, all the aspects of these two approaches are compared in this study. Finally, among these new sensitivity approaches developed in this study, the adjoint variable method based on the penalty formulation is demonstrated to be able to perform sensitivity analysis for complex multibody systems by applying this approach to a 14-DOF full vehicle.

1.3.2 Dynamic response optimization

State of the art

With the outputs of the sensitivity analysis, many gradient-based optimization packages can be used to perform the dynamic response optimization, most of them are described in

[79]. Dynamic response optimization is defined as the optimization of the dynamic response of multibody systems, which is still an open topic in the dynamics community. More specifically, it means how are the dynamic outputs such as displacement, velocity, acceleration, angle, and frequency, affected by the system parameters, such as geometry, trajectory, inertia parameters, stiffness, and damping coefficient. For a large system, it could be a very complicated problem if a large number of parameters are optimized.

An extensive literature can be found in the area of dynamical optimization. Here only some representative works will be reviewed. One of the initial pieces of work is presented by Besselink and Van Asperen in 1994 [80]. In this paper, they made a point of the importance of sensitivity analysis on the numerical optimization. They optimized the comfort response of a tractor vehicle by using numerical sensitivities. Pagalday and Avello, 1997 [46], used the direct differentiation method based on the penalty formulation to perform sensitivity analysis and optimization. Eberhard, Schiehlen and Bestle, 1999 [81], firstly employed stochastic methods instead of gradient-based methods to search for a global minimum. Andersson and Eriksson, 2004 [55], used Adams to perform optimization of vehicle ride and vehicle handling response of a intercity bus. Gonalves and Ambrósio, 2005 [82], optimized the handling response of a sports car with respect to suspension parameters. In their paper, the flexible chassis was studied to a large extent. In Alfonso's PhD thesis, 2013 [2], optimization of ride response and handling response of a bus were performed. In his thesis, automatic differentiation technology is applied to compute all the derivatives for sensitivity analysis. For more details about dynamic response optimization, the reader is referred to [83–98].

Contributions

One of the purposes of this study is to perform vehicle ride optimization and vehicle handling optimization. There are several ways to improve the handling response and ride comfort of vehicles. For example, the active suspension systems are implemented in some vehicles in order to control the vehicle roll, pitch and bounce motion during different ma-

neuvers. On the other hand, for the passive suspension systems, the suspension parameters are adjusted before vehicles have been manufactured, which does not allow big changes. No matter what kind of suspension system is chosen, it's good to adjust the suspension parameters as much as possible during the design stage. Therefore, the optimization of the dynamic response with respect to suspension parameters becomes very important in the design stage. In addition, for the active suspension system, several methods have been developed to change the dynamic response by changing the suspension parameters, such as [99] and [100]. How to adjust the parameters during different maneuvers? Which parameter has the greatest impact to the dynamic response? How to lower the cost of adjusting the parameters? To answer these questions, optimization with respect to suspension parameters should be performed.

This study focuses mainly on the vehicle handling optimization and vehicle ride optimization of a Bombardier Iltis vehicle model [1] with respect to six suspension parameters, demonstrating the capability of the adjoint variable method using the penalty formulation to perform sensitivity analysis for large and complex multibody systems with respect to multiple design parameters. Double-lane change maneuver, four-post test maneuver, and speed bump test are employed as standard maneuvers for the virtual experiments.

1.4 Multibody dynamics software development

Contributions

There are many commercial packages that can model multibody systems, such as MSC ADAMS, SIMPACK, SimMechanics, LMS VirtualLab Motion, and RecurDyn. One of the drawbacks of these commercial packages is that they don't release the source code to the users. Moreover, commercial multibody packages focus primarily on kinematics and dynamics capabilities, their sensitivity analysis and optimization capabilities are not efficient. There are also some packages developed in academia, but these packages focus on specific

applications and algorithms.

This study developed a modular multibody package MBSVT (Multibody Systems at Virginia Tech) as a software library with forward kinematics and dynamics, sensitivity analysis, and optimization capabilities. MBSVT is a package for education and research, which allows access to the source code for customization. In MBSVT, ewton's method, explicit and implicit Runge-Kutta method are used to perform kinematic analysis and dynamic analysis, respectively. Moreover, the adjoint variable method based on the penalty formulation is employed to calculate the sensitivities. Furthermore, with the outputs of sensitivity analysis, MBSVT uses L-BFGS-B [101], a very popular optimization package, to perform gradient-based optimization. Finally, several different contact and friction models, which can be used to model point contact and surface contact, are developed and included in MBSVT. These models include a static friction model, Ambrósio dry friction model [102], Kelvin-Voigt viscous-elastic model [103], and a simplified tire model.

To show the functionality of the library, the application of MBSVT to a full vehicle and a passive legged robot are discussed in section 5.2 and section 5.3 respectively.

Manual differentiation

Unlike kinematic analysis and dynamic analysis, the computation of sensitivities requires the differentiation of cost function and the EOM, which generates many derivatives. Thus, an efficient, accurate, and general-purpose tool is required to compute these derivatives. At present, the use of symbolic mathematical codes is one of the methods to compute derivatives generated from sensitivity analysis, such as Maple or Mathematica. This method requires the whole codes to use symbolic computation, thus it's not efficient. A better method is to use symbolic codes to derive the expressions for these derivatives. However, sometimes these symbolic codes generate highly complex expressions, complicating the problem. In the last few decades automatic differentiation (AD) tools become popular for the sensitivities computation of multibody systems [104–106]. Alfonso, 2013 [2], proposed automatic differ-

entiation technology to calculate the derivatives for sensitivity analysis in vehicle dynamics problems.

The other method to compute the derivatives is manual differentiation (MD), which is employed in MBSVT. By using manual differentiation, the expressions of all the derivatives are simplified to a very large extent, avoiding complicated expressions generated by symbolic tools and making the computation become very efficient. Complex finite difference method [107] is used to validate all the derivatives before implementing them in the software library, which guarantees the computational accuracy. Both AD and MD are very good methods to compute the derivatives for sensitivity analysis. The comparison of efficiency and accuracy between AD and MD is still an open topic.

Chapter 2

Multibody dynamics formulations

Forward dynamics of multibody systems is a crucial step that provides the basis of sensitivity analysis and dynamical optimization carried out later. In this chapter, several basic formulations are introduced, including DAE formulations with different index and ODE formulations. Direct integration method is presented for DAE formulations.

2.1 Index-3 DAE formulation

The equations of motion of multibody system written in dependent coordinates constitute an index-3 DAE formulation. Assume that the configuration of a multibody system is given by a set of n coordinates $\mathbf{q} \in \mathbb{R}^n$, related by a set of m holonomic constraint equations

$$\boldsymbol{\Phi}\left(t, \mathbf{q}, \boldsymbol{\rho}\right) = \mathbf{0} \in \mathbb{R}^m \tag{2.1}$$

where $\boldsymbol{\rho} \in \mathbb{R}^p$ is a vector of parameters of the system. Some parameters may describe the geometry of the system and therefore affect the constraints (2.1).

The constraint equations (2.1) allow to obtain several kinematic relations. Any virtual displacements of the system coordinates $\delta\mathbf{q}^*$, with the time held fixed, have to satisfy the following equations:

$$\boldsymbol{\Phi}_{\mathbf{q}}\, \delta\mathbf{q}^* = \mathbf{0} \tag{2.2a}$$

The velocities and accelerations of the system have to fulfill the following equations

$$\dot{\mathbf{\Phi}} = \mathbf{\Phi_q}\dot{\mathbf{q}} + \mathbf{\Phi}_t = \mathbf{0} \Rightarrow \mathbf{\Phi_q}\dot{\mathbf{q}} = -\mathbf{\Phi}_t = \mathbf{b} \tag{2.2b}$$

$$\ddot{\mathbf{\Phi}} = \mathbf{\Phi_q}\ddot{\mathbf{q}} + \dot{\mathbf{\Phi}}_\mathbf{q}\dot{\mathbf{q}} + \dot{\mathbf{\Phi}}_t = \mathbf{0} \Rightarrow \mathbf{\Phi_q}\ddot{\mathbf{q}} = -\dot{\mathbf{\Phi}}_\mathbf{q}\dot{\mathbf{q}} - \dot{\mathbf{\Phi}}_t = \mathbf{c} \tag{2.2c}$$

The equations of motion constitute an index-3 system of $n + m$ DAE:

$$\mathbf{M}\ddot{\mathbf{q}} + \mathbf{\Phi}_\mathbf{q}^{\mathrm{T}}\boldsymbol{\lambda} = \mathbf{Q} \tag{2.3a}$$

$$\mathbf{\Phi} = \mathbf{0} \tag{2.3b}$$

where $\mathbf{M} = \mathbf{M}(\mathbf{q}, \boldsymbol{\rho}) \in \mathbb{R}^{n\times n}$ is the mass matrix, $\mathbf{Q} = \mathbf{Q}(t, \mathbf{q}, \dot{\mathbf{q}}, \boldsymbol{\rho}) \in \mathbb{R}^n$ contains the generalized forces and may also include the Coriolis and centrifugal effects (if the formulation needs them), $\mathbf{\Phi_q} \in \mathbb{R}^{m\times n}$ is the Jacobian matrix of the constraints (2.3b), and $\boldsymbol{\lambda} \in \mathbb{R}^m$ are the Lagrange multipliers associated with the constraints.

The direct numerical integration of the index-3 DAE EOM poses a number of numerical difficulties, including ill-conditioning for small time steps and instability problems that make the direct solution of the equations not recommendable in general.

The instability problems come from the fact that only the position level constraints are considered in the formulation and this causes an unstable behavior when integrated with the classical time-stepping schemes. The explanation for it is that the EOM only impose the satisfaction of the constraints themselves, but no integrator can automatically guarantee that the velocities and the accelerations will remain onto their respective manifolds of the constraints derivatives since they are weak invariants of the EOM [108]. To avoid this problem the authors recommend to combine the direct integration of the equations (2.3) with the use of projection techniques like proposed in [109] or to reformulate the problem as an augmented Lagrangian approach with projections like in [110], to solve the forward dynamics.

The ill-conditioning problem was also addressed in the past by several authors and it

will be presented in this section particularized for an integrator belonging to the Newmark family. It will be illustrative since the sensitivity equations developed later in this study will inherit the same issue. Let

$$\mathbf{f} = \begin{bmatrix} \mathbf{M}\ddot{\mathbf{q}} + \mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}}\boldsymbol{\lambda} - \mathbf{Q} \\ \mathbf{\Phi} \end{bmatrix} \tag{2.4}$$

be the residual of the EOM (2.3), meaning that equations (2.4) have to be equal to zero.

Using an implicit integrator (e.g., trapezoidal rule) to integrate the previous equations as described in [109]

$$\dot{\mathbf{q}}_{n+1} = \frac{2}{h}\mathbf{q}_{n+1} + \hat{\dot{\mathbf{q}}}_n \; ; \quad \hat{\dot{\mathbf{q}}}_n = -\left(\frac{2}{h}\mathbf{q}_n + \dot{\mathbf{q}}_n\right) \tag{2.5a}$$

$$\ddot{\mathbf{q}}_{n+1} = \frac{4}{h^2}\mathbf{q}_{n+1} + \hat{\ddot{\mathbf{q}}}_n \; ; \quad \hat{\ddot{\mathbf{q}}}_n = -\left(\frac{4}{h^2}\mathbf{q}_n + \frac{4}{h}\dot{\mathbf{q}}_n + \ddot{\mathbf{q}}_n\right) \tag{2.5b}$$

where n is the time step index and h the time step.

Replacing equations (2.5) in (2.4), a nonlinear system of algebraic equations in $n+1$ is obtained that can be solved using following Newton's method

$$\left[\frac{\partial \mathbf{f}}{\partial \mathbf{y}}\right]^{(i)} \Delta\mathbf{y}_{n+1}^{(i+1)} = -\mathbf{f}^{(i)} \tag{2.6a}$$

$$\left[\frac{\partial \mathbf{f}}{\partial \mathbf{y}}\right] = \begin{bmatrix} \dfrac{4}{h^2}\mathbf{M} + \dfrac{2}{h}\mathbf{C} + \mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} + \mathbf{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\boldsymbol{\lambda} + \mathbf{K} & \mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}} \\ \mathbf{\Phi}_{\mathbf{q}} & \mathbf{0} \end{bmatrix} \tag{2.6b}$$

where $\mathbf{y} = \begin{bmatrix} \mathbf{q} & \boldsymbol{\lambda} \end{bmatrix}^{\mathrm{T}}$.

The solution of equations (2.6a) for small time steps poses severe issues: it was reported in [111] that the propagation of errors in the solution of the Lagrange multipliers is of order $\mathcal{O}\left(h^{-2}\right)$ and the condition number of the tangent matrix (2.6b) is of order $\mathcal{O}\left(h^{-4}\right)$ (ill conditioned for small time steps). Several authors proposed the scaling of the equations to alleviate these problems [111–113].

In [111] a specific scaling for equations (2.6a) with Newmark integrators is proposed, which can be easily particularized for the particular case of the trapezoidal rule by scaling the first n equations in the residual (2.4) and the Lagrange multipliers by a factor of $h^2/4$, leading to the following scaled equations and states

$$\left[\frac{\partial \bar{\mathbf{f}}}{\partial \bar{\mathbf{y}}}\right]^{(i)} \Delta \bar{\mathbf{y}}_{n+1}^{(i+1)} = -\bar{\mathbf{f}}^{(i)} \tag{2.7a}$$

$$\bar{\mathbf{f}} = \begin{bmatrix} \frac{h^2}{4}\left(\mathbf{M}\ddot{\mathbf{q}} + \mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}}\boldsymbol{\lambda} - \mathbf{Q}\right) \\ \mathbf{\Phi} \end{bmatrix} \tag{2.7b}$$

$$\left[\frac{\partial \bar{\mathbf{f}}}{\partial \bar{\mathbf{y}}}\right] = \begin{bmatrix} \mathbf{M} + \frac{h}{2}\mathbf{C} + \frac{h^2}{4}\left(\mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} + \mathbf{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\boldsymbol{\lambda} + \mathbf{K}\right) & \mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}} \\ \mathbf{\Phi}_{\mathbf{q}} & \mathbf{0} \end{bmatrix} \tag{2.7c}$$

where $\bar{\mathbf{y}} = \begin{bmatrix} \mathbf{q} & \bar{\boldsymbol{\lambda}} \end{bmatrix}^{\mathrm{T}}$ are the scaled states and $\bar{\boldsymbol{\lambda}} = \left(h^2/4\right)\boldsymbol{\lambda}$ are the scaled Lagrange multipliers.

The suggested scaling leads both the propagation of errors in the solution of the Lagrange multipliers and the condition number of the tangent matrix (2.6b) to order $\mathcal{O}\left(h^0\right)$.

2.2 Index-1 DAE formulation

The difficulties mentioned before to numerically solve equations (2.3) recommend reformulating the problem to formulations that are easier to solve. Index reduction is a common technique [11]. Differentiating (2.3b) twice leads to the following index-1 DAE system:

$$\mathbf{M}\ddot{\mathbf{q}} + \mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}}\boldsymbol{\lambda} = \mathbf{Q} \tag{2.8a}$$

$$\mathbf{\Phi}_{\mathbf{q}}\ddot{\mathbf{q}} = -\dot{\mathbf{\Phi}}_{\mathbf{q}}\dot{\mathbf{q}} - \dot{\mathbf{\Phi}}_{t} = \mathbf{c} \tag{2.8b}$$

The direct numerical integration of the index-1 DAE equations follows the same steps as we did in the previous section. Let

$$\mathbf{f} = \begin{bmatrix} \mathbf{M}\ddot{\mathbf{q}} + \mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}}\boldsymbol{\lambda} - \mathbf{Q} \\ \mathbf{\Phi}_{\mathbf{q}}\ddot{\mathbf{q}} - \mathbf{c} \end{bmatrix} \tag{2.9}$$

be the residual of the EOM (2.8), meaning that equations (2.9) have to be equal to zero. Replacing equations (2.5) in (2.9), a nonlinear system of algebraic equations in $n+1$ is obtained that can be solved using following Newton's method

$$\left[\frac{\partial \mathbf{f}}{\partial \mathbf{y}}\right]^{(i)} \Delta \mathbf{y}_{n+1}^{(i+1)} = -\mathbf{f}^{(i)} \tag{2.10a}$$

$$\left[\frac{\partial \mathbf{f}}{\partial \mathbf{y}}\right] = \begin{bmatrix} \frac{4}{h^2}\mathbf{M} + \frac{2}{h}\mathbf{C} + \mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} + \mathbf{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\boldsymbol{\lambda} + \mathbf{K} & \mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}} \\ \frac{4}{h^2}\mathbf{\Phi}_{\mathbf{q}} - \frac{2}{h}\mathbf{c}_{\dot{\mathbf{q}}} + \mathbf{\Phi}_{\mathbf{qq}}\ddot{\mathbf{q}} - \mathbf{c}_{\mathbf{q}} & \mathbf{0} \end{bmatrix} \tag{2.10b}$$

where $\mathbf{y} = \begin{bmatrix} \mathbf{q} & \boldsymbol{\lambda} \end{bmatrix}^{\mathrm{T}}$.

Scaling the first n equations in the residual (2.9) and the Lagrange multipliers by a factor of $h^2/4$, the following scaled equations and states can be solved

$$\left[\frac{\partial \bar{\mathbf{f}}}{\partial \bar{\mathbf{y}}}\right]^{(i)} \Delta \bar{\mathbf{y}}_{n+1}^{(i+1)} = -\bar{\mathbf{f}}^{(i)} \tag{2.11a}$$

$$\bar{\mathbf{f}} = \begin{bmatrix} \frac{h^2}{4}\left(\mathbf{M}\ddot{\mathbf{q}} + \mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}}\boldsymbol{\lambda} - \mathbf{Q}\right) \\ \mathbf{c} \end{bmatrix} \tag{2.11b}$$

$$\left[\frac{\partial \bar{\mathbf{f}}}{\partial \bar{\mathbf{y}}}\right] = \begin{bmatrix} \mathbf{M} + \frac{h}{2}\mathbf{C} + \frac{h^2}{4}\left(\mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} + \mathbf{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\boldsymbol{\lambda} + \mathbf{K}\right) & \mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}} \\ \frac{4}{h^2}\mathbf{\Phi}_{\mathbf{q}} - \frac{2}{h}\mathbf{c}_{\dot{\mathbf{q}}} + \mathbf{\Phi}_{\mathbf{qq}}\ddot{\mathbf{q}} - \mathbf{c}_{\mathbf{q}} & \mathbf{0} \end{bmatrix} \tag{2.11c}$$

where $\bar{\mathbf{y}} = \begin{bmatrix} \mathbf{q} & \bar{\boldsymbol{\lambda}} \end{bmatrix}^{\mathrm{T}}$ are the scaled states and $\bar{\boldsymbol{\lambda}} = \left(h^2/4\right)\boldsymbol{\lambda}$ are the scaled Lagrange multipliers.

2.3 Baumgarte stabilization method

A direct numerical solution of (2.8) suffers from drift-off [19], meaning that any small perturbation in the acceleration constraints leads to an error in the position constraints that grows quadratically with time: $\ddot{\mathbf{\Phi}} = \epsilon_1 \Rightarrow \dot{\mathbf{\Phi}} = \epsilon_1 t + \epsilon_2 \Rightarrow \mathbf{\Phi} = \epsilon_1 t^2/2 + \epsilon_2 t + \epsilon_3$.

Baumgarte, 1972, presented a stabilization method [20] that replaces the differential constraint

equations (2.8b) by the following equations:

$$\ddot{\mathbf{\Phi}} + 2\alpha\dot{\mathbf{\Phi}} + \beta^2\mathbf{\Phi} = \mathbf{0} \tag{2.12}$$

where α and β are positive constants. Assume the general solution of (2.12) is:

$$\mathbf{\Phi} = \mathbf{a}_1 e^{b_1 t} + \mathbf{a}_2 e^{b_2 t} \tag{2.13}$$

where $\mathbf{a}_1$ and $\mathbf{a}_2$ are constant vectors that depend on the initial conditions, b_1 and b_2 are the unknowns. Replace (2.13) in (2.12), we get the expressions of b_1 and b_2:

$$b_1, b_2 = -\alpha \pm \sqrt{\alpha^2 + \beta^2} \tag{2.14}$$

Since α and β are positive constants, the real part of b_1 and b_2 is negative, which makes $\mathbf{\Phi}$ converge to $\mathbf{0}$ as t grows up, stabilizing the general solutions.

After replacing (2.8b) with (2.12), the index-1 DAE system (2.8) are transformed into:

$$\mathbf{M}\ddot{\mathbf{q}} + \mathbf{\Phi}_\mathbf{q}^\mathrm{T}\boldsymbol{\lambda} = \mathbf{Q} \tag{2.15a}$$

$$\mathbf{\Phi}_\mathbf{q}\ddot{\mathbf{q}} = \mathbf{c} - 2\alpha\dot{\mathbf{\Phi}} - \beta^2\mathbf{\Phi} \tag{2.15b}$$

The direct numerical integration of this system is similar to what we did in the previous section.

2.4 The penalty formulation

In [21] the penalty formulation of multibody dynamics is introduced. The equations constitute a system of ODE of dimension n:

$$\mathbf{M}\ddot{\mathbf{q}} + \mathbf{\Phi}_\mathbf{q}^\mathrm{T}\boldsymbol{\alpha}\left(\ddot{\mathbf{\Phi}} + 2\xi\omega\dot{\mathbf{\Phi}} + \omega^2\mathbf{\Phi}\right) = \mathbf{Q} \tag{2.16}$$

where $\boldsymbol{\alpha}$ is the penalty matrix and ξ, ω are coefficients of the method. An equivalent formulation is:

$$\bar{\mathbf{M}}\ddot{\mathbf{q}} = \bar{\mathbf{Q}} \tag{2.17a}$$

$$\bar{\mathbf{M}} \equiv \left(\mathbf{M} + \mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}}\boldsymbol{\alpha}\mathbf{\Phi}_{\mathbf{q}}\right) \tag{2.17b}$$

$$\bar{\mathbf{Q}} \equiv \mathbf{Q} - \mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}}\boldsymbol{\alpha}\left(\dot{\mathbf{\Phi}}_{\mathbf{q}}\dot{\mathbf{q}} + \dot{\mathbf{\Phi}}_t + 2\xi\omega\dot{\mathbf{\Phi}} + \omega^2\mathbf{\Phi}\right) \tag{2.17c}$$

where the following kinematic identities hold

$$\dot{\mathbf{\Phi}} = \mathbf{\Phi}_{\mathbf{q}}\dot{\mathbf{q}} + \mathbf{\Phi}_t \tag{2.18}$$

$$\ddot{\mathbf{\Phi}} = \mathbf{\Phi}_{\mathbf{q}}\ddot{\mathbf{q}} + \dot{\mathbf{\Phi}}_{\mathbf{q}}\dot{\mathbf{q}} + \dot{\mathbf{\Phi}}_t \tag{2.19}$$

Comparing (2.16) and (2.3a) it can be seen that the penalty equations approximate the Lagrange multipliers with the following term

$$\boldsymbol{\lambda}^* = \boldsymbol{\alpha}\left(\ddot{\mathbf{\Phi}} + 2\xi\omega\dot{\mathbf{\Phi}} + \omega^2\mathbf{\Phi}\right) \tag{2.20}$$

where the star means approximated or fictitious Lagrange multipliers, as opposed to the real ones that arise in DAE formulations.

Assuming that $\boldsymbol{\alpha}$ is a diagonal matrix with the penalty factors for each constraint on the diagonal, each entry of $\boldsymbol{\lambda}^*$ in (2.20) is the equation of a one-DOF oscillatory system. The coefficients of the oscillatory system are usually selected as $\omega = 10$ and $\xi = 1$, which corresponds to a critically damped system.

The penalty equations (2.16) are equivalent to the original DAE for infinite penalty factors. In floating point computing loss-of-significance errors appear in the first parenthesis of equation (2.17a) for large penalties; thus, the selection of the penalty factors is a sensitive issue that the analyst has to solve. In practice, this formulation does not satisfy exactly either constraint equation ($\mathbf{\Phi} = \mathbf{0}$, $\dot{\mathbf{\Phi}} = \mathbf{0}$, or $\ddot{\mathbf{\Phi}} = \mathbf{0}$) but it approximately satisfies the equation $\ddot{\mathbf{\Phi}} + 2\xi\omega\dot{\mathbf{\Phi}} + \omega^2\mathbf{\Phi} = \mathbf{0}$.

2.5 Maggi's formulation

All the formulations introduced above are in dependent coordinates. Maggi's formulation [17], which can be also called state-space formulation based on the projection Matrix R, is one of the well-known formulation in independent coordinates. The starting point is the d'Alembert's principle in dependent coordinates

$$\delta\mathbf{q}^{*\mathrm{T}}\left(\mathbf{M}\left(\mathbf{q},\boldsymbol{\rho}\right)\ddot{\mathbf{q}}-\mathbf{Q}\left(\mathbf{q},\dot{\mathbf{q}},t,\boldsymbol{\rho}\right)\right)=\mathbf{0} \tag{2.21}$$

where $\delta\mathbf{q}^*$ constitutes a set of n dependent virtual displacements, and the rest of the terms are described in the nomenclature.

In addition to Eqn. (2.21), the actual positions, velocities, accelerations, and the virtual displacements have to fulfill the constraint equations and their derivatives, i.e., the following kinematic relations:

$$\boldsymbol{\Phi}\left(\mathbf{q},t,\boldsymbol{\rho}\right)=\mathbf{0} \tag{2.22a}$$

$$\boldsymbol{\Phi}_{\mathbf{q}}\dot{\mathbf{q}}=-\boldsymbol{\Phi}_t=\mathbf{b} \tag{2.22b}$$

$$\boldsymbol{\Phi}_{\mathbf{q}}\ddot{\mathbf{q}}=-\dot{\boldsymbol{\Phi}}_{\mathbf{q}}\dot{\mathbf{q}}-\dot{\boldsymbol{\Phi}}_t=\mathbf{c} \tag{2.22c}$$

$$\boldsymbol{\Phi}_{\mathbf{q}}\delta\mathbf{q}^*=\mathbf{0} \tag{2.22d}$$

where the derivatives of the constraints vector are calculated as described in the nomenclature.

Note that the EOM (2.21) and the constraints (2.22) are dependent on some design parameters $\boldsymbol{\rho}\in\mathbb{R}^p$ (typically masses, lengths, or other parameters as selected by engineers).

Consider the matrices $\mathbf{R}\in\mathbb{R}^{n\times d}$ and $\mathbf{S}\in\mathbb{R}^{n\times m}$. As it was indicated in the nomenclature, n is the number of dependent coordinates, m is the number of constraints and d is the number of *DOF* of the system. Without loss of generality, the constraints are assumed to be independent and therefore $n=m+d$. The dependent velocities, accelerations, and virtual displacements can

be expressed in terms of the independent *DOF* by means of these matrices, as

$$\dot{\mathbf{q}} = \mathbf{R}\dot{\mathbf{z}} + \mathbf{S}\mathbf{b} \tag{2.23a}$$

$$\ddot{\mathbf{q}} = \mathbf{R}\ddot{\mathbf{z}} + \mathbf{S}\mathbf{c} \tag{2.23b}$$

$$\delta\mathbf{q}^* = \mathbf{R}\delta\mathbf{z}^* \tag{2.23c}$$

Using (2.23) in (2.21), and taking into account that the virtual displacements $\delta\mathbf{z}^*$ are independent, one obtains

$$\left(\mathbf{R}^{\mathrm{T}}\mathbf{M}\mathbf{R}\right)\ddot{\mathbf{z}} = \mathbf{R}^{\mathrm{T}}\left(\mathbf{Q} - \mathbf{M}\mathbf{S}\mathbf{c}\right) \tag{2.24}$$

or, in a more compact form

$$\bar{\mathbf{M}}\left(\mathbf{z}, \boldsymbol{\rho}\right)\ddot{\mathbf{z}} = \bar{\mathbf{Q}}\left(\mathbf{z}, \dot{\mathbf{z}}, t, \boldsymbol{\rho}\right) \tag{2.25a}$$

$$\bar{\mathbf{M}}\left(\mathbf{z}, \boldsymbol{\rho}\right) = \mathbf{R}^{\mathrm{T}}\mathbf{M}\mathbf{R} \tag{2.25b}$$

$$\bar{\mathbf{Q}}\left(\mathbf{z}, \dot{\mathbf{z}}, t, \boldsymbol{\rho}\right) = \mathbf{R}^{\mathrm{T}}\left(\mathbf{Q} - \mathbf{M}\mathbf{S}\mathbf{c}\right) \tag{2.25c}$$

Equations (2.25) constitute a second-order state-space ODE. Note that the mass matrix and generalized forces vector in (2.25) are dependent on the design parameter set $\boldsymbol{\rho}$, and therefore $\mathbf{z} = \mathbf{z}\left(t, \boldsymbol{\rho}\right)$, $\dot{\mathbf{z}} = \dot{\mathbf{z}}\left(t, \boldsymbol{\rho}\right)$, $\ddot{\mathbf{z}} = \ddot{\mathbf{z}}\left(t, \boldsymbol{\rho}\right)$.

The matrices $\mathbf{R}$ and $\mathbf{S}$ can be calculated in different ways. One possibility is to write the following kinematic systems of equations derived from (2.22)

$$\begin{bmatrix} \boldsymbol{\Phi}_{\mathbf{q}} \\ \mathbf{B} \end{bmatrix} \dot{\mathbf{q}} = \begin{bmatrix} \mathbf{b} \\ \dot{\mathbf{z}} \end{bmatrix} \tag{2.26a}$$

$$\begin{bmatrix} \boldsymbol{\Phi}_{\mathbf{q}} \\ \mathbf{B} \end{bmatrix} \ddot{\mathbf{q}} = \begin{bmatrix} \mathbf{c} \\ \ddot{\mathbf{z}} \end{bmatrix} \tag{2.26b}$$

$$\begin{bmatrix} \boldsymbol{\Phi}_{\mathbf{q}} \\ \mathbf{B} \end{bmatrix} \delta\mathbf{q}^* = \begin{bmatrix} \mathbf{0} \\ \delta\mathbf{z}^* \end{bmatrix} \tag{2.26c}$$

where $\mathbf{B}$ is a matrix composed of zeros and ones that imposes a set of additional constraints over a subset of independent coordinates $\mathbf{z}$, selected by the user as *DOF* from the entries of $\mathbf{q}$. Thus,

Eq. (2.23) are directly obtained from (2.26) by inverting and partitioning the leading matrix

$$\begin{bmatrix} \mathbf{S} & \mathbf{R} \end{bmatrix} = \begin{bmatrix} \mathbf{\Phi_q} \\ \mathbf{B} \end{bmatrix}^{-1} \Rightarrow \begin{bmatrix} \mathbf{\Phi_q} \\ \mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{S} & \mathbf{R} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_m & \mathbf{0}_{m\times d} \\ \mathbf{0}_{d\times m} & \mathbf{I}_d \end{bmatrix} \tag{2.27}$$

Here $\mathbf{I}$ stands for the identity matrix and $\mathbf{0}$ for a matrix with all of its elements equal to zero. The matrices $\mathbf{S}$ and $\mathbf{R}$ are the first m and the last d columns of the inverse matrix, respectively. Since the set of constraints is assumed to be independent $\mathbf{\Phi_q}$ is a $m \times n$ matrix with rank m and the full leading matrix is always invertible, provided that a valid set of d *DOF* are selected via $\mathbf{B}$. For selecting a proper set of *DOF* and for algebraic techniques to deal with redundant constraints, the reader is referred to [114].

The explicit calculation of matrix $\mathbf{S}$ is rarely needed for the forward dynamics so, from the computational point of view, the calculation of the inverse in (2.27) could be avoided for that case. Nevertheless, it will be shown next that the matrix $\mathbf{S}$ plays an important role in the sensitivities calculation.

Chapter 3

Sensitivity analysis of multibody systems

In this chapter, we bring new contributions to the state-of-the-art in analytical approaches to perform sensitivity analysis of multibody systems. A direct differentiation method and an adjoint variable method are developed in the context of several different multibody dynamics formulations: index-1 and index-3 DAE formulations, the penalty formulation and Maggi's formulation. Direct integration method is incorporated into the direct differentiation methods with index-1 and index-3 DAE formulations. The computed sensitivities are validated by finite difference methods. Finally ,the resulting sensitivities will be applied to perform dynamical optimization of different multibody systems in the next chapter.

3.1 Direct differentiation method

In this section, the sensitivity equations for the EOM presented in chapter 2 are derived based on the direct differentiation method.

Consider the case where the EOM dependent on the vector of parameters $\boldsymbol{\rho} \in \mathbb{R}^p$. The following objective function is defined in terms of the parameters, on the states $\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}} \in \mathbb{R}^n$, and on the Lagrange multipliers $\boldsymbol{\lambda} \in \mathbb{R}^m$

$$\psi = w\left(\mathbf{q}_F, \dot{\mathbf{q}}_F, \ddot{\mathbf{q}}_F, \boldsymbol{\rho}_F, \boldsymbol{\lambda}_F\right) + \int_{t_0}^{t_F} g\left(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}}, \boldsymbol{\lambda}, \boldsymbol{\rho}\right) \mathrm{dt} \tag{3.1}$$

where the subindex F means evaluation at the final time t_F. The sensitivity analysis techniques

discussed herein will evaluate the gradient of the objective function with respect to parameters:

$$\nabla_{\boldsymbol{\rho}}\psi = (d\psi/d\boldsymbol{\rho})^{\mathrm{T}} \tag{3.2}$$

3.1.1 Index-3 DAE formulation

The direct differentiation method for the sensitivity analysis using the index-3 formulation in chapter 2 was developed in [41] for objective functions dependent on $\mathbf{q}, \dot{\mathbf{q}}, \boldsymbol{\rho}, \boldsymbol{\lambda}$ and variable time limits t_F. In this section, the results are revisited also for objective functions depending on $\ddot{\mathbf{q}}$ but fixed end time t_F.

Differentiating (3.1)

$$\nabla_{\boldsymbol{\rho}}\psi^{\mathrm{T}} = (w_{\mathbf{q}}\mathbf{q}_{\boldsymbol{\rho}} + w_{\dot{\mathbf{q}}}\dot{\mathbf{q}}_{\boldsymbol{\rho}} + w_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}} + w_{\boldsymbol{\lambda}}\boldsymbol{\lambda}_{\boldsymbol{\rho}} + w_{\boldsymbol{\rho}})_F + \int_{t_0}^{t_F} (g_{\mathbf{q}}\mathbf{q}_{\boldsymbol{\rho}} + g_{\dot{\mathbf{q}}}\dot{\mathbf{q}}_{\boldsymbol{\rho}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}} + g_{\boldsymbol{\lambda}}\boldsymbol{\lambda}_{\boldsymbol{\rho}} + g_{\boldsymbol{\rho}})\,\mathrm{dt} \tag{3.3}$$

In equation (3.3) the derivatives of functions w and g are known, since the objective function has a known expression. The derivatives $\mathbf{q}_{\boldsymbol{\rho}}$, $\dot{\mathbf{q}}_{\boldsymbol{\rho}}$, $\ddot{\mathbf{q}}_{\boldsymbol{\rho}}$ and $\boldsymbol{\lambda}_{\boldsymbol{\rho}}$ are the sensitivities of the solution of the dynamical system. These can be obtained by differentiating (2.3) with respect to each one of the parameters:

$$\frac{\mathrm{d}\mathbf{M}}{\mathrm{d}\rho_k}\ddot{\mathbf{q}} + \mathbf{M}\frac{\partial\ddot{\mathbf{q}}}{\partial\rho_k} + \frac{\mathrm{d}\boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}}{\mathrm{d}\rho_k}\boldsymbol{\lambda} + \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\frac{\partial\boldsymbol{\lambda}}{\partial\rho_k} = \frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}\rho_k} \tag{3.4}$$

$$\frac{\mathrm{d}\boldsymbol{\Phi}}{\mathrm{d}\rho_k} = \mathbf{0}, \quad k = 1, \ldots, m \tag{3.5}$$

Expanding the total derivatives and grouping them together in matrix notation, leads to the following set of p DAE, each one of them is called a tangent linear model (TLM):

$$\mathbf{M}\ddot{\mathbf{q}}_{\boldsymbol{\rho}} + \mathbf{C}\dot{\mathbf{q}}_{\boldsymbol{\rho}} + \left(\mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\boldsymbol{\lambda} + \mathbf{K}\right)\mathbf{q}_{\boldsymbol{\rho}} + \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\boldsymbol{\lambda}_{\boldsymbol{\rho}} = \mathbf{Q}_{\boldsymbol{\rho}} - \mathbf{M}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} - \boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}^{\mathrm{T}}\boldsymbol{\lambda} \tag{3.6a}$$

$$\boldsymbol{\Phi}_{\mathbf{q}}\mathbf{q}_{\boldsymbol{\rho}} = -\boldsymbol{\Phi}_{\boldsymbol{\rho}} \tag{3.6b}$$

where $\mathbf{K} = -\mathbf{Q}_{\mathbf{q}}$, $\mathbf{C} = -\mathbf{Q}_{\dot{\mathbf{q}}}$, and the following terms are tensor-vector products: $\mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} \equiv \mathbf{M}_{\mathbf{q}} \otimes \ddot{\mathbf{q}}$, $\boldsymbol{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\boldsymbol{\lambda} \equiv \boldsymbol{\Phi}_{\mathbf{qq}}^{\mathrm{T}} \otimes \boldsymbol{\lambda}$, $\mathbf{M}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} \equiv \mathbf{M}_{\boldsymbol{\rho}} \otimes \ddot{\mathbf{q}}$, $\boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}^{\mathrm{T}}\boldsymbol{\lambda} \equiv \boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}^{\mathrm{T}} \otimes \boldsymbol{\lambda}$.

The TLM (3.6) needs for the following $2np$ initial conditions

$$\mathbf{q}_{\boldsymbol{\rho}}\left(t_0\right) = \mathbf{q}_{\boldsymbol{\rho}0} \tag{3.7a}$$

$$\dot{\mathbf{q}}_{\boldsymbol{\rho}}\left(t_0\right) = \dot{\mathbf{q}}_{\boldsymbol{\rho}0} \tag{3.7b}$$

The initial conditions (3.7) are not independent, since they have to satisfy the following constraint equations

$$\frac{\mathrm{d}\boldsymbol{\Phi}\left(t_0\right)}{\mathrm{d}\boldsymbol{\rho}} = \mathbf{0} \quad \rightarrow \quad \left[\boldsymbol{\Phi}_{\mathbf{q}}\mathbf{q}_{\boldsymbol{\rho}}\right]_0 = -\boldsymbol{\Phi}_{\boldsymbol{\rho}0} \tag{3.8a}$$

$$\frac{\mathrm{d}\dot{\boldsymbol{\Phi}}\left(t_0\right)}{\mathrm{d}\boldsymbol{\rho}} = \mathbf{0} \quad \rightarrow \quad \left[\boldsymbol{\Phi}_{\mathbf{q}}\dot{\mathbf{q}}_{\boldsymbol{\rho}}\right]_0 = -\left[\left(\boldsymbol{\Phi}_{\mathbf{qq}}\dot{\mathbf{q}} + \boldsymbol{\Phi}_{t\mathbf{q}}\right)\mathbf{q}_{\boldsymbol{\rho}} + \boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}\dot{\mathbf{q}} + \boldsymbol{\Phi}_{t\boldsymbol{\rho}}\right]_0 \tag{3.8b}$$

where the sub-index 0 means evaluation at the initial time t_0, as indicated in the nomenclature.

Consequently, $n - \mathrm{rank}\left(\boldsymbol{\Phi}_{\mathbf{q}}\right)$ independent sensitivities can be chosen from (3.8a) and $n - \mathrm{rank}\left(\boldsymbol{\Phi}_{\mathbf{q}}\right)$ independent "velocity" sensitivities from (3.8b). That means that the impact of the parameters on the initial configuration of the system, by means of a subset of *DOF*, can be decided as an input to the problem.

The TLM DAE in equations (3.6) can be directly integrated in the same way that the EOM in section 2.1. The trapezoidal rule equations for the sensitivities

$$\dot{\mathbf{q}}_{\boldsymbol{\rho}}^{n+1} = \frac{2}{h}\mathbf{q}_{\boldsymbol{\rho}}^{n+1} + \hat{\dot{\mathbf{q}}}_{\boldsymbol{\rho}}^{n}\,; \quad \hat{\dot{\mathbf{q}}}_{\boldsymbol{\rho}}^{n} = -\left(\frac{2}{h}\mathbf{q}_{\boldsymbol{\rho}}^{n} + \dot{\mathbf{q}}_{\boldsymbol{\rho}}^{n}\right) \tag{3.9a}$$

$$\ddot{\mathbf{q}}_{\boldsymbol{\rho}}^{n+1} = \frac{4}{h^2}\mathbf{q}_{\boldsymbol{\rho}}^{n+1} + \hat{\ddot{\mathbf{q}}}_{\boldsymbol{\rho}}^{n}\,; \quad \hat{\ddot{\mathbf{q}}}_{\boldsymbol{\rho}}^{n} = -\left(\frac{4}{h^2}\mathbf{q}_{\boldsymbol{\rho}}^{n} + \frac{4}{h}\dot{\mathbf{q}}_{\boldsymbol{\rho}}^{n} + \ddot{\mathbf{q}}_{\boldsymbol{\rho}}^{n}\right) \tag{3.9b}$$

where the superindex n means time step. Replacing the integrator equations in (3.6) and solving for $\mathbf{q}_{\boldsymbol{\rho}}^{n+1}$ and $\boldsymbol{\lambda}^{n+1}$

$$\begin{bmatrix} \dfrac{4}{h^2}\mathbf{M} + \dfrac{2}{h}\mathbf{C} + \mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\boldsymbol{\lambda} + \mathbf{K} & \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}} \\ \boldsymbol{\Phi}_{\mathbf{q}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{q}_{\boldsymbol{\rho}}^{n+1} \\ \boldsymbol{\lambda}_{\boldsymbol{\rho}}^{n+1} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_{\boldsymbol{\rho}} - \mathbf{M}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} - \boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}^{\mathrm{T}}\boldsymbol{\lambda} - \mathbf{M}\hat{\ddot{\mathbf{q}}}_{\boldsymbol{\rho}}^{n} - \mathbf{C}\hat{\dot{\mathbf{q}}}_{\boldsymbol{\rho}}^{n} \\ -\boldsymbol{\Phi}_{\boldsymbol{\rho}} \end{bmatrix} \tag{3.10}$$

Observe that the leading matrix in equation (3.10) is identical to the tangent matrix (2.6b)

and therefore exactly the same scaling applies here

$$\begin{bmatrix} \mathbf{M} + \frac{h}{2}\mathbf{C} + \frac{h^2}{4}\left(\mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} + \mathbf{\Phi}^{\mathrm{T}}_{\mathbf{qq}}\boldsymbol{\lambda} + \mathbf{K}\right) & \mathbf{\Phi}^{\mathrm{T}}_{\mathbf{q}} \\ \mathbf{\Phi}_{\mathbf{q}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{q}^{n+1}_{\boldsymbol{\rho}} \\ \bar{\boldsymbol{\lambda}}^{n+1}_{\boldsymbol{\rho}} \end{bmatrix} = \begin{bmatrix} \frac{h^2}{4}\left(\mathbf{Q}_{\boldsymbol{\rho}} - \mathbf{M}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} - \mathbf{\Phi}^{\mathrm{T}}_{\mathbf{q}\boldsymbol{\rho}}\boldsymbol{\lambda} - \mathbf{M}\hat{\ddot{\mathbf{q}}}^{n}_{\boldsymbol{\rho}} - \mathbf{C}\hat{\dot{\mathbf{q}}}^{n}_{\boldsymbol{\rho}}\right) \\ -\mathbf{\Phi}_{\boldsymbol{\rho}} \end{bmatrix} \tag{3.11}$$

where $\bar{\boldsymbol{\lambda}}_{\boldsymbol{\rho}} = \left(h^2/4\right)\boldsymbol{\lambda}_{\boldsymbol{\rho}}$ are the Lagrange multipliers scaled sensitivities.

Observe that the systems (3.11) are cheap to solve, since they don't need to be iterated and the previous factorization of the leading matrix, from the dynamics, can be employed to solve it.

3.1.2 Index-1 DAE formulation

The direct differentiation method is now applied to the index-1 formulation discussed in chapter 2. The gradient of the objective function is given again by (3.3) and the derivatives $\mathbf{q}_{\boldsymbol{\rho}}$, $\dot{\mathbf{q}}_{\boldsymbol{\rho}}$, $\ddot{\mathbf{q}}_{\boldsymbol{\rho}}$ and $\boldsymbol{\lambda}_{\boldsymbol{\rho}}$ are the sensitivities of the solution of the dynamical equations (2.8). These sensitivities are obtained differentiating (2.8) with respect to each one of the parameters as follows:

$$\frac{\mathrm{d}\mathbf{M}}{\mathrm{d}\rho_k}\ddot{\mathbf{q}} + \mathbf{M}\frac{\partial\ddot{\mathbf{q}}}{\partial\rho_k} + \frac{\mathrm{d}\mathbf{\Phi}^{\mathrm{T}}_{\mathbf{q}}}{\mathrm{d}\rho_k}\boldsymbol{\lambda} + \mathbf{\Phi}^{\mathrm{T}}_{\mathbf{q}}\frac{\partial\boldsymbol{\lambda}}{\partial\rho_k} = \frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}\rho_k} \tag{3.12a}$$

$$\frac{\mathrm{d}\ddot{\mathbf{\Phi}}}{\mathrm{d}\rho_k} = \mathbf{0}, \quad k = 1, \ldots, p \tag{3.12b}$$

Expanding the total derivatives and grouping them together in matrix notation, leads to the following TLM DAE system:

$$\mathbf{M}\ddot{\mathbf{q}}_{\boldsymbol{\rho}} + \mathbf{C}\dot{\mathbf{q}}_{\boldsymbol{\rho}} + \left(\mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} + \mathbf{\Phi}^{\mathrm{T}}_{\mathbf{qq}}\boldsymbol{\lambda} + \mathbf{K}\right)\mathbf{q}_{\boldsymbol{\rho}} + \mathbf{\Phi}^{\mathrm{T}}_{\mathbf{q}}\boldsymbol{\lambda}_{\boldsymbol{\rho}} = \mathbf{Q}_{\boldsymbol{\rho}} - \mathbf{M}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} - \mathbf{\Phi}^{\mathrm{T}}_{\mathbf{q}\boldsymbol{\rho}}\boldsymbol{\lambda} \tag{3.13a}$$

$$\mathbf{\Phi}_{\mathbf{q}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}} - \mathbf{c}_{\dot{\mathbf{q}}}\dot{\mathbf{q}}_{\boldsymbol{\rho}} + \left(\mathbf{\Phi}_{\mathbf{qq}}\ddot{\mathbf{q}} - \mathbf{c}_{\mathbf{q}}\right)\mathbf{q}_{\boldsymbol{\rho}} = \mathbf{c}_{\boldsymbol{\rho}} - \mathbf{\Phi}_{\mathbf{q}\boldsymbol{\rho}}\ddot{\mathbf{q}} \tag{3.13b}$$

where

$$\mathbf{c_q} = -\left(\dot{\mathbf{\Phi}}_\mathbf{q}\right)_\mathbf{q}\dot{\mathbf{q}} - \left(\dot{\mathbf{\Phi}}_t\right)_\mathbf{q} \tag{3.13c}$$

$$\mathbf{c}_{\dot{\mathbf{q}}} = -\left(\dot{\mathbf{\Phi}}_\mathbf{q}\right)_{\dot{\mathbf{q}}}\dot{\mathbf{q}} - \dot{\mathbf{\Phi}}_\mathbf{q} - \left(\dot{\mathbf{\Phi}}_t\right)_{\dot{\mathbf{q}}} = -\mathbf{\Phi_{qq}}\dot{\mathbf{q}} - \dot{\mathbf{\Phi}}_\mathbf{q} - \mathbf{\Phi}_{t\mathbf{q}} \tag{3.13d}$$

$$\mathbf{c}_{\boldsymbol{\rho}} = -\left(\dot{\mathbf{\Phi}}_\mathbf{q}\right)_{\boldsymbol{\rho}}\dot{\mathbf{q}} - \left(\dot{\mathbf{\Phi}}_t\right)_{\boldsymbol{\rho}} \tag{3.13e}$$

The identities $\left(\dot{\mathbf{\Phi}}_\mathbf{q}\right)_{\dot{\mathbf{q}}} = \mathbf{\Phi_{qq}}$ and $\left(\dot{\mathbf{\Phi}}_t\right)_{\dot{\mathbf{q}}} = \mathbf{\Phi}_{t\mathbf{q}}$ were used, and the tensor-vector product rules, with the operator $\otimes$, were applied to the terms involving $\mathbf{\Phi_{qq}}$, $\mathbf{\Phi_{q\rho}}$, $\left(\dot{\mathbf{\Phi}}_\mathbf{q}\right)_\mathbf{q}$ and $\left(\dot{\mathbf{\Phi}}_\mathbf{q}\right)_{\boldsymbol{\rho}}$.

The approach presented for the index-3 TLM can be employed here to obtain the initial conditions of the index-1 TLM.

The tangent linear model DAE in equations (3.13) can be directly integrated in the same way that the EOM in section 3.1.1. Replacing (3.9) in (3.13) and solving for $\mathbf{q}_{\boldsymbol{\rho}}^{n+1}$ and $\boldsymbol{\lambda}^{n+1}$

$$\begin{bmatrix} \frac{4}{h^2}\mathbf{M} + \frac{2}{h}\mathbf{C} + \mathbf{M_q}\ddot{\mathbf{q}} + \mathbf{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\boldsymbol{\lambda} + \mathbf{K} & \mathbf{\Phi}_\mathbf{q}^{\mathrm{T}} \\ \frac{4}{h^2}\mathbf{\Phi_q} - \frac{2}{h}\mathbf{c}_{\dot{\mathbf{q}}} + \mathbf{\Phi_{qq}}\ddot{\mathbf{q}} - \mathbf{c_q} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{q}_{\boldsymbol{\rho}}^{n+1} \\ \boldsymbol{\lambda}_{\boldsymbol{\rho}}^{n+1} \end{bmatrix} = \begin{bmatrix} \mathbf{Q_\rho} - \mathbf{M_\rho}\ddot{\mathbf{q}} - \mathbf{\Phi}_{\mathbf{q}\rho}^{\mathrm{T}}\boldsymbol{\lambda} - \mathbf{M}\hat{\ddot{\mathbf{q}}}_{\boldsymbol{\rho}}^{n} - \mathbf{C}\hat{\dot{\mathbf{q}}}_{\boldsymbol{\rho}}^{n} \\ \mathbf{c_\rho} - \mathbf{\Phi}_{\mathbf{q}\rho}\ddot{\mathbf{q}} - \mathbf{\Phi_q}\hat{\ddot{\mathbf{q}}}_{\boldsymbol{\rho}}^{n} - \mathbf{c}_{\dot{\mathbf{q}}}\hat{\dot{\mathbf{q}}}_{\boldsymbol{\rho}}^{n} \end{bmatrix} \tag{3.14}$$

Scaling applies here

$$\begin{bmatrix} \mathbf{M} + \frac{h}{2}\mathbf{C} + \frac{h^2}{4}\left(\mathbf{M_q}\ddot{\mathbf{q}} + \mathbf{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\boldsymbol{\lambda} + \mathbf{K}\right) & \mathbf{\Phi}_\mathbf{q}^{\mathrm{T}} \\ \frac{4}{h^2}\mathbf{\Phi_q} - \frac{2}{h}\mathbf{c}_{\dot{\mathbf{q}}} + \mathbf{\Phi_{qq}}\ddot{\mathbf{q}} - \mathbf{c_q} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{q}_{\boldsymbol{\rho}}^{n+1} \\ \bar{\boldsymbol{\lambda}}_{\boldsymbol{\rho}}^{n+1} \end{bmatrix} = \begin{bmatrix} \frac{h^2}{4}\left(\mathbf{Q_\rho} - \mathbf{M_\rho}\ddot{\mathbf{q}} - \mathbf{\Phi}_{\mathbf{q}\rho}^{\mathrm{T}}\boldsymbol{\lambda} - \mathbf{M}\hat{\ddot{\mathbf{q}}}_{\boldsymbol{\rho}}^{n} - \mathbf{C}\hat{\dot{\mathbf{q}}}_{\boldsymbol{\rho}}^{n}\right) \\ \mathbf{c_\rho} - \mathbf{\Phi}_{\mathbf{q}\rho}\ddot{\mathbf{q}} - \mathbf{\Phi_q}\hat{\ddot{\mathbf{q}}}_{\boldsymbol{\rho}}^{n} - \mathbf{c}_{\dot{\mathbf{q}}}\hat{\dot{\mathbf{q}}}_{\boldsymbol{\rho}}^{n} \end{bmatrix} \tag{3.15}$$

3.1.3 The penalty formulation

The direct differentiation method for the sensitivity analysis using the penalty formulation was initially developed in [46, 115]. We seek to obtain the sensitivities of the objective function

$$\psi = w\left(\mathbf{q}_F, \dot{\mathbf{q}}_F, \ddot{\mathbf{q}}_F, \boldsymbol{\rho}_F, \boldsymbol{\lambda}_F^*\right) + \int_{t_0}^{t_F} g\left(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}}, \boldsymbol{\rho}, \boldsymbol{\lambda}^*\right) \mathrm{d}t \tag{3.16}$$

The Lagrange multipliers $\boldsymbol{\lambda}$ in (3.1) are replaced by the approximate Lagrange multipliers $\boldsymbol{\lambda}^*$ in (3.16). These variables don't show up in the penalty equations of motion (2.16) and they have to be approximated by (2.20). The initial and final times are considered here independent of the parameters, but more general objective functions are possible.

The gradient of the objective function (3.16) can be obtained by the following expressions

$$\nabla_{\boldsymbol{\rho}}\psi^{\mathrm{T}} = \frac{\mathrm{d}\psi}{\mathrm{d}\boldsymbol{\rho}} = \left[w_{\mathbf{q}}\mathbf{q}_{\boldsymbol{\rho}} + w_{\dot{\mathbf{q}}}\dot{\mathbf{q}}_{\boldsymbol{\rho}} + w_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}} + w_{\boldsymbol{\rho}} + w_{\boldsymbol{\lambda}^*}\alpha \left(\frac{\mathrm{d}\ddot{\boldsymbol{\Phi}}}{\mathrm{d}\boldsymbol{\rho}} + 2\xi\omega\frac{\mathrm{d}\dot{\boldsymbol{\Phi}}}{\mathrm{d}\boldsymbol{\rho}} + \omega^2\frac{\mathrm{d}\boldsymbol{\Phi}}{\mathrm{d}\boldsymbol{\rho}} \right) \right]_F +$$
$$\int_{t_0}^{t_F} \left[g_{\mathbf{q}}\mathbf{q}_{\boldsymbol{\rho}} + g_{\dot{\mathbf{q}}}\dot{\mathbf{q}}_{\boldsymbol{\rho}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}} + g_{\boldsymbol{\rho}} + g_{\boldsymbol{\lambda}^*}\alpha \left(\frac{\mathrm{d}\ddot{\boldsymbol{\Phi}}}{\mathrm{d}\boldsymbol{\rho}} + 2\xi\omega\frac{\mathrm{d}\dot{\boldsymbol{\Phi}}}{\mathrm{d}\boldsymbol{\rho}} + \omega^2\frac{\mathrm{d}\boldsymbol{\Phi}}{\mathrm{d}\boldsymbol{\rho}} \right) \right] \mathrm{dt} \tag{3.17a}$$

$$\frac{\mathrm{d}\ddot{\boldsymbol{\Phi}}}{\mathrm{d}\boldsymbol{\rho}} = \boldsymbol{\Phi}_{\mathbf{q}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}} + \left(\boldsymbol{\Phi}_{\mathbf{qq}}\dot{\mathbf{q}} + \dot{\boldsymbol{\Phi}}_{\mathbf{q}} + \boldsymbol{\Phi}_{t\mathbf{q}} \right) \dot{\mathbf{q}}_{\boldsymbol{\rho}} + \left(\boldsymbol{\Phi}_{\mathbf{qq}}\ddot{\mathbf{q}} + \left(\dot{\boldsymbol{\Phi}}_{\mathbf{q}} \right)_{\mathbf{q}} \dot{\mathbf{q}} + \left(\dot{\boldsymbol{\Phi}}_{t} \right)_{\mathbf{q}} \right) \mathbf{q}_{\boldsymbol{\rho}} +$$
$$\boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}\ddot{\mathbf{q}} + \left(\dot{\boldsymbol{\Phi}}_{\mathbf{q}} \right)_{\boldsymbol{\rho}} \dot{\mathbf{q}} + \left(\dot{\boldsymbol{\Phi}}_{t} \right)_{\boldsymbol{\rho}} \tag{3.17b}$$

$$\frac{\mathrm{d}\dot{\boldsymbol{\Phi}}}{\mathrm{d}\boldsymbol{\rho}} = \boldsymbol{\Phi}_{\mathbf{q}}\dot{\mathbf{q}}_{\boldsymbol{\rho}} + \left(\boldsymbol{\Phi}_{\mathbf{qq}}\dot{\mathbf{q}} + \boldsymbol{\Phi}_{t\mathbf{q}} \right) \mathbf{q}_{\boldsymbol{\rho}} + \boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}\dot{\mathbf{q}} + \boldsymbol{\Phi}_{t\boldsymbol{\rho}} \tag{3.17c}$$

$$\frac{\mathrm{d}\boldsymbol{\Phi}}{\mathrm{d}\boldsymbol{\rho}} = \boldsymbol{\Phi}_{\mathbf{q}}\mathbf{q}_{\boldsymbol{\rho}} + \boldsymbol{\Phi}_{\boldsymbol{\rho}} \tag{3.17d}$$

where the identities $\left(\dot{\boldsymbol{\Phi}}_{\mathbf{q}} \right)_{\dot{\mathbf{q}}} = \boldsymbol{\Phi}_{\mathbf{qq}}$ and $\left(\dot{\boldsymbol{\Phi}}_{t} \right)_{\dot{\mathbf{q}}} = \boldsymbol{\Phi}_{t\mathbf{q}}$ were used, and the tensor-vector product rules mentioned in the nomenclature, with the operator $\otimes$, were applied to the terms involving $\boldsymbol{\Phi}_{\mathbf{qq}}$, $\boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}$, $\left(\dot{\boldsymbol{\Phi}}_{\mathbf{q}} \right)_{\mathbf{q}}$ and $\left(\dot{\boldsymbol{\Phi}}_{\mathbf{q}} \right)_{\boldsymbol{\rho}}$.

In equations (3.17) the derivatives of functions w and g are known, since the objective function has a known expression. The derivatives $\mathbf{q}_{\boldsymbol{\rho}}$, $\dot{\mathbf{q}}_{\boldsymbol{\rho}}$ and $\ddot{\mathbf{q}}_{\boldsymbol{\rho}}$ are the sensitivities of the solution of the dynamical system, and are obtained by differentiating equations (2.16) with respect to each one of the parameters :

$$\frac{\mathrm{d}\bar{\mathbf{M}}}{\mathrm{d}\rho_k}\ddot{\mathbf{q}} + \bar{\mathbf{M}}\frac{\partial\ddot{\mathbf{q}}}{\partial\rho_k} = \frac{\mathrm{d}\bar{\mathbf{Q}}}{\mathrm{d}\rho_k}\,, \quad k = 1, \ldots, p \tag{3.18}$$

Expanding the total derivatives and grouping them together in matrix notation leads to the

following TLM ODE system:

$$\bar{\mathbf{M}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}} + \bar{\mathbf{C}}\dot{\mathbf{q}}_{\boldsymbol{\rho}} + \left(\bar{\mathbf{K}} + \bar{\mathbf{M}}_{\mathbf{q}}\ddot{\mathbf{q}}\right)\mathbf{q}_{\boldsymbol{\rho}} = \bar{\mathbf{Q}}_{\boldsymbol{\rho}} - \bar{\mathbf{M}}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} \tag{3.19a}$$

$$\mathbf{q}_{\boldsymbol{\rho}}(t_0) = \mathbf{q}_{\boldsymbol{\rho}0} \tag{3.19b}$$

$$\dot{\mathbf{q}}_{\boldsymbol{\rho}}(t_0) = \dot{\mathbf{q}}_{\boldsymbol{\rho}0} \tag{3.19c}$$

Observe that each TLM (3.19) is a system of only n equations, compared to the $n+m$ equations of the index-3 and index-1 tangent linear DAE in sections 3.1.1 and 3.1.2 respectively. Equations (3.19b) and (3.19c) are the initial conditions for the tangent linear ODE and can be obtained by an approach similar to the one in Section 3.1.1. The following terms appear in (3.19a):

$$\bar{\mathbf{K}} = -\frac{\partial\bar{\mathbf{Q}}}{\partial\mathbf{q}} = \mathbf{K} + \boldsymbol{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\alpha\left(\dot{\boldsymbol{\Phi}}_{\mathbf{q}}\dot{\mathbf{q}} + \dot{\boldsymbol{\Phi}}_t + 2\xi\omega\dot{\boldsymbol{\Phi}} + \omega^2\boldsymbol{\Phi}\right) +$$
$$\boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\alpha\left(\left(\dot{\boldsymbol{\Phi}}_{\mathbf{q}}\dot{\mathbf{q}}\right)_{\mathbf{q}} + \left(\dot{\boldsymbol{\Phi}}_t\right)_{\mathbf{q}} + 2\xi\omega\left(\boldsymbol{\Phi}_{\mathbf{qq}}\dot{\mathbf{q}} + \boldsymbol{\Phi}_{t\mathbf{q}}\right) + \omega^2\boldsymbol{\Phi}_{\mathbf{q}}\right) \tag{3.20a}$$

$$\bar{\mathbf{C}} = -\frac{\partial\bar{\mathbf{Q}}}{\partial\dot{\mathbf{q}}} = \mathbf{C} + \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\alpha\left(\boldsymbol{\Phi}_{\mathbf{qq}}\dot{\mathbf{q}} + \dot{\boldsymbol{\Phi}}_{\mathbf{q}} + \boldsymbol{\Phi}_{tq} + 2\xi\omega\boldsymbol{\Phi}_{\mathbf{q}}\right) \tag{3.20b}$$

$$\bar{\mathbf{Q}}_{\boldsymbol{\rho}} = \frac{\partial\bar{\mathbf{Q}}}{\partial\boldsymbol{\rho}} = \mathbf{Q}_{\boldsymbol{\rho}} - \boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}^{\mathrm{T}}\alpha\left(\dot{\boldsymbol{\Phi}}_{\mathbf{q}}\dot{\mathbf{q}} + \dot{\boldsymbol{\Phi}}_t + 2\xi\omega\dot{\boldsymbol{\Phi}} + \omega^2\boldsymbol{\Phi}\right) -$$
$$\boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\alpha\left(\left(\dot{\boldsymbol{\Phi}}_{\mathbf{q}}\dot{\mathbf{q}}\right)_{\boldsymbol{\rho}} + \dot{\boldsymbol{\Phi}}_{t\boldsymbol{\rho}} + 2\xi\omega\dot{\boldsymbol{\Phi}}_{\boldsymbol{\rho}} + \omega^2\boldsymbol{\Phi}_{\boldsymbol{\rho}}\right) \tag{3.20c}$$

$$\bar{\mathbf{M}}_{\mathbf{q}}\ddot{\mathbf{q}} = \mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\left(\alpha\boldsymbol{\Phi}_{\mathbf{q}}\ddot{\mathbf{q}}\right) + \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\alpha\left(\boldsymbol{\Phi}_{\mathbf{qq}}\ddot{\mathbf{q}}\right) \tag{3.20d}$$

$$\bar{\mathbf{M}}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} = \mathbf{M}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}^{\mathrm{T}}\left(\alpha\boldsymbol{\Phi}_{\mathbf{q}}\ddot{\mathbf{q}}\right) + \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\alpha\left(\boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}\ddot{\mathbf{q}}\right) \tag{3.20e}$$

We have $\mathbf{K} = -\mathbf{Q}_{\mathbf{q}}$ in equation (3.20a) and $\mathbf{C} = -\mathbf{Q}_{\dot{\mathbf{q}}}$ in (3.20b). In equations (3.20d) and (3.20e) the following tensor-vector products are calculated: $\mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} = \mathbf{M}_{\mathbf{q}} \otimes \ddot{\mathbf{q}}$ and $\mathbf{M}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} = \mathbf{M}_{\boldsymbol{\rho}} \otimes \ddot{\mathbf{q}}$.

3.1.4 Maggi's formulation

The direct differentiation method of this section is useful when the number of parameters p is small. If the number of parameters is large, this approach can become computationally expensive.

One seeks to obtain the sensitivity of a cost function defined in terms of the states and the

design parameters of the system

$$\psi = \int_{t_0}^{t_F} g\left(\mathbf{q}, \dot{\mathbf{q}}, \boldsymbol{\rho}\right) \mathrm{dt} \tag{3.21}$$

The proposed expressions can be extended to include more general cost functions. Note that the cost function (3.21) is supposed to depend explicitly on the full set of states $\mathbf{q}$ and their derivatives, instead of depending on the independent ones $\mathbf{z}$ and their derivatives.

The gradient of the cost function (3.21) is:

$$\nabla_{\boldsymbol{\rho}} \psi^{\mathrm{T}} = \frac{\mathrm{d}\psi}{\mathrm{d}\boldsymbol{\rho}} = \int_{t_0}^{t_F} \left(\left(\frac{\partial g}{\partial \mathbf{q}} \frac{\partial \mathbf{q}}{\partial \mathbf{z}} + \frac{\partial g}{\partial \dot{\mathbf{q}}} \frac{\partial \dot{\mathbf{q}}}{\partial \mathbf{z}} \right) \frac{\partial \mathbf{z}}{\partial \boldsymbol{\rho}} + \frac{\partial g}{\partial \dot{\mathbf{q}}} \frac{\partial \dot{\mathbf{q}}}{\partial \dot{\mathbf{z}}} \frac{\partial \dot{\mathbf{z}}}{\partial \boldsymbol{\rho}} + \frac{\partial g}{\partial \boldsymbol{\rho}} \right) \mathrm{dt} \tag{3.22}$$

The following equation is derived from (2.26a) and (2.27)

$$\begin{aligned}
&\begin{bmatrix} \boldsymbol{\Phi}_{\mathbf{q}} \\ \mathbf{B} \end{bmatrix} \frac{\partial \dot{\mathbf{q}}}{\partial \mathbf{q}} + \begin{bmatrix} \boldsymbol{\Phi}_{\mathbf{qq}} \\ \mathbf{0} \end{bmatrix} \dot{\mathbf{q}} = \begin{bmatrix} \mathbf{b}_{\mathbf{q}} \\ \mathbf{0} \end{bmatrix} \Rightarrow \\
&\frac{\partial \dot{\mathbf{q}}}{\partial \mathbf{q}} = \begin{bmatrix} \boldsymbol{\Phi}_{\mathbf{q}} \\ \mathbf{B} \end{bmatrix}^{-1} \begin{bmatrix} -\boldsymbol{\Phi}_{t\mathbf{q}} - \boldsymbol{\Phi}_{\mathbf{qq}} \dot{\mathbf{q}} \\ \mathbf{0} \end{bmatrix} \Rightarrow \\
&\frac{\partial \dot{\mathbf{q}}}{\partial \mathbf{q}} = -\mathbf{S}\left(\boldsymbol{\Phi}_{\mathbf{qq}} \dot{\mathbf{q}} + \boldsymbol{\Phi}_{t\mathbf{q}}\right)
\end{aligned} \tag{3.23}$$

From Eqns. (2.23a), (2.23c), and (3.23) one obtains:

$$\frac{\partial \mathbf{q}}{\partial \mathbf{z}} = \frac{\partial \dot{\mathbf{q}}}{\partial \dot{\mathbf{z}}} = \mathbf{R} \tag{3.24a}$$

$$\frac{\partial \dot{\mathbf{q}}}{\partial \mathbf{z}} = \frac{\partial \dot{\mathbf{q}}}{\partial \mathbf{q}} \frac{\partial \mathbf{q}}{\partial \mathbf{z}} = -\mathbf{S}\left(\boldsymbol{\Phi}_{\mathbf{qq}} \dot{\mathbf{q}} + \boldsymbol{\Phi}_{t\mathbf{q}}\right) \mathbf{R} = -\mathbf{S}\dot{\boldsymbol{\Phi}}_{\mathbf{q}} \mathbf{R} \tag{3.24b}$$

Using (3.24) the gradient (3.22) becomes:

$$\nabla_{\boldsymbol{\rho}} \psi^{\mathrm{T}} = \int_{t_0}^{t_F} \left(\left(g_{\mathbf{q}} - g_{\dot{\mathbf{q}}} \mathbf{S} \dot{\boldsymbol{\Phi}}_{\mathbf{q}} \right) \mathbf{R} \mathbf{z}_{\boldsymbol{\rho}} + g_{\dot{\mathbf{q}}} \mathbf{R} \dot{\mathbf{z}}_{\boldsymbol{\rho}} + g_{\boldsymbol{\rho}} \right) \mathrm{dt} \tag{3.25}$$

In Eq. (3.25) the derivatives of the function g are known, since the objective function has a known expression, and the derivatives $\mathbf{z}_{\boldsymbol{\rho}}$ and $\dot{\mathbf{z}}_{\boldsymbol{\rho}}$ are the sensitivities of the solution of the dynamical equations (2.25). They are obtained by differentiating (2.25) with respect to each one

of the parameters, as follows:

$$\frac{\partial \bar{\mathbf{M}}}{\partial \rho_i}\ddot{\mathbf{z}} + \bar{\mathbf{M}}_{\mathbf{z}}\ddot{\mathbf{z}}\frac{\partial \mathbf{z}}{\partial \rho_i} + \bar{\mathbf{M}}\frac{\partial \ddot{\mathbf{z}}}{\partial \rho_i} = \frac{\partial \bar{\mathbf{Q}}}{\partial \mathbf{z}}\frac{\partial \mathbf{z}}{\partial \rho_i} + \frac{\partial \bar{\mathbf{Q}}}{\partial \dot{\mathbf{z}}}\frac{\partial \dot{\mathbf{z}}}{\partial \rho_i} + \frac{\partial \bar{\mathbf{Q}}}{\partial \rho_i} \tag{3.26}$$

where the term $\bar{\mathbf{M}}_{\mathbf{z}}\ddot{\mathbf{z}} \equiv \bar{\mathbf{M}}_{\mathbf{z}} \otimes \ddot{\mathbf{z}}$ is a tensor-vector product, described in the nomenclature, and is calculated later in this section.

Grouping all the derivatives with respect to parameters together in (3.26) leads to the following TLM ODE system:

$$\bar{\mathbf{M}}\ddot{\mathbf{z}}_{\boldsymbol{\rho}} + \bar{\mathbf{C}}\dot{\mathbf{z}}_{\boldsymbol{\rho}} + \left(\bar{\mathbf{K}} + \bar{\mathbf{M}}_{\mathbf{z}}\ddot{\mathbf{z}}\right)\mathbf{z}_{\boldsymbol{\rho}} = \bar{\mathbf{Q}}_{\boldsymbol{\rho}} - \bar{\mathbf{M}}_{\boldsymbol{\rho}}\ddot{\mathbf{z}} \tag{3.27a}$$

$$\mathbf{z}_{\boldsymbol{\rho}}(t_0) = \mathbf{z}_{\boldsymbol{\rho}0} \tag{3.27b}$$

$$\dot{\mathbf{z}}_{\boldsymbol{\rho}}(t_0) = \dot{\mathbf{z}}_{\boldsymbol{\rho}0} \tag{3.27c}$$

where $\mathbf{z}_{\boldsymbol{\rho}}$ is the sensitivity matrix of size $d \times p$. The matricial TLM (3.27) can be interpreted as a set of p classical ODE systems of size d, one per parameter. Since the initial conditions (3.27b) and (3.27c) are the initial sensitivities of the *DOF* of the system, their values are assigned based on their physical interpretation.

In Eq. (3.27a), $\bar{\mathbf{K}}$, $\bar{\mathbf{C}}$ and $\bar{\mathbf{Q}}_{\boldsymbol{\rho}}$ are derivatives of the projected vector of generalized forces $\bar{\mathbf{Q}}$, given by

$$\bar{\mathbf{K}} = -\frac{\partial \bar{\mathbf{Q}}}{\partial \mathbf{z}} = -\left(\frac{\partial \bar{\mathbf{Q}}}{\partial \mathbf{q}}\frac{\partial \mathbf{q}}{\partial \mathbf{z}} + \frac{\partial \bar{\mathbf{Q}}}{\partial \dot{\mathbf{q}}}\frac{\partial \dot{\mathbf{q}}}{\partial \mathbf{z}}\right) = -\left(\bar{\mathbf{Q}}_{\mathbf{q}} - \bar{\mathbf{Q}}_{\dot{\mathbf{q}}}\mathbf{S}\dot{\mathbf{\Phi}}_{\mathbf{q}}\right)\mathbf{R} \tag{3.28a}$$

$$\bar{\mathbf{C}} = -\frac{\partial \bar{\mathbf{Q}}}{\partial \dot{\mathbf{z}}} = -\frac{\partial \bar{\mathbf{Q}}}{\partial \dot{\mathbf{q}}}\frac{\partial \dot{\mathbf{q}}}{\partial \dot{\mathbf{z}}} = -\bar{\mathbf{Q}}_{\dot{\mathbf{q}}}\mathbf{R} \tag{3.28b}$$

$$\bar{\mathbf{Q}}_{\boldsymbol{\rho}} = -\frac{\partial \bar{\mathbf{Q}}}{\partial \boldsymbol{\rho}} = \mathbf{R}^{\mathrm{T}}\left(\mathbf{Q}_{\boldsymbol{\rho}} - \mathbf{M}_{\boldsymbol{\rho}}\mathbf{S}\mathbf{c}\right) \tag{3.28c}$$

where equation (3.24b) was used inside (3.28a).

The terms $\bar{\mathbf{M}}_{\mathbf{z}}\ddot{\mathbf{z}}$ and $\bar{\mathbf{M}}_{\boldsymbol{\rho}}\ddot{\mathbf{z}}$ in (3.27a) are derivatives of the projected mass matrix $\bar{\mathbf{M}}$ times the

vector of accelerations, and are obtained by

$$\bar{\mathbf{M}}_{\mathbf{z}}\ddot{\mathbf{z}} = \mathbf{R}_{\mathbf{z}}^{\mathrm{T}}\mathbf{M}\mathbf{R}\ddot{\mathbf{z}} + \mathbf{R}^{\mathrm{T}}\mathbf{M}_{\mathbf{z}}\mathbf{R}\ddot{\mathbf{z}} + \mathbf{R}^{\mathrm{T}}\mathbf{M}\mathbf{R}_{\mathbf{z}}\ddot{\mathbf{z}} \tag{3.29a}$$

$$\bar{\mathbf{M}}_{\boldsymbol{\rho}}\ddot{\mathbf{z}} = \mathbf{R}^{\mathrm{T}}\mathbf{M}_{\boldsymbol{\rho}}\mathbf{R}\ddot{\mathbf{z}} \tag{3.29b}$$

In Eq. (3.28c) and Eq. (3.29b) the following products must be calculated with the tensorial operator $\otimes$: $\mathbf{M}_{\boldsymbol{\rho}}\mathbf{S}\mathbf{c} = \mathbf{M}_{\boldsymbol{\rho}} \otimes \mathbf{S}\mathbf{c}$ and $\mathbf{M}_{\boldsymbol{\rho}}\mathbf{R}\ddot{\mathbf{z}} = \mathbf{M}_{\boldsymbol{\rho}} \otimes \mathbf{R}\ddot{\mathbf{z}}$. The derivatives $\bar{\mathbf{Q}}_{\mathbf{q}}$ and $\bar{\mathbf{Q}}_{\dot{\mathbf{q}}}$ of Eq. (3.28a) and Eq. (3.28b) are calculated as follows:

$$\bar{\mathbf{Q}}_{\mathbf{q}} = \mathbf{R}_{\mathbf{q}}^{\mathrm{T}}\left(\mathbf{Q} - \mathbf{M}\mathbf{S}\mathbf{c}\right) - \mathbf{R}^{\mathrm{T}}\left(\mathbf{K} + \mathbf{M}_{\mathbf{q}}\mathbf{S}\mathbf{c} + \mathbf{M}\frac{\partial \mathbf{S}\mathbf{c}}{\partial \mathbf{q}}\right) \tag{3.30a}$$

$$\bar{\mathbf{Q}}_{\dot{\mathbf{q}}} = -\mathbf{R}^{\mathrm{T}}\left(\mathbf{C} + \mathbf{M}\mathbf{S}\mathbf{c}_{\dot{\mathbf{q}}}\right) \tag{3.30b}$$

where $\mathbf{K} = -\mathbf{Q}_{\mathbf{q}}$, $\mathbf{C} = -\mathbf{Q}_{\dot{\mathbf{q}}}$ and the following products involve tensors: $\mathbf{R}_{\mathbf{q}}^{\mathrm{T}}\left(\mathbf{Q} - \mathbf{M}\mathbf{S}\mathbf{c}\right) = \mathbf{R}_{\mathbf{q}}^{\mathrm{T}} \otimes \left(\mathbf{Q} - \mathbf{M}\mathbf{S}\mathbf{c}\right)$ and $\mathbf{M}_{\mathbf{q}}\mathbf{S}\mathbf{c} = \mathbf{M}_{\mathbf{q}} \otimes \mathbf{S}\mathbf{c}$. The following extra terms, derived from (2.27), are needed for Eq. (3.30a) and Eq. (3.30b):

$$\underbrace{\frac{\partial \mathbf{R}}{\partial q_j}}_{n\times d} = \begin{bmatrix} \boldsymbol{\Phi}_{\mathbf{q}} \\ \mathbf{B} \end{bmatrix}^{-1} \begin{bmatrix} -\dfrac{\partial \boldsymbol{\Phi}_{\mathbf{q}}}{\partial q_j}\mathbf{R} \\ \mathbf{0} \end{bmatrix} = -\mathbf{S}\frac{\partial \boldsymbol{\Phi}_{\mathbf{q}}}{\partial q_j}\mathbf{R} \tag{3.31}$$

$$\underbrace{\mathbf{R}_{\mathbf{q}}^{\mathrm{T}}}_{d\times n\times n} = -\mathbf{R}^{\mathrm{T}} \otimes \begin{bmatrix} \boldsymbol{\Phi}_{\mathbf{qq}}^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \otimes \begin{bmatrix} \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}} & \mathbf{B}^{\mathrm{T}} \end{bmatrix}^{-1} = -\mathbf{R}^{\mathrm{T}} \otimes \boldsymbol{\Phi}_{\mathbf{qq}}^{\mathrm{T}} \otimes \mathbf{S}^{\mathrm{T}} = -\mathbf{R}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}} \tag{3.32}$$

$$\frac{\partial \mathbf{S}\mathbf{c}}{\partial \mathbf{q}} = \mathbf{S}\left(-\boldsymbol{\Phi}_{\mathbf{qq}}\mathbf{S}\mathbf{c} + \mathbf{c}_{\mathbf{q}}\right) \tag{3.33}$$

$$\mathbf{c}_{\mathbf{q}} = -\left(\dot{\boldsymbol{\Phi}}_{\mathbf{q}}\right)_{\mathbf{q}}\dot{\mathbf{q}} - \left(\dot{\boldsymbol{\Phi}}_{t}\right)_{\mathbf{q}} \tag{3.34}$$

$$\mathbf{c}_{\dot{\mathbf{q}}} = -\left(\dot{\boldsymbol{\Phi}}_{\mathbf{q}}\right)_{\dot{\mathbf{q}}}\dot{\mathbf{q}} - \dot{\boldsymbol{\Phi}}_{\mathbf{q}} - \left(\dot{\boldsymbol{\Phi}}_{t}\right)_{\dot{\mathbf{q}}} = -\boldsymbol{\Phi}_{\mathbf{qq}}\dot{\mathbf{q}} - \dot{\boldsymbol{\Phi}}_{\mathbf{q}} - \boldsymbol{\Phi}_{t\mathbf{q}} \tag{3.35}$$

In Eqns. (3.33), (3.34), and (3.35), the terms $\boldsymbol{\Phi}_{\mathbf{qq}}\mathbf{S}\mathbf{c} \equiv \boldsymbol{\Phi}_{\mathbf{qq}} \otimes \mathbf{S}\mathbf{c}$, $\left(\dot{\boldsymbol{\Phi}}_{\mathbf{q}}\right)_{\mathbf{q}}\dot{\mathbf{q}} \equiv \left(\dot{\boldsymbol{\Phi}}_{\mathbf{q}}\right)_{\mathbf{q}} \otimes \dot{\mathbf{q}}$, and $\boldsymbol{\Phi}_{\mathbf{qq}}\dot{\mathbf{q}} \equiv \boldsymbol{\Phi}_{\mathbf{qq}} \otimes \dot{\mathbf{q}}$ are tensor products and $\left(\dot{\boldsymbol{\Phi}}_{\mathbf{q}}\right)_{\mathbf{q}}$, $\boldsymbol{\Phi}_{\mathbf{qq}}$ are the tensor derivatives of $\dot{\boldsymbol{\Phi}}_{\mathbf{q}}$ and $\boldsymbol{\Phi}_{\mathbf{q}}$ with respect to $\mathbf{q}$ respectively, obtained as explained in the nomenclature. Moreover, the following

identities from (2.27) were used to obtain the derivatives (3.31), (3.32) and (3.33),

$$\begin{bmatrix} \mathbf{\Phi_q} \\ \mathbf{B} \end{bmatrix} \mathbf{R} = \begin{bmatrix} \mathbf{0}_{m\times d} \\ \mathbf{I}_d \end{bmatrix} \tag{3.36a}$$

$$\mathbf{R}^{\mathrm{T}} \begin{bmatrix} \mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}} & \mathbf{B}^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{d\times m} & \mathbf{I}_d \end{bmatrix} \tag{3.36b}$$

$$\begin{bmatrix} \mathbf{\Phi_q} \\ \mathbf{B} \end{bmatrix} \mathbf{Sc} = \begin{bmatrix} \mathbf{c} \\ \mathbf{0}_{d\times 1} \end{bmatrix} \tag{3.36c}$$

Equation (3.29a) involves several tensor product terms: $\mathbf{M_z R\ddot{z}} \equiv \mathbf{M_z} \otimes \mathbf{R\ddot{z}}$, $\mathbf{R_z^{\mathrm{T}} M R\ddot{z}} \equiv \mathbf{R_z^{\mathrm{T}}} \otimes \mathbf{MR\ddot{z}}$, and $\mathbf{R_z \ddot{z}} \equiv \mathbf{R_z} \otimes \mathbf{\ddot{z}}$. Using the identity $\partial q_j / \partial z_i = R_{ji}$ and Eq. (3.31), the components of $\mathbf{R_z}$ and $\mathbf{M_z}$ are

$$\frac{\partial \mathbf{R}}{\partial z_i} = \sum_{j=1}^{n} \frac{\partial \mathbf{R}}{\partial q_j} \frac{\partial q_j}{\partial z_i} = -\mathbf{S} \sum_{j=1}^{n} \left(\frac{\partial \mathbf{\Phi_q}}{\partial q_j} R_{ji} \right) \mathbf{R} \tag{3.37}$$

$$\frac{\partial \mathbf{M}}{\partial z_i} = \sum_{j=1}^{n} \frac{\partial \mathbf{M}}{\partial q_j} \frac{\partial q_j}{\partial z_i} = \sum_{j=1}^{n} \frac{\partial \mathbf{M}}{\partial q_j} R_{ji} \tag{3.38}$$

3.2 Adjoint variable method

In this section, the sensitivity equations for the EOM presented in chapter 2 are derived based on the adjoint variable method.

3.2.1 Index-3 DAE formulation

The adjoint variable method for the sensitivity analysis using the index-3 formulation in chapter 2 was developed in [41] for objective functions dependent on $\mathbf{q}, \dot{\mathbf{q}}, \boldsymbol{\rho}, \boldsymbol{\lambda}$ and variable time limits t_F. In this section, the results are revisited also for objective functions depending on $\ddot{\mathbf{q}}$ but fixed end time t_F.

Considering the EOM (2.3) the gradient (3.2) can be indirectly obtained via the Lagrangian

function

$$\mathcal{L}(\boldsymbol{\rho}) = w(\mathbf{q}_F, \dot{\mathbf{q}}_F, \ddot{\mathbf{q}}_F, \boldsymbol{\rho}_F, \boldsymbol{\lambda}_F) + \int_{t_0}^{t_F} g(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}}, \boldsymbol{\lambda}, \boldsymbol{\rho})\,\mathrm{dt} - \int_{t_0}^{t_F} \boldsymbol{\mu}^{\mathrm{T}}\left(\mathbf{M}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\boldsymbol{\lambda} - \mathbf{Q}\right)\mathrm{dt} - \int_{t_0}^{t_F} \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\boldsymbol{\Phi}\mathrm{dt} \tag{3.39}$$

and the following identity that holds along any solution of the EOM:

$$\nabla_{\boldsymbol{\rho}}\psi = \nabla_{\boldsymbol{\rho}}\mathcal{L} \tag{3.40}$$

Infinitesimal variations of $\mathcal{L}$ under infinitesimal variations $\delta\boldsymbol{\rho}$ are as follows (note that the computation of $\delta\boldsymbol{\mu}$ and $\delta\boldsymbol{\lambda}$ is not needed):

$$\begin{aligned}
\delta\mathcal{L} &= \left[w_{\mathbf{q}}\delta\mathbf{q} + w_{\dot{\mathbf{q}}}\delta\dot{\mathbf{q}} + w_{\ddot{\mathbf{q}}}\delta\ddot{\mathbf{q}} + w_{\boldsymbol{\lambda}}\delta\boldsymbol{\lambda} + w_{\boldsymbol{\rho}}\delta\boldsymbol{\rho}\right]_F \\
&+ \int_{t_0}^{t_F}\left(g_{\mathbf{q}} - \boldsymbol{\mu}^{\mathrm{T}}\left(\mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\boldsymbol{\lambda} - \mathbf{Q}_{\mathbf{q}}\right) - \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathbf{q}}\right)\delta\mathbf{q}\mathrm{dt} \\
&+ \int_{t_0}^{t_F}\left(g_{\dot{\mathbf{q}}} + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{Q}_{\dot{\mathbf{q}}}\right)\delta\dot{\mathbf{q}}\mathrm{dt} + \int_{t_0}^{t_F}\left(g_{\ddot{\mathbf{q}}} - \boldsymbol{\mu}^{\mathrm{T}}\mathbf{M}\right)\delta\ddot{\mathbf{q}}\mathrm{dt} + \int_{t_0}^{t_F}\left(g_{\boldsymbol{\lambda}} - \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\right)\delta\boldsymbol{\lambda}\mathrm{dt} \\
&+ \int_{t_0}^{t_F}\left(g_{\boldsymbol{\rho}} - \boldsymbol{\mu}^{\mathrm{T}}\left(\mathbf{M}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}^{\mathrm{T}}\boldsymbol{\lambda} - \mathbf{Q}_{\boldsymbol{\rho}}\right) - \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\boldsymbol{\Phi}_{\boldsymbol{\rho}}\right)\delta\boldsymbol{\rho}\mathrm{dt}
\end{aligned} \tag{3.41}$$

Integrating by parts the integrals involving $\delta\dot{\mathbf{q}}, \delta\ddot{\mathbf{q}}$:

$$\begin{aligned}
\delta\mathcal{L} &= \left[w_{\mathbf{q}}\delta\mathbf{q} + w_{\dot{\mathbf{q}}}\delta\dot{\mathbf{q}} + w_{\ddot{\mathbf{q}}}\delta\ddot{\mathbf{q}} + w_{\boldsymbol{\lambda}}\delta\boldsymbol{\lambda} + w_{\boldsymbol{\rho}}\delta\boldsymbol{\rho}\right]_F \\
&+ \int_{t_0}^{t_F}\left(g_{\mathbf{q}} - \boldsymbol{\mu}^{\mathrm{T}}\left(\mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\boldsymbol{\lambda} + \mathbf{K}\right) - \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathbf{q}}\right)\delta\mathbf{q}\mathrm{dt} + \left(g_{\dot{\mathbf{q}}} - \boldsymbol{\mu}^{\mathrm{T}}\mathbf{C}\right)\delta\mathbf{q}\Big|_{t0}^{tF} \\
&- \int_{t_0}^{t_F}\left(\frac{\mathrm{d}g_{\dot{\mathbf{q}}}}{\mathrm{dt}} - \dot{\boldsymbol{\mu}}^{\mathrm{T}}\mathbf{C} - \boldsymbol{\mu}^{\mathrm{T}}\dot{\mathbf{C}}\right)\delta\mathbf{q}\mathrm{dt} + \left(g_{\ddot{\mathbf{q}}} - \boldsymbol{\mu}^{\mathrm{T}}\mathbf{M}\right)\delta\dot{\mathbf{q}}\Big|_{t0}^{tF} - \left(\frac{\mathrm{d}g_{\ddot{\mathbf{q}}}}{\mathrm{dt}} - \dot{\boldsymbol{\mu}}^{\mathrm{T}}\mathbf{M} - \boldsymbol{\mu}^{\mathrm{T}}\dot{\mathbf{M}}\right)\delta\mathbf{q}\Big|_{t0}^{tF} \\
&+ \int_{t_0}^{t_F}\left(\frac{\mathrm{d}^2 g_{\ddot{\mathbf{q}}}}{\mathrm{dt}^2} - \ddot{\boldsymbol{\mu}}^{\mathrm{T}}\mathbf{M} - 2\dot{\boldsymbol{\mu}}^{\mathrm{T}}\dot{\mathbf{M}} - \boldsymbol{\mu}^{\mathrm{T}}\ddot{\mathbf{M}}\right)\delta\mathbf{q}\mathrm{dt} + \int_{t_0}^{t_F}\left(g_{\boldsymbol{\lambda}} - \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\right)\delta\boldsymbol{\lambda}\mathrm{dt} + \\
&\int_{t_0}^{t_F}\left(g_{\boldsymbol{\rho}} - \boldsymbol{\mu}^{\mathrm{T}}\left(\mathbf{M}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}^{\mathrm{T}}\boldsymbol{\lambda} - \mathbf{Q}_{\boldsymbol{\rho}}\right) - \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\boldsymbol{\Phi}_{\boldsymbol{\rho}}\right)\delta\boldsymbol{\rho}\mathrm{dt}
\end{aligned} \tag{3.42}$$

When the following equations are satisfied all the integral terms that involve $\delta\mathbf{q}$ and $\delta\boldsymbol{\lambda}$ cancel:

$$g_{\mathbf{q}} - \boldsymbol{\mu}^{\mathrm{T}}\left(\mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\boldsymbol{\lambda} + \mathbf{K}\right) - \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathbf{q}} - \frac{\mathrm{d}g_{\dot{\mathbf{q}}}}{\mathrm{dt}} + \dot{\boldsymbol{\mu}}^{\mathrm{T}}\mathbf{C} + \boldsymbol{\mu}^{\mathrm{T}}\dot{\mathbf{C}} + \frac{\mathrm{d}^2 g_{\ddot{\mathbf{q}}}}{\mathrm{dt}^2} - \ddot{\boldsymbol{\mu}}^{\mathrm{T}}\mathbf{M} - 2\dot{\boldsymbol{\mu}}^{\mathrm{T}}\dot{\mathbf{M}} - \boldsymbol{\mu}^{\mathrm{T}}\ddot{\mathbf{M}} = \mathbf{0} \tag{3.43a}$$

$$\boldsymbol{\Phi}_{\mathbf{q}}\boldsymbol{\mu} = g_{\boldsymbol{\lambda}}^{\mathrm{T}} \tag{3.43b}$$

The initial conditions can be obtained by canceling the additional terms at the final time appearing in (3.42): $\delta\mathbf{q}_F$, $\delta\dot{\mathbf{q}}_F$, $\delta\boldsymbol{\lambda}_F$

$$\begin{aligned}&\left[w_{\mathbf{q}} + g_{\dot{\mathbf{q}}} - \boldsymbol{\mu}^{\mathrm{T}}\mathbf{C} - \frac{\mathrm{d}g_{\ddot{\mathbf{q}}}}{\mathrm{dt}} + \dot{\boldsymbol{\mu}}^{\mathrm{T}}\mathbf{M} + \boldsymbol{\mu}^{\mathrm{T}}\dot{\mathbf{M}}\right]_F \delta\mathbf{q}_F + \\ &\left[w_{\dot{\mathbf{q}}} + g_{\ddot{\mathbf{q}}} - \boldsymbol{\mu}^{\mathrm{T}}\mathbf{M}\right]_F \delta\dot{\mathbf{q}}_F + \left[w_{\ddot{\mathbf{q}}}\delta\ddot{\mathbf{q}}\right]_F + \left[w_{\boldsymbol{\lambda}}\delta\boldsymbol{\lambda}\right]_F = \mathbf{0}\end{aligned} \tag{3.44}$$

An additional initial condition is given by (3.43b) that holds at any time, in particular at $t = t_F$.

It can be shown that the resulting adjoint system is an index-3 DAE in $\boldsymbol{\mu}$ and $\boldsymbol{\mu}_{\boldsymbol{\Phi}}$:

$$\begin{aligned}&\mathbf{M}^{\mathrm{T}}\ddot{\boldsymbol{\mu}} + \left(2\dot{\mathbf{M}} - \mathbf{C}\right)^{\mathrm{T}}\dot{\boldsymbol{\mu}} + \left(\mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\boldsymbol{\lambda} + \mathbf{K} + \ddot{\mathbf{M}} - \dot{\mathbf{C}}\right)^{\mathrm{T}}\boldsymbol{\mu} + \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\boldsymbol{\mu}_{\boldsymbol{\Phi}} = \\ &g_{\mathbf{q}}^{\mathrm{T}} - \frac{\mathrm{d}g_{\dot{\mathbf{q}}}}{\mathrm{dt}}^{\mathrm{T}} + \frac{\mathrm{d}^2 g_{\ddot{\mathbf{q}}}}{\mathrm{dt}^2}^{\mathrm{T}}\end{aligned} \tag{3.45a}$$

$$\boldsymbol{\Phi}_{\mathbf{q}}\boldsymbol{\mu} = g_{\boldsymbol{\lambda}}^{\mathrm{T}} \tag{3.45b}$$

and has the following initial conditions

$$\left[\mathbf{M}^{\mathrm{T}}\dot{\boldsymbol{\mu}} + \left(\dot{\mathbf{M}} - \mathbf{C}\right)^{\mathrm{T}}\boldsymbol{\mu}\right]_F = \left[\frac{\mathrm{d}g_{\ddot{\mathbf{q}}}}{\mathrm{dt}}^{\mathrm{T}} - g_{\dot{\mathbf{q}}}^{\mathrm{T}} - w_{\mathbf{q}}^{\mathrm{T}}\right]_F \tag{3.45c}$$

$$\left[\mathbf{M}^{\mathrm{T}}\boldsymbol{\mu}\right]_F = \left[g_{\dot{\mathbf{q}}}^{\mathrm{T}} + w_{\dot{\mathbf{q}}}^{\mathrm{T}}\right]_F \tag{3.45d}$$

$$\left[w_{\ddot{\mathbf{q}}}\right]_F = \mathbf{0} \tag{3.45e}$$

$$\left[w_{\boldsymbol{\lambda}}|\right]_F = \mathbf{0} \tag{3.45f}$$

$$\left[\boldsymbol{\Phi}_{\mathbf{q}}\boldsymbol{\mu}\right]_F = \left[g_{\boldsymbol{\lambda}}^{\mathrm{T}}\right]_F \tag{3.45g}$$

Equations (3.45e) and (3.45e) express incompatible conditions for the objective function, mean-

ing that the objective function cannot depend on the acceleration or Lagrange multipliers in the final time. Moreover, if $g_{\ddot{\mathbf{q}}}$, $w_{\dot{\mathbf{q}}}$ or $g_{\boldsymbol{\lambda}}$ are not equal to zero, equations (3.45d) and (3.45g) constitute an incompatible system of equations in $\boldsymbol{\mu}$ that can be solved, according to [41], adding the following terms, evaluated at the final time, to the Lagrangian (3.39).

$$\boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{\Phi}\left(t_F, \mathbf{q}_F, \boldsymbol{\rho}_F\right) \tag{3.46}$$

$$\boldsymbol{\eta}^{\mathrm{T}}\dot{\boldsymbol{\Phi}}\left(t_F, \mathbf{q}_F, \dot{\mathbf{q}}_F, \boldsymbol{\rho}_F\right) \tag{3.47}$$

where $\boldsymbol{\gamma} \in \mathbb{R}^m$ and $\boldsymbol{\mu} \in \mathbb{R}^m$ are new adjoint variables that need to be determined and the constraint equations are given in equations (2.1) and (2.22b).

The addition of condition (3.46) is convenient, since the constraint equations are imposed by the index-3 formulation. Nevertheless, the addition of (3.47) can be problematic since they are hidden constraints not explicitly imposed by the index-3 formulation unless, for example, projection techniques are employed, as proposed in [109] and suggested in section 3.1.1.

The addition of equations (3.46) and (3.47) to the Lagrangian (3.39), contribute with the following terms to the variation (3.41)

$$\boldsymbol{\gamma}^{\mathrm{T}}\left[\boldsymbol{\Phi}_{\mathbf{q}}\delta\mathbf{q} + \boldsymbol{\Phi}_{\boldsymbol{\rho}}\delta\boldsymbol{\rho}\right]_F \tag{3.48}$$

$$\boldsymbol{\eta}^{\mathrm{T}}\left[\left(\boldsymbol{\Phi}_{\mathbf{qq}}\dot{\mathbf{q}} + \boldsymbol{\Phi}_{t\mathbf{q}}\right)\delta\mathbf{q} + \boldsymbol{\Phi}_{\mathbf{q}}\delta\dot{\mathbf{q}} + \left(\boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}\dot{\mathbf{q}} + \boldsymbol{\Phi}_{t\boldsymbol{\rho}}\right)\delta\boldsymbol{\rho}\right]_F \tag{3.49}$$

The final adjoint is not affected by these additions and equations (3.45) are still valid, nevertheless the initial conditions for the adjoint at the final time become the following

$$\left[\mathbf{M}^{\mathrm{T}}\dot{\boldsymbol{\mu}} + \left(\dot{\mathbf{M}} - \mathbf{C}\right)^{\mathrm{T}}\boldsymbol{\mu} + \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\boldsymbol{\gamma} + \left(\boldsymbol{\Phi}_{\mathbf{qq}}\dot{\mathbf{q}} + \boldsymbol{\Phi}_{t\mathbf{q}}\right)^{\mathrm{T}}\boldsymbol{\eta}\right]_F = \left[\frac{\mathrm{d}g_{\ddot{\mathbf{q}}}{}^{\mathrm{T}}}{\mathrm{dt}} - g_{\dot{\mathbf{q}}}^{\mathrm{T}} - w_{\mathbf{q}}^{\mathrm{T}}\right]_F \tag{3.50a}$$

$$\left[\mathbf{M}^{\mathrm{T}}\boldsymbol{\mu} + \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\boldsymbol{\eta}\right]_F = \left[g_{\ddot{\mathbf{q}}}^{\mathrm{T}} + w_{\dot{\mathbf{q}}}^{\mathrm{T}}\right]_F \tag{3.50b}$$

$$\left[\boldsymbol{\Phi}_{\mathbf{q}}\boldsymbol{\mu}\right]_F = \left[g_{\boldsymbol{\lambda}}^{\mathrm{T}}\right]_F \tag{3.50c}$$

$$\left[\boldsymbol{\Phi}_{\mathbf{q}}\dot{\boldsymbol{\mu}} + \dot{\boldsymbol{\Phi}}_{\mathbf{q}}\boldsymbol{\mu}\right]_F = \left[\frac{\mathrm{d}g_{\boldsymbol{\lambda}}^{\mathrm{T}}}{\mathrm{d}t}\right]_F \tag{3.50d}$$

where equation (3.50d) is the derivative of (3.50c) and it was added to complete a full set of

equations to determine the initial adjoint variables. The process was described by [41] and consist of solving the two sets of equations formed with (3.50b), (3.50c) to obtain $\boldsymbol{\mu}_F$ and $\boldsymbol{\eta}_F$ and (3.50a), (3.50d) to obtain $\dot{\boldsymbol{\mu}}_F$ and $\boldsymbol{\gamma}_F$.

The gradient of the objective function is obtained with the remaining terms not canceled out in the variational equation (3.42) and additional terms (3.48), (3.49):

$$\nabla_{\boldsymbol{\rho}}\psi^{\mathrm{T}} = \left[w_{\boldsymbol{\rho}} + \boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{\Phi}_{\boldsymbol{\rho}} + \boldsymbol{\eta}^{\mathrm{T}}\left(\boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}\dot{\mathbf{q}} + \boldsymbol{\Phi}_{t\boldsymbol{\rho}}\right)\right]_F + \left[\left(\frac{\mathrm{d}g_{\ddot{\mathbf{q}}}}{\mathrm{dt}} - g_{\dot{\mathbf{q}}} - \dot{\boldsymbol{\mu}}^{\mathrm{T}}\mathbf{M} - \boldsymbol{\mu}^{\mathrm{T}}\left(\dot{\mathbf{M}} - \mathbf{C}\right)\right)\mathbf{q}_{\boldsymbol{\rho}}\right]_0 +$$
$$- \left[\left(g_{\ddot{\mathbf{q}}} - \boldsymbol{\mu}^{\mathrm{T}}\mathbf{M}\right)\dot{\mathbf{q}}_{\boldsymbol{\rho}}\right]_0 + \int_{t_0}^{t_F}\left(g_{\boldsymbol{\rho}} - \boldsymbol{\mu}^{\mathrm{T}}\left(\mathbf{M}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}^{\mathrm{T}}\boldsymbol{\lambda} - \mathbf{Q}_{\boldsymbol{\rho}}\right) - \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\boldsymbol{\Phi}_{\boldsymbol{\rho}}\right)\mathrm{dt} \tag{3.51}$$

The gradient depends on the solution of the EOM and on the adjoint differential and algebraic variables $\boldsymbol{\mu}$ and $\boldsymbol{\mu}_{\boldsymbol{\Phi}}$. The adjoint variables $\boldsymbol{\mu}$ and $\boldsymbol{\mu}_{\boldsymbol{\Phi}}$ are the solution of the adjoint system. The adjoint system derived from the index-3 EOM is an index-3 DAE.

This formulation of the sensitivity equations is not very convenient for the case of objective functions depending on $\ddot{\mathbf{q}}$ or $\boldsymbol{\lambda}$ due to the difficulty to obtain the derivatives of these terms, involving higher order derivatives of the states or Lagrange multipliers, normally not calculated.

The adjoint index-3 DAE (3.45) can be directly integrated backward in time:

$$\mathbf{M}^{\mathrm{T}}\ddot{\boldsymbol{\mu}} + \mathbf{A}^{\mathrm{T}}\dot{\boldsymbol{\mu}} + \mathbf{B}^{\mathrm{T}}\boldsymbol{\mu} + \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\boldsymbol{\mu}_{\boldsymbol{\Phi}} = g_{\mathbf{q}}^{\mathrm{T}} - \frac{\mathrm{d}g_{\dot{\mathbf{q}}}}{\mathrm{dt}}^{\mathrm{T}} + \frac{\mathrm{d}^2 g_{\ddot{\mathbf{q}}}}{\mathrm{dt}^2}^{\mathrm{T}} \tag{3.52a}$$
$$\boldsymbol{\Phi}_{\mathbf{q}}\boldsymbol{\mu} = g_{\boldsymbol{\lambda}}^{\mathrm{T}} \tag{3.52b}$$

The implicit trapezoidal rule equations for backward integration can be expressed as

$$\dot{\boldsymbol{\mu}}_*^n = -\frac{2}{h}\boldsymbol{\mu}_*^n + \hat{\dot{\boldsymbol{\mu}}}_*^{n+1}\,; \qquad \hat{\dot{\boldsymbol{\mu}}}_*^{n+1} = \frac{2}{h}\boldsymbol{\mu}_*^{n+1} - \dot{\boldsymbol{\mu}}_*^{n+1} \tag{3.53a}$$
$$\ddot{\boldsymbol{\mu}}_*^n = -\frac{4}{h^2}\boldsymbol{\mu}_*^n + \hat{\ddot{\boldsymbol{\mu}}}_*^{n+1}\,; \qquad \hat{\ddot{\boldsymbol{\mu}}}_*^{n+1} = -\frac{4}{h^2}\boldsymbol{\mu}_*^{n+1} + \frac{4}{h}\dot{\boldsymbol{\mu}}_*^{n+1} - \ddot{\boldsymbol{\mu}}_*^{n+1} \tag{3.53b}$$

where the subindex $*$ means that the same equations hold for $\boldsymbol{\mu}$ and $\boldsymbol{\mu}_{\boldsymbol{\Phi}}$.

Replacing equations (3.53) in (3.52)

$$\begin{bmatrix} \frac{4}{h^2}\mathbf{M}^{\mathrm{T}} + \frac{2}{h}\mathbf{A}^{\mathrm{T}} + \mathbf{B}^{\mathrm{T}} & \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}} \\ \boldsymbol{\Phi}_{\mathbf{q}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\mu}^n \\ \boldsymbol{\mu}_{\boldsymbol{\Phi}}^n \end{bmatrix} = \begin{bmatrix} g_{\mathbf{q}}^{\mathrm{T}} - \dot{g}_{\dot{\mathbf{q}}}^{\mathrm{T}} + \ddot{g}_{\ddot{\mathbf{q}}}^{\mathrm{T}} - \mathbf{M}^{\mathrm{T}}\ddot{\boldsymbol{\mu}}^{n+1} - \mathbf{A}^{\mathrm{T}}\dot{\boldsymbol{\mu}}^{n+1} \\ \mathbf{0} \end{bmatrix} \tag{3.54}$$

Observe that the system of equations (3.54) is very similar in its structure to (2.6a), except by the fact that (3.54) is linear in $\boldsymbol{\mu}^n$, $\boldsymbol{\mu}_{\boldsymbol{\Phi}}^n$ and therefore it doesn't need to be iterated. This similarity suggests the same scaling proposed before, leading to the following scaled system

$$\begin{bmatrix} \mathbf{M}^{\mathrm{T}} + \frac{h}{2}\mathbf{A}^{\mathrm{T}} + \frac{h^2}{4}\mathbf{B}^{\mathrm{T}} & \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}} \\ \boldsymbol{\Phi}_{\mathbf{q}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\mu}^n \\ \bar{\boldsymbol{\mu}}_{\boldsymbol{\Phi}}^n \end{bmatrix} = \begin{bmatrix} \frac{h^2}{4}\left(g_{\mathbf{q}}^{\mathrm{T}} - \dot{g}_{\dot{\mathbf{q}}}^{\mathrm{T}} + \ddot{g}_{\ddot{\mathbf{q}}}^{\mathrm{T}} - \mathbf{M}^{\mathrm{T}}\ddot{\boldsymbol{\mu}}^{n+1} - \mathbf{A}^{\mathrm{T}}\dot{\boldsymbol{\mu}}^{n+1}\right) \\ \mathbf{0} \end{bmatrix} \tag{3.55}$$

where $\bar{\boldsymbol{\mu}}_{\boldsymbol{\Phi}}^n = \left(h^2/4\right)\boldsymbol{\mu}_{\boldsymbol{\Phi}}^n$ are the scaled adjoint variables associated to the constraint equations.

3.2.2 Index-1 DAE formulation

Approach 1

The adjoint variable method for the index-1 formulation was obtained in [43] for the EOM considered as a first order system and objective functions with variable time limits t_F but not dependent either on $\boldsymbol{\lambda}$ or $\ddot{\mathbf{q}}_F$. In this section, the results are revisited also for objective functions depending on $\ddot{\mathbf{q}}_F$ and $\boldsymbol{\lambda}$ but fixed end time t_F.

Considering the EOM (2.8), the gradient (3.2) can be obtained like in section 3.2.1

$$\begin{aligned} \mathcal{L}(\boldsymbol{\rho}) &= w\left(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}}, \boldsymbol{\lambda}, \boldsymbol{\rho}\right)_{t_F} + \int_{t_0}^{t_F} g\left(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}}, \boldsymbol{\lambda}, \boldsymbol{\rho}\right) \mathrm{dt} \\ &- \int_{t_0}^{t_F} \boldsymbol{\mu}^{\mathrm{T}}\left(\mathbf{M}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\boldsymbol{\lambda} - \mathbf{Q}\right) \mathrm{dt} - \int_{t_0}^{t_F} \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\left(\boldsymbol{\Phi}_{\mathbf{q}}\ddot{\mathbf{q}} - \mathbf{c}\right) \mathrm{dt} \end{aligned} \tag{3.56}$$

Compute infinitesimal variations of $\mathcal{L}$ under infinitesimal variations $\delta\boldsymbol{\rho}$ (computation of $\delta\boldsymbol{\mu}$

and $\delta\boldsymbol{\lambda}$ not needed):

$$\begin{aligned}
&\delta\mathcal{L} = \left[w_{\mathbf{q}}\delta\mathbf{q} + w_{\dot{\mathbf{q}}}\delta\dot{\mathbf{q}} + w_{\ddot{\mathbf{q}}}\delta\ddot{\mathbf{q}} + w_{\boldsymbol{\lambda}}\delta\boldsymbol{\lambda} + w_{\boldsymbol{\rho}}\delta\boldsymbol{\rho}\right]_{t_F} \\
&+ \int_{t_0}^{t_F} \left(g_{\mathbf{q}} - \boldsymbol{\mu}^{\mathrm{T}}\left(\mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\boldsymbol{\lambda} + \mathbf{K}\right) - \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\left(\boldsymbol{\Phi}_{\mathbf{qq}}\ddot{\mathbf{q}} - \mathbf{c}_{\mathbf{q}}\right)\right)\delta\mathbf{q}\mathrm{d}t + \\
&\int_{t_0}^{t_F} \left(g_{\dot{\mathbf{q}}} - \boldsymbol{\mu}^{\mathrm{T}}\mathbf{C} + \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\mathbf{c}_{\dot{\mathbf{q}}}\right)\delta\dot{\mathbf{q}}\mathrm{d}t + \int_{t_0}^{t_F} \left(g_{\ddot{\mathbf{q}}} - \boldsymbol{\mu}^{\mathrm{T}}\mathbf{M} - \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathbf{q}}\right)\delta\ddot{\mathbf{q}}\mathrm{d}t + \int_{t_0}^{t_F} \left(g_{\boldsymbol{\lambda}} - \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\right)\delta\boldsymbol{\lambda}\mathrm{d}t + \\
&\int_{t_0}^{t_F} \left(g_{\boldsymbol{\rho}} - \boldsymbol{\mu}^{\mathrm{T}}\left(\mathbf{M}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}^{\mathrm{T}}\boldsymbol{\lambda} - \mathbf{Q}_{\boldsymbol{\rho}}\right) - \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\left(\boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}\ddot{\mathbf{q}} - \mathbf{c}_{\boldsymbol{\rho}}\right)\right)\delta\boldsymbol{\rho}\mathrm{d}t
\end{aligned} \tag{3.57}$$

Integrating by parts the integrals involving $\delta\dot{\mathbf{q}}, \delta\ddot{\mathbf{q}}$ leads to:

$$\begin{aligned}
&\delta\mathcal{L} = \left[w_{\mathbf{q}}\delta\mathbf{q} + w_{\dot{\mathbf{q}}}\delta\dot{\mathbf{q}} + w_{\ddot{\mathbf{q}}}\delta\ddot{\mathbf{q}} + w_{\boldsymbol{\lambda}}\delta\boldsymbol{\lambda} + w_{\boldsymbol{\rho}}\delta\boldsymbol{\rho}\right]_{t_F} + \\
&\int_{t_0}^{t_F} \left(g_{\mathbf{q}} - \boldsymbol{\mu}^{\mathrm{T}}\left(\mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\boldsymbol{\lambda} + \mathbf{K}\right) - \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\left(\boldsymbol{\Phi}_{\mathbf{qq}}\ddot{\mathbf{q}} - \mathbf{c}_{\mathbf{q}}\right)\right)\delta\mathbf{q}\mathrm{d}t + \\
&\left(g_{\dot{\mathbf{q}}} - \boldsymbol{\mu}^{\mathrm{T}}\mathbf{C} + \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\mathbf{c}_{\dot{\mathbf{q}}}\right)\delta\mathbf{q}\Big|_{t0}^{tF} - \int_{t_0}^{t_F} \left(\frac{\mathrm{d}g_{\dot{\mathbf{q}}}}{\mathrm{d}t} - \dot{\boldsymbol{\mu}}^{\mathrm{T}}\mathbf{C} - \boldsymbol{\mu}^{\mathrm{T}}\dot{\mathbf{C}} + \dot{\boldsymbol{\mu}}_{\boldsymbol{\Phi}}^{\mathrm{T}}\mathbf{c}_{\dot{\mathbf{q}}} + \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\frac{\mathrm{d}\mathbf{c}_{\dot{\mathbf{q}}}}{\mathrm{d}t}\right)\delta\mathbf{q}\mathrm{d}t + \\
&\left(g_{\ddot{\mathbf{q}}} - \boldsymbol{\mu}^{\mathrm{T}}\mathbf{M} - \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathbf{q}}\right)\delta\dot{\mathbf{q}}\Big|_{t0}^{tF} - \left(\frac{\mathrm{d}g_{\ddot{\mathbf{q}}}}{\mathrm{d}t} - \dot{\boldsymbol{\mu}}^{\mathrm{T}}\mathbf{M} - \boldsymbol{\mu}^{\mathrm{T}}\dot{\mathbf{M}} - \dot{\boldsymbol{\mu}}_{\boldsymbol{\Phi}}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathbf{q}} - \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\dot{\boldsymbol{\Phi}}_{\mathbf{q}}\right)\delta\mathbf{q}\Big|_{t0}^{tF} + \\
&\int_{t_0}^{t_F} \left(\frac{\mathrm{d}^2 g_{\ddot{\mathbf{q}}}}{\mathrm{d}t^2} - \ddot{\boldsymbol{\mu}}^{\mathrm{T}}\mathbf{M} - 2\dot{\boldsymbol{\mu}}^{\mathrm{T}}\dot{\mathbf{M}} - \boldsymbol{\mu}^{\mathrm{T}}\ddot{\mathbf{M}} - \ddot{\boldsymbol{\mu}}_{\boldsymbol{\Phi}}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathbf{q}} - 2\dot{\boldsymbol{\mu}}_{\boldsymbol{\Phi}}^{\mathrm{T}}\dot{\boldsymbol{\Phi}}_{\mathbf{q}} - \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\ddot{\boldsymbol{\Phi}}_{\mathbf{q}}\right)\delta\mathbf{q}\mathrm{d}t + \\
&\int_{t_0}^{t_F} \left(g_{\boldsymbol{\lambda}} - \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\right)\delta\boldsymbol{\lambda}\mathrm{d}t + \\
&\int_{t_0}^{t_F} \left(g_{\boldsymbol{\rho}} - \boldsymbol{\mu}^{\mathrm{T}}\left(\mathbf{M}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}^{\mathrm{T}}\boldsymbol{\lambda} - \mathbf{Q}_{\boldsymbol{\rho}}\right) - \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\left(\boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}\ddot{\mathbf{q}} - \mathbf{c}_{\boldsymbol{\rho}}\right)\right)\delta\boldsymbol{\rho}\mathrm{d}t
\end{aligned} \tag{3.58}$$

Canceling all the integral terms that involve $\delta\mathbf{q}$ and $\delta\boldsymbol{\lambda}$, the following adjoint DAE is obtained

$$\begin{aligned}
&\mathbf{M}^{\mathrm{T}}\ddot{\boldsymbol{\mu}} + \left(2\dot{\mathbf{M}} - \mathbf{C}\right)^{\mathrm{T}}\dot{\boldsymbol{\mu}} + \left(\mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\boldsymbol{\lambda} + \mathbf{K} - \dot{\mathbf{C}} + \ddot{\mathbf{M}}\right)^{\mathrm{T}}\boldsymbol{\mu} + \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\ddot{\boldsymbol{\mu}}_{\boldsymbol{\Phi}} + \left(\mathbf{c}_{\dot{\mathbf{q}}} + 2\dot{\boldsymbol{\Phi}}_{\mathbf{q}}\right)^{\mathrm{T}}\dot{\boldsymbol{\mu}}_{\boldsymbol{\Phi}} \\
&+ \left(\boldsymbol{\Phi}_{\mathbf{qq}}\ddot{\mathbf{q}} + \ddot{\boldsymbol{\Phi}}_{\mathbf{q}} - \mathbf{c}_{\mathbf{q}} + \frac{\mathrm{d}\mathbf{c}_{\dot{\mathbf{q}}}}{\mathrm{d}t}\right)^{\mathrm{T}}\boldsymbol{\mu}_{\boldsymbol{\Phi}} = g_{\mathbf{q}}^{\mathrm{T}} - \frac{\mathrm{d}g_{\dot{\mathbf{q}}}^{\mathrm{T}}}{\mathrm{d}t} + \frac{\mathrm{d}^2 g_{\ddot{\mathbf{q}}}^{\mathrm{T}}}{\mathrm{d}t^2}
\end{aligned} \tag{3.59a}$$

$$\boldsymbol{\Phi}_{\mathbf{q}}\boldsymbol{\mu} = g_{\boldsymbol{\lambda}} \tag{3.59b}$$

The adjoint system (3.59) can be easily proved to be an index-1 DAE in $\boldsymbol{\mu}$ and $\boldsymbol{\mu}_{\boldsymbol{\Phi}}$. The initial

conditions for the adjoint DAE are:

$$\left[\mathbf{M}^{\mathrm{T}}\dot{\boldsymbol{\mu}} + \left(\dot{\mathbf{M}} - \mathbf{C}\right)^{\mathrm{T}}\boldsymbol{\mu} + \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\dot{\boldsymbol{\mu}}_{\boldsymbol{\Phi}} + \left(\dot{\boldsymbol{\Phi}}_{\mathbf{q}} + \mathbf{c}_{\dot{\mathbf{q}}}\right)^{\mathrm{T}}\boldsymbol{\mu}_{\boldsymbol{\Phi}}\right]_F = \left[-w_{\mathbf{q}}^{\mathrm{T}} - g_{\dot{\mathbf{q}}}^{\mathrm{T}} + \frac{\mathrm{d}g_{\ddot{\mathbf{q}}}^{\mathrm{T}}}{\mathrm{d}t}\right]_F \tag{3.60a}$$

$$\left[\mathbf{M}^{\mathrm{T}}\boldsymbol{\mu} + \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\boldsymbol{\mu}_{\boldsymbol{\Phi}}\right]_F = \left[\left(w_{\dot{\mathbf{q}}} + g_{\ddot{\mathbf{q}}}\right)^{\mathrm{T}}\right]_F \tag{3.60b}$$

$$\left[w_{\ddot{\mathbf{q}}}\right]_F = \mathbf{0} \tag{3.60c}$$

$$\left[w_{\boldsymbol{\lambda}}\right]_F = \mathbf{0} \tag{3.60d}$$

$$\left[\boldsymbol{\Phi}_{\mathbf{q}}\boldsymbol{\mu}\right]_F = \left[g_{\boldsymbol{\lambda}}^{\mathrm{T}}\right]_F \tag{3.60e}$$

$$\left[\boldsymbol{\Phi}_{\mathbf{q}}\dot{\boldsymbol{\mu}} + \dot{\boldsymbol{\Phi}}_{\mathbf{q}}\boldsymbol{\mu}\right]_F = \left[\frac{\mathrm{d}g_{\boldsymbol{\lambda}}^{\mathrm{T}}}{\mathrm{d}t}\right]_F \tag{3.60f}$$

Observe that two incompatibility conditions (3.60c) and (3.60d), arise for the objective functionals whose final time term cannot depend on $\ddot{\mathbf{q}}$ or $\boldsymbol{\lambda}$. Moreover, the extra condition (3.60e) and its derivative (3.60f) were taken from (3.59b) particularized for the final time to complete a full set of equations, which allow to obtain the initial values of $\boldsymbol{\mu}$, $\dot{\boldsymbol{\mu}}$, $\boldsymbol{\mu}_{\boldsymbol{\Phi}}$ and $\dot{\boldsymbol{\mu}}_{\boldsymbol{\Phi}}$ at the final time t_F.

The following gradient of the objective function is obtained from the variational equation (3.58):

$$\begin{aligned}\nabla_{\boldsymbol{\rho}}\psi = {} & \left[w_{\boldsymbol{\rho}}^{\mathrm{T}}\right]_F - \left[\frac{\partial \mathbf{q}}{\partial \boldsymbol{\rho}}^{\mathrm{T}}\left(g_{\dot{\mathbf{q}}}^{\mathrm{T}} - \frac{\mathrm{d}g_{\ddot{\mathbf{q}}}^{\mathrm{T}}}{\mathrm{d}t} + \left(\dot{\mathbf{M}} - \mathbf{C}\right)^{\mathrm{T}}\boldsymbol{\mu} + \mathbf{M}^{\mathrm{T}}\dot{\boldsymbol{\mu}} + \left(\mathbf{c}_{\dot{\mathbf{q}}} + \dot{\boldsymbol{\Phi}}_{\mathbf{q}}\right)^{\mathrm{T}}\boldsymbol{\mu}_{\boldsymbol{\Phi}} + \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\dot{\boldsymbol{\mu}}_{\boldsymbol{\Phi}}\right)\right]_0 \\ & - \left[\frac{\partial \dot{\mathbf{q}}}{\partial \boldsymbol{\rho}}^{\mathrm{T}}\left(g_{\ddot{\mathbf{q}}}^{\mathrm{T}} - \mathbf{M}^{\mathrm{T}}\boldsymbol{\mu} - \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\boldsymbol{\mu}_{\boldsymbol{\Phi}}\right)\right]_0 \\ & + \int_{t_0}^{t_F}\left(g_{\boldsymbol{\rho}}^{\mathrm{T}} - \left(\mathbf{M}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}^{\mathrm{T}}\boldsymbol{\lambda} - \mathbf{Q}_{\boldsymbol{\rho}}\right)^{\mathrm{T}}\boldsymbol{\mu} - \left(\boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}\ddot{\mathbf{q}} - \mathbf{c}_{\boldsymbol{\rho}}\right)^{\mathrm{T}}\boldsymbol{\mu}_{\boldsymbol{\Phi}}\right)\mathrm{dt}\end{aligned} \tag{3.61}$$

The gradient depends on the solution of the EOM and on the adjoint differential and algebraic variables $\boldsymbol{\mu}$ and $\boldsymbol{\mu}_{\boldsymbol{\Phi}}$ that are the solution of the adjoint system.

Approach 2

Equations (3.59) and (3.61), are of theoretical interest, but of lower practical value because they can involve higher order derivatives of the state variables, as pointed out in [116]. A more convenient approach is to use equations (2.8) to express the variations $\delta\ddot{\mathbf{q}}$ and $\delta\boldsymbol{\lambda}$ in terms of the

variations of the states $\delta\mathbf{q}$ and $\delta\dot{\mathbf{q}}$. Taking the variation of (2.8)

$$\begin{bmatrix} \delta\ddot{\mathbf{q}} \\ \delta\boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{M} & \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}} \\ \boldsymbol{\Phi}_{\mathbf{q}} & \mathbf{0} \end{bmatrix}^{-1} \left(\begin{bmatrix} -\mathbf{K} - \mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} - \boldsymbol{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\boldsymbol{\lambda} \\ \mathbf{c}_{\mathbf{q}} - \boldsymbol{\Phi}_{\mathbf{qq}}\ddot{\mathbf{q}} \end{bmatrix} \delta\mathbf{q} + \begin{bmatrix} -\mathbf{C} \\ \mathbf{c}_{\dot{\mathbf{q}}} \end{bmatrix} \delta\dot{\mathbf{q}} \right)$$
$$+ \begin{bmatrix} \mathbf{M} & \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}} \\ \boldsymbol{\Phi}_{\mathbf{q}} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{Q}_{\boldsymbol{\rho}} - \mathbf{M}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} - \boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}^{\mathrm{T}}\boldsymbol{\lambda} \\ \mathbf{c}_{\boldsymbol{\rho}} - \boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}\ddot{\mathbf{q}} \end{bmatrix} \delta\boldsymbol{\rho} \tag{3.62}$$

leads to

$$\delta\ddot{\mathbf{q}} = \ddot{\mathbf{q}}_{\mathbf{q}}\delta\mathbf{q} + \ddot{\mathbf{q}}_{\dot{\mathbf{q}}}\delta\dot{\mathbf{q}} + \ddot{\mathbf{q}}_{\boldsymbol{\rho}}\delta\boldsymbol{\rho} \tag{3.63a}$$

$$\delta\boldsymbol{\lambda} = \boldsymbol{\lambda}_{\mathbf{q}}\delta\mathbf{q} + \boldsymbol{\lambda}_{\dot{\mathbf{q}}}\delta\dot{\mathbf{q}} + \boldsymbol{\lambda}_{\boldsymbol{\rho}}\delta\boldsymbol{\rho} \tag{3.63b}$$

where

$$\begin{bmatrix} \ddot{\mathbf{q}}_{\mathbf{q}} \\ \boldsymbol{\lambda}_{\mathbf{q}} \end{bmatrix} = \begin{bmatrix} \mathbf{M} & \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}} \\ \boldsymbol{\Phi}_{\mathbf{q}} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} -\mathbf{K} - \mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} - \boldsymbol{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\boldsymbol{\lambda} \\ \mathbf{c}_{\mathbf{q}} - \boldsymbol{\Phi}_{\mathbf{qq}}\ddot{\mathbf{q}} \end{bmatrix} \tag{3.64}$$

$$\begin{bmatrix} \ddot{\mathbf{q}}_{\dot{\mathbf{q}}} \\ \boldsymbol{\lambda}_{\dot{\mathbf{q}}} \end{bmatrix} = \begin{bmatrix} \mathbf{M} & \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}} \\ \boldsymbol{\Phi}_{\mathbf{q}} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} -\mathbf{C} \\ \mathbf{c}_{\dot{\mathbf{q}}} \end{bmatrix} \tag{3.65}$$

$$\begin{bmatrix} \ddot{\mathbf{q}}_{\boldsymbol{\rho}} \\ \boldsymbol{\lambda}_{\boldsymbol{\rho}} \end{bmatrix} = \begin{bmatrix} \mathbf{M} & \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}} \\ \boldsymbol{\Phi}_{\mathbf{q}} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{Q}_{\boldsymbol{\rho}} - \mathbf{M}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} - \boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}^{\mathrm{T}}\boldsymbol{\lambda} \\ \mathbf{c}_{\boldsymbol{\rho}} - \boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}\ddot{\mathbf{q}} \end{bmatrix} \tag{3.66}$$

Replacing equations (3.63a) and (3.63b) in (3.57) yields

$$\begin{aligned}
\delta\mathcal{L} = {} & \left[\left(w_{\mathbf{q}} + w_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\mathbf{q}} + w_{\boldsymbol{\lambda}}\boldsymbol{\lambda}_{\mathbf{q}}\right)\delta\mathbf{q} + \left(w_{\dot{\mathbf{q}}} + w_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}} + w_{\boldsymbol{\lambda}}\boldsymbol{\lambda}_{\dot{\mathbf{q}}}\right)\delta\dot{\mathbf{q}} + \left(w_{\boldsymbol{\rho}} + w_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}} + w_{\boldsymbol{\lambda}}\boldsymbol{\lambda}_{\boldsymbol{\rho}}\right)\delta\boldsymbol{\rho}\right]_{t_F} + \\
& \int_{t_0}^{t_F} \left(g_{\mathbf{q}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\mathbf{q}} + g_{\boldsymbol{\lambda}}\boldsymbol{\lambda}_{\mathbf{q}} - \boldsymbol{\mu}^{\mathrm{T}}\left(\mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\boldsymbol{\lambda} + \mathbf{K}\right) - \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\left(\boldsymbol{\Phi}_{\mathbf{qq}}\ddot{\mathbf{q}} - \mathbf{c}_{\mathbf{q}}\right)\right)\delta\mathbf{q}\mathrm{d}t + \\
& \int_{t_0}^{t_F} \left(g_{\dot{\mathbf{q}}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}} + g_{\boldsymbol{\lambda}}\boldsymbol{\lambda}_{\dot{\mathbf{q}}} - \boldsymbol{\mu}^{\mathrm{T}}\mathbf{C} + \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\mathbf{c}_{\dot{\mathbf{q}}}\right)\delta\dot{\mathbf{q}}\mathrm{d}t - \int_{t_0}^{t_F} \left(\boldsymbol{\mu}^{\mathrm{T}}\mathbf{M} + \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathbf{q}}\right)\delta\ddot{\mathbf{q}}\mathrm{d}t - \int_{t_0}^{t_F} \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\delta\boldsymbol{\lambda}\mathrm{d}t + \\
& \int_{t_0}^{t_F} \left(g_{\boldsymbol{\rho}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}} + g_{\boldsymbol{\lambda}}\boldsymbol{\lambda}_{\boldsymbol{\rho}} - \boldsymbol{\mu}^{\mathrm{T}}\left(\mathbf{M}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}^{\mathrm{T}}\boldsymbol{\lambda} - \mathbf{Q}_{\boldsymbol{\rho}}\right) - \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\left(\boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}\ddot{\mathbf{q}} - \mathbf{c}_{\boldsymbol{\rho}}\right)\right)\delta\boldsymbol{\rho}\mathrm{d}t
\end{aligned} \tag{3.67}$$

Observe that only the variations $\delta\ddot{\mathbf{q}}$ and $\delta\boldsymbol{\lambda}$ associated with the objective function terms w and g were eliminated in (3.67). Integration by parts can be applied to the remaining $\delta\ddot{\mathbf{q}}$ and $\delta\dot{\mathbf{q}}$

terms

$$
\begin{aligned}
&\delta\mathcal{L} = \left[\left(w_{\mathbf{q}} + w_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\mathbf{q}} + w_{\boldsymbol{\lambda}}\boldsymbol{\lambda}_{\mathbf{q}}\right)\delta\mathbf{q} + \left(w_{\dot{\mathbf{q}}} + w_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}} + w_{\boldsymbol{\lambda}}\boldsymbol{\lambda}_{\dot{\mathbf{q}}}\right)\delta\dot{\mathbf{q}} + \left(w_{\boldsymbol{\rho}} + w_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}} + w_{\boldsymbol{\lambda}}\boldsymbol{\lambda}_{\boldsymbol{\rho}}\right)\delta\boldsymbol{\rho}\right]_{t_F} \\
&\int_{t_0}^{t_F}\left(g_{\mathbf{q}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\mathbf{q}} + g_{\boldsymbol{\lambda}}\boldsymbol{\lambda}_{\mathbf{q}} - \boldsymbol{\mu}^{\mathrm{T}}\left(\mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\boldsymbol{\lambda} + \mathbf{K}\right) - \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\left(\boldsymbol{\Phi}_{\mathbf{qq}}\ddot{\mathbf{q}} - \mathbf{c}_{\mathbf{q}}\right)\right)\delta\mathbf{q}\mathrm{d}t + \\
&\left(g_{\dot{\mathbf{q}}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}} + g_{\boldsymbol{\lambda}}\boldsymbol{\lambda}_{\dot{\mathbf{q}}} - \boldsymbol{\mu}^{\mathrm{T}}\mathbf{C} + \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\mathbf{c}_{\dot{\mathbf{q}}}\right)\delta\mathbf{q}\Big|_{t0}^{tF} - \\
&\int_{t_0}^{t_F}\left(\frac{\mathrm{d}\left(g_{\dot{\mathbf{q}}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}} + g_{\boldsymbol{\lambda}}\boldsymbol{\lambda}_{\dot{\mathbf{q}}}\right)}{\mathrm{d}t} - \dot{\boldsymbol{\mu}}^{\mathrm{T}}\mathbf{C} - \boldsymbol{\mu}^{\mathrm{T}}\dot{\mathbf{C}} + \dot{\boldsymbol{\mu}}_{\boldsymbol{\Phi}}^{\mathrm{T}}\mathbf{c}_{\dot{\mathbf{q}}} + \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\frac{\mathrm{d}\mathbf{c}_{\dot{\mathbf{q}}}}{\mathrm{d}t}\right)\delta\mathbf{q}\mathrm{d}t - \\
&\left(\boldsymbol{\mu}^{\mathrm{T}}\mathbf{M} + \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathbf{q}}\right)\delta\dot{\mathbf{q}}\Big|_{t0}^{tF} + \left(\dot{\boldsymbol{\mu}}^{\mathrm{T}}\mathbf{M} + \boldsymbol{\mu}^{\mathrm{T}}\dot{\mathbf{M}} + \dot{\boldsymbol{\mu}}_{\boldsymbol{\Phi}}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathbf{q}} + \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\dot{\boldsymbol{\Phi}}_{\mathbf{q}}\right)\delta\mathbf{q}\Big|_{t0}^{tF} - \\
&\int_{t_0}^{t_F}\left(\ddot{\boldsymbol{\mu}}^{\mathrm{T}}\mathbf{M} + 2\dot{\boldsymbol{\mu}}^{\mathrm{T}}\dot{\mathbf{M}} + \boldsymbol{\mu}^{\mathrm{T}}\ddot{\mathbf{M}} + \ddot{\boldsymbol{\mu}}_{\boldsymbol{\Phi}}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathbf{q}} + 2\dot{\boldsymbol{\mu}}_{\boldsymbol{\Phi}}^{\mathrm{T}}\dot{\boldsymbol{\Phi}}_{\mathbf{q}} + \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\ddot{\boldsymbol{\Phi}}_{\mathbf{q}}\right)\delta\mathbf{q}\mathrm{d}t - \\
&\int_{t_0}^{t_F}\left(\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\right)\delta\boldsymbol{\lambda}\mathrm{d}t + \int_{t_0}^{t_F}\left(g_{\boldsymbol{\rho}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}} + g_{\boldsymbol{\lambda}}\boldsymbol{\lambda}_{\boldsymbol{\rho}} - \boldsymbol{\mu}^{\mathrm{T}}\left(\mathbf{M}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}^{\mathrm{T}}\boldsymbol{\lambda} - \mathbf{Q}_{\boldsymbol{\rho}}\right) - \boldsymbol{\mu}_{\boldsymbol{\Phi}}^{\mathrm{T}}\left(\boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}\ddot{\mathbf{q}} - \mathbf{c}_{\boldsymbol{\rho}}\right)\right)\delta\boldsymbol{\rho}\mathrm{d}t
\end{aligned}
\tag{3.68}
$$

Canceling all the integral terms that involve $\delta\mathbf{q}$ and $\delta\boldsymbol{\lambda}$ leads to the following adjoint DAE

$$
\begin{aligned}
&\mathbf{M}^{\mathrm{T}}\ddot{\boldsymbol{\mu}} + \left(2\dot{\mathbf{M}} - \mathbf{C}\right)^{\mathrm{T}}\dot{\boldsymbol{\mu}} + \left(\mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{qq}}^{\mathrm{T}}\boldsymbol{\lambda} + \mathbf{K} - \dot{\mathbf{C}} + \ddot{\mathbf{M}}\right)^{\mathrm{T}}\boldsymbol{\mu} + \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\ddot{\boldsymbol{\mu}}_{\boldsymbol{\Phi}} + \left(\mathbf{c}_{\dot{\mathbf{q}}} + 2\dot{\boldsymbol{\Phi}}_{\mathbf{q}}\right)^{\mathrm{T}}\dot{\boldsymbol{\mu}}_{\boldsymbol{\Phi}} \\
&+ \left(\boldsymbol{\Phi}_{\mathbf{qq}}\ddot{\mathbf{q}} + \ddot{\boldsymbol{\Phi}}_{\mathbf{q}} - \mathbf{c}_{\mathbf{q}} + \frac{\mathrm{d}\mathbf{c}_{\dot{\mathbf{q}}}}{\mathrm{d}t}\right)^{\mathrm{T}}\boldsymbol{\mu}_{\boldsymbol{\Phi}} = \left(g_{\mathbf{q}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\mathbf{q}} + g_{\boldsymbol{\lambda}}\boldsymbol{\lambda}_{\mathbf{q}}\right)^{\mathrm{T}} - \frac{\mathrm{d}\left(g_{\dot{\mathbf{q}}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}} + g_{\boldsymbol{\lambda}}\boldsymbol{\lambda}_{\dot{\mathbf{q}}}\right)^{\mathrm{T}}}{\mathrm{d}t} \qquad (3.69a) \\
&\boldsymbol{\Phi}_{\mathbf{q}}\boldsymbol{\mu} = \mathbf{0} \qquad (3.69b)
\end{aligned}
$$

Observe that the last term in equation (3.69a) can be difficult to obtain, because the temporal derivatives of a functional that depends on the accelerations and Lagrange multipliers can involve temporal derivatives of the accelerations and temporal derivatives of the Lagrange multipliers that normally are not calculated by the integrator.

The initial conditions for the adjoint are the following

$$\left[\mathbf{M}^{\mathrm{T}}\dot{\boldsymbol{\mu}} + \left(\dot{\mathbf{M}} - \mathbf{C}\right)^{\mathrm{T}}\boldsymbol{\mu} + \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\dot{\boldsymbol{\mu}}_{\boldsymbol{\Phi}} + \left(\dot{\boldsymbol{\Phi}}_{\mathbf{q}} + \mathbf{c}_{\dot{\mathbf{q}}}\right)^{\mathrm{T}}\boldsymbol{\mu}_{\boldsymbol{\Phi}}\right]_F =$$
$$-\left[\left(w_{\mathbf{q}} + w_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\mathbf{q}} + w_{\boldsymbol{\lambda}}\boldsymbol{\lambda}_{\mathbf{q}} + g_{\dot{\mathbf{q}}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}} + g_{\boldsymbol{\lambda}}\boldsymbol{\lambda}_{\dot{\mathbf{q}}}\right)^{\mathrm{T}}\right]_F \tag{3.70a}$$
$$\left[\mathbf{M}^{\mathrm{T}}\boldsymbol{\mu} + \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\boldsymbol{\mu}_{\boldsymbol{\Phi}}\right]_F = \left[\left(w_{\dot{\mathbf{q}}} + w_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}} + w_{\boldsymbol{\lambda}}\boldsymbol{\lambda}_{\dot{\mathbf{q}}}\right)^{\mathrm{T}}\right]_F \tag{3.70b}$$
$$\left[\boldsymbol{\Phi}_{\mathbf{q}}\boldsymbol{\mu}\right]_F = \mathbf{0} \tag{3.70c}$$
$$\left[\boldsymbol{\Phi}_{\mathbf{q}}\dot{\boldsymbol{\mu}} + \dot{\boldsymbol{\Phi}}_{\mathbf{q}}\boldsymbol{\mu}\right]_F = \mathbf{0} \tag{3.70d}$$

Finally, the gradient

$$\nabla_{\boldsymbol{\rho}}\psi = \left[\left(w_{\boldsymbol{\rho}} + w_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}} + w_{\boldsymbol{\lambda}}\boldsymbol{\lambda}_{\boldsymbol{\rho}}\right)^{\mathrm{T}}\right]_F -$$
$$\left[\frac{\partial \mathbf{q}}{\partial \boldsymbol{\rho}}^{\mathrm{T}}\left(\left(g_{\dot{\mathbf{q}}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}} + g_{\boldsymbol{\lambda}}\boldsymbol{\lambda}_{\dot{\mathbf{q}}}\right)^{\mathrm{T}} + \left(\dot{\mathbf{M}} - \mathbf{C}\right)^{\mathrm{T}}\boldsymbol{\mu} + \mathbf{M}^{\mathrm{T}}\dot{\boldsymbol{\mu}} + \left(\mathbf{c}_{\dot{\mathbf{q}}} + \dot{\boldsymbol{\Phi}}_{\mathbf{q}}\right)^{\mathrm{T}}\boldsymbol{\mu}_{\boldsymbol{\Phi}} + \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\dot{\boldsymbol{\mu}}_{\boldsymbol{\Phi}}\right)\right]_0 +$$
$$\left[\frac{\partial \dot{\mathbf{q}}}{\partial \boldsymbol{\rho}}^{\mathrm{T}}\left(\mathbf{M}^{\mathrm{T}}\boldsymbol{\mu} + \boldsymbol{\Phi}_{\mathbf{q}}^{\mathrm{T}}\boldsymbol{\mu}_{\boldsymbol{\Phi}}\right)\right]_0 +$$
$$\int_{t_0}^{t_F}\left(\left(g_{\boldsymbol{\rho}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}} + g_{\boldsymbol{\lambda}}\boldsymbol{\lambda}_{\boldsymbol{\rho}}\right)^{\mathrm{T}} - \left(\mathbf{M}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}^{\mathrm{T}}\boldsymbol{\lambda} - \mathbf{Q}_{\boldsymbol{\rho}}\right)^{\mathrm{T}}\boldsymbol{\mu} - \left(\boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}\ddot{\mathbf{q}} - \mathbf{c}_{\boldsymbol{\rho}}\right)^{\mathrm{T}}\boldsymbol{\mu}_{\boldsymbol{\Phi}}\right)\mathrm{dt} \tag{3.71}$$

3.2.3 The penalty formulation

The adjoint variable method seeks to obtain the sensitivity of a cost function, ψ, with respect to the set of parameters $\boldsymbol{\rho}$. For practical applications, very general cost functions depend not only on positions and velocities, but also on accelerations and reaction forces:

$$\psi = w\left(\mathbf{q}_F, \dot{\mathbf{q}}_F, \ddot{\mathbf{q}}_F, \boldsymbol{\rho}, \boldsymbol{\lambda}_F^*\right) + \int_{t_0}^{t_F} g\left(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}}, \boldsymbol{\rho}, \boldsymbol{\lambda}^*\right)\mathrm{dt} \tag{3.72}$$

Equations of motion written as a first-order ODE system

The system (2.17a) can be transformed into a first order system by simply defining a new set of variables by the relation $\dot{\mathbf{q}} = \mathbf{v}$,

$$\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{M}} \end{bmatrix} \begin{bmatrix} \dot{\mathbf{q}} \\ \dot{\mathbf{v}} \end{bmatrix} = \begin{bmatrix} \mathbf{v} \\ \bar{\mathbf{Q}} \end{bmatrix} \quad \Leftrightarrow \quad \hat{\mathbf{M}}(\mathbf{y}, \boldsymbol{\rho})\, \dot{\mathbf{y}} = \hat{\mathbf{Q}}(t, \mathbf{y}, \boldsymbol{\rho}) \tag{3.73}$$

In (3.73), the new state vector is $\mathbf{y} = \begin{bmatrix} \mathbf{q}^{\mathrm{T}} & \mathbf{v}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$. Taking the inverse of the leading matrix, the system (3.73) can be expressed as a first order explicit ODE

$$\dot{\mathbf{y}} = \hat{\mathbf{M}}^{-1}(\mathbf{y}, \boldsymbol{\rho})\, \hat{\mathbf{Q}}(t, \mathbf{y}, \boldsymbol{\rho}) = \mathbf{f}(t, \mathbf{y}, \boldsymbol{\rho}) \tag{3.74}$$

Similarly, the objective function (3.72) can be expressed as a function of the first order states

$$\psi = w(\mathbf{y}_F, \dot{\mathbf{y}}_F, \boldsymbol{\rho}_F, \boldsymbol{\lambda}_F^*) + \int_{t_0}^{t_F} g(\mathbf{y}, \dot{\mathbf{y}}, \boldsymbol{\rho}, \boldsymbol{\lambda}^*)\, \mathrm{dt} \tag{3.75}$$

Following [117], we consider the following Lagrangian, given by the cost function subject to the EOM constraints

$$\mathcal{L}(\boldsymbol{\rho}) = \psi - \int_{t_0}^{t_F} \boldsymbol{\mu}^{\mathrm{T}} \left(\dot{\mathbf{y}} - \mathbf{f}(t, \mathbf{y}, \boldsymbol{\rho})\right) \mathrm{dt} \tag{3.76}$$

where $\boldsymbol{\mu}$ is the vector of Lagrange multipliers or adjoint variables. Applying variational calculus

$$\delta\mathcal{L} = \delta\psi - \int_{t_0}^{t_F} \delta\boldsymbol{\mu}^{\mathrm{T}} \left(\dot{\mathbf{y}} - \mathbf{f}(t, \mathbf{y}, \boldsymbol{\rho})\right) \mathrm{dt} - \int_{t_0}^{t_F} \boldsymbol{\mu}^{\mathrm{T}} \left(\delta\dot{\mathbf{y}} - \mathbf{f}_{\mathbf{y}}\delta\mathbf{y} - \mathbf{f}_{\boldsymbol{\rho}}\delta\boldsymbol{\rho}\right) \mathrm{dt} \tag{3.77}$$

The central term vanishes if the EOM are fulfilled at each time step.

The variation of the cost function is

$$\delta\psi \quad = \quad \left(w_{\mathbf{y}}\delta\mathbf{y} + w_{\dot{\mathbf{y}}}\delta\dot{\mathbf{y}} + w_{\boldsymbol{\rho}}\delta\boldsymbol{\rho} + w_{\boldsymbol{\lambda}^*}\delta\boldsymbol{\lambda}^*\right)_F \; + \; \int_{t_0}^{t_F} \left(g_{\mathbf{y}}\delta\mathbf{y} + g_{\dot{\mathbf{y}}}\delta\dot{\mathbf{y}} + g_{\boldsymbol{\rho}}\delta\boldsymbol{\rho} + g_{\boldsymbol{\lambda}^*}\delta\boldsymbol{\lambda}^*\right) \mathrm{dt} \tag{3.78}$$

From Eqn. (2.20)

$$\delta\boldsymbol{\lambda}^* = \alpha\left(\delta\ddot{\boldsymbol{\Phi}} + 2\xi\omega\delta\dot{\boldsymbol{\Phi}} + \omega^2\delta\boldsymbol{\Phi}\right) \tag{3.79a}$$

where

$$\delta\ddot{\mathbf{\Phi}} = \mathbf{\Phi_q}\delta\ddot{\mathbf{q}} + \left(\mathbf{\Phi_{qq}}\dot{\mathbf{q}} + \dot{\mathbf{\Phi}}_{\mathbf{q}} + \mathbf{\Phi}_{t\mathbf{q}}\right)\delta\dot{\mathbf{q}} + \left(\mathbf{\Phi_{qq}}\ddot{\mathbf{q}} + \left(\dot{\mathbf{\Phi}}_{\mathbf{q}}\right)_{\mathbf{q}}\dot{\mathbf{q}} + \left(\dot{\mathbf{\Phi}}_{t}\right)_{\mathbf{q}}\right)\delta\mathbf{q}$$
$$+ \left(\mathbf{\Phi_{q\rho}}\ddot{\mathbf{q}} + \left(\dot{\mathbf{\Phi}}_{\mathbf{q}}\right)_{\boldsymbol{\rho}}\dot{\mathbf{q}} + \left(\dot{\mathbf{\Phi}}_{t}\right)_{\boldsymbol{\rho}}\right)\delta\boldsymbol{\rho} \tag{3.79b}$$

$$\delta\dot{\mathbf{\Phi}} = \mathbf{\Phi_q}\delta\dot{\mathbf{q}} + \left(\mathbf{\Phi_{qq}}\dot{\mathbf{q}} + \mathbf{\Phi}_{t\mathbf{q}}\right)\delta\mathbf{q} + \left(\mathbf{\Phi_{q\rho}}\dot{\mathbf{q}} + \mathbf{\Phi}_{t\boldsymbol{\rho}}\right)\delta\boldsymbol{\rho} \tag{3.79c}$$

$$\delta\mathbf{\Phi} = \mathbf{\Phi_q}\delta\mathbf{q} + \mathbf{\Phi}_{\boldsymbol{\rho}}\delta\boldsymbol{\rho} \tag{3.79d}$$

Grouping together the terms associated to $\delta\ddot{\mathbf{q}}$, $\delta\dot{\mathbf{q}}$, $\delta\mathbf{q}$, $\delta\boldsymbol{\rho}$ and taking into account that $\mathbf{y} = \begin{bmatrix} \mathbf{q}^{\mathrm{T}} & \mathbf{v}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$, Eqn. (3.79a) becomes

$$\delta\boldsymbol{\lambda}^* = \boldsymbol{\lambda}^*_{\dot{\mathbf{y}}}\delta\dot{\mathbf{y}} + \boldsymbol{\lambda}^*_{\mathbf{y}}\delta\mathbf{y} + \boldsymbol{\lambda}^*_{\boldsymbol{\rho}}\delta\boldsymbol{\rho} \tag{3.80}$$

Identifying the common terms in (3.79a) and (3.80) and using the identity $\mathbf{v} = \dot{\mathbf{q}}$ one obtains

$$\boldsymbol{\lambda}_{\mathbf{y}} = \begin{bmatrix} \boldsymbol{\lambda}^*_{\mathbf{q}} & \boldsymbol{\lambda}^*_{\mathbf{v}} \end{bmatrix} \tag{3.81a}$$

$$\boldsymbol{\lambda}_{\dot{\mathbf{y}}} = \begin{bmatrix} \mathbf{0} & \boldsymbol{\lambda}^*_{\dot{\mathbf{v}}} \end{bmatrix} \tag{3.81b}$$

$$\boldsymbol{\lambda}^*_{\dot{\mathbf{v}}} = \alpha\mathbf{\Phi_q} \tag{3.81c}$$

$$\boldsymbol{\lambda}^*_{\mathbf{v}} = \alpha\left[\mathbf{\Phi_{qq}}\mathbf{v} + \dot{\mathbf{\Phi}}_{\mathbf{q}} + \mathbf{\Phi}_{t\mathbf{q}} + 2\xi\omega\mathbf{\Phi_q}\right] \tag{3.81d}$$

$$\boldsymbol{\lambda}^*_{\mathbf{q}} = \alpha\left[\mathbf{\Phi_{qq}}\dot{\mathbf{v}} + \left(\dot{\mathbf{\Phi}}_{\mathbf{q}}\right)_{\mathbf{q}}\mathbf{v} + \left(\dot{\mathbf{\Phi}}_{t}\right)_{\mathbf{q}} + 2\xi\omega\left(\mathbf{\Phi_{qq}}\mathbf{v} + \mathbf{\Phi}_{t\mathbf{q}}\right) + \omega^2\mathbf{\Phi_q}\right] \tag{3.81e}$$

$$\boldsymbol{\lambda}^*_{\boldsymbol{\rho}} = \alpha\left[\mathbf{\Phi_{q\rho}}\dot{\mathbf{v}} + \left(\dot{\mathbf{\Phi}}_{\mathbf{q}}\right)_{\boldsymbol{\rho}}\mathbf{v} + \left(\dot{\mathbf{\Phi}}_{t}\right)_{\boldsymbol{\rho}} + 2\xi\omega\left(\mathbf{\Phi_{q\rho}}\mathbf{v} + \mathbf{\Phi}_{t\boldsymbol{\rho}}\right) + \omega^2\mathbf{\Phi}_{\boldsymbol{\rho}}\right] \tag{3.81f}$$

Replacing (3.80) in (3.78)

$$\delta\psi = \left[\left(w_{\mathbf{y}} + w_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\mathbf{y}}\right)\delta\mathbf{y} + \left(w_{\dot{\mathbf{y}}} + w_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\dot{\mathbf{y}}}\right)\delta\dot{\mathbf{y}} + \left(w_{\boldsymbol{\rho}} + w_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\boldsymbol{\rho}}\right)\delta\boldsymbol{\rho}\right]_F + \int_{t_0}^{t_F}\left[\left(g_{\mathbf{y}} + g_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\mathbf{y}}\right)\delta\mathbf{y}\right.$$
$$\left. + \left(g_{\dot{\mathbf{y}}} + g_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\dot{\mathbf{y}}}\right)\delta\dot{\mathbf{y}} + \left(g_{\boldsymbol{\rho}} + g_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\boldsymbol{\rho}}\right)\delta\boldsymbol{\rho}\right]\mathrm{dt} \tag{3.82}$$

For convenience, $\delta\dot{\mathbf{y}}$ in (3.82) can be expressed as a function of $\delta\mathbf{y}$. Differentiating Eqn. (3.74)

$$\delta\dot{\mathbf{y}} = \mathbf{f}_{\mathbf{y}}\delta\mathbf{y} + \mathbf{f}_{\boldsymbol{\rho}}\delta\boldsymbol{\rho} \tag{3.83}$$

and replacing Eqn. (3.83) in (3.82) leads to

$$\delta\psi = \left[\left(w_{\mathbf{y}} + w_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\mathbf{y}} + \left(w_{\dot{\mathbf{y}}} + w_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\dot{\mathbf{y}}}\right)\mathbf{f}_{\mathbf{y}}\right)\delta\mathbf{y} + \left(w_{\boldsymbol{\rho}} + w_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\boldsymbol{\rho}} + \left(w_{\dot{\mathbf{y}}} + w_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\dot{\mathbf{y}}}\right)\mathbf{f}_{\boldsymbol{\rho}}\right)\delta\boldsymbol{\rho}\right]_F$$
$$+ \int_{t_0}^{t_F}\left[\left(g_{\mathbf{y}} + g_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\mathbf{y}} + \left(g_{\dot{\mathbf{y}}} + g_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\dot{\mathbf{y}}}\right)\mathbf{f}_{\mathbf{y}}\right)\delta\mathbf{y} + \left(g_{\boldsymbol{\rho}} + g_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\boldsymbol{\rho}} + \left(g_{\dot{\mathbf{y}}} + g_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\dot{\mathbf{y}}}\right)\mathbf{f}_{\boldsymbol{\rho}}\right)\delta\boldsymbol{\rho}\right]\mathrm{dt} \quad (3.84)$$

The variation of the full Lagrangian (3.77) can be obtained by replacing (3.84) in (3.77)

$$\delta\mathcal{L} = \left[\left(w_{\mathbf{y}} + w_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\mathbf{y}} + \left(w_{\dot{\mathbf{y}}} + w_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\dot{\mathbf{y}}}\right)\mathbf{f}_{\mathbf{y}}\right)\delta\mathbf{y} + \left(w_{\boldsymbol{\rho}} + w_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\boldsymbol{\rho}} + \left(w_{\dot{\mathbf{y}}} + w_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\dot{\mathbf{y}}}\right)\mathbf{f}_{\boldsymbol{\rho}}\right)\delta\boldsymbol{\rho}\right]_F +$$
$$\int_{t_0}^{t_F}\left[\left(g_{\mathbf{y}} + g_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\mathbf{y}} + \left(\boldsymbol{\mu}^{\mathrm{T}} + g_{\dot{\mathbf{y}}} + g_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\dot{\mathbf{y}}}\right)\mathbf{f}_{\mathbf{y}}\right)\delta\mathbf{y} + \left(g_{\boldsymbol{\rho}} + g_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\boldsymbol{\rho}} + \left(\boldsymbol{\mu}^{\mathrm{T}} + g_{\dot{\mathbf{y}}} + g_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\dot{\mathbf{y}}}\right)\mathbf{f}_{\boldsymbol{\rho}}\right)\delta\boldsymbol{\rho} - \boldsymbol{\mu}^{\mathrm{T}}\delta\dot{\mathbf{y}}\right]\mathrm{dt} \quad (3.85)$$

In Eqn. (3.85), the variation of the parameters $\delta\boldsymbol{\rho}$ is known, and variations $\delta\mathbf{y}$ and $\delta\dot{\mathbf{y}}$ could be calculated by solving the linearized form of the EOM (3.74), but this is computationally expensive. Instead of calculating them, the idea is to cancel these variations. Integrating by parts the integral terms involving $\delta\dot{\mathbf{y}}$ the variation can be removed from the integral

$$\int_{t_0}^{t_F} -\boldsymbol{\mu}^{\mathrm{T}}\delta\dot{\mathbf{y}}\mathrm{dt} = -\boldsymbol{\mu}^{\mathrm{T}}\delta\mathbf{y}\big|_{t_0}^{t_F} + \int_{t_0}^{t_F}\dot{\boldsymbol{\mu}}^{\mathrm{T}}\delta\mathbf{y}\mathrm{dt} \quad (3.86)$$

Therefore

$$\delta\mathcal{L} = \left[\left(w_{\mathbf{y}} + w_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\mathbf{y}} + \left(w_{\dot{\mathbf{y}}} + w_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\dot{\mathbf{y}}}\right)\mathbf{f}_{\mathbf{y}} - \boldsymbol{\mu}^{\mathrm{T}}\right)\delta\mathbf{y} + \left(w_{\boldsymbol{\rho}} + w_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\boldsymbol{\rho}} + \left(w_{\dot{\mathbf{y}}} + w_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\dot{\mathbf{y}}}\right)\mathbf{f}_{\boldsymbol{\rho}}\right)\delta\boldsymbol{\rho}\right]_F + \left[\boldsymbol{\mu}^{\mathrm{T}}\delta\mathbf{y}\right]_0$$
$$+ \int_{t_0}^{t_F}\left[\left(\dot{\boldsymbol{\mu}}^{\mathrm{T}} + g_{\mathbf{y}} + g_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\mathbf{y}} + \left(\boldsymbol{\mu}^{\mathrm{T}} + g_{\dot{\mathbf{y}}} + g_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\dot{\mathbf{y}}}\right)\mathbf{f}_{\mathbf{y}}\right)\delta\mathbf{y} + \left(g_{\boldsymbol{\rho}} + g_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\boldsymbol{\rho}} + \left(\boldsymbol{\mu}^{\mathrm{T}} + g_{\dot{\mathbf{y}}} + g_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\dot{\mathbf{y}}}\right)\mathbf{f}_{\boldsymbol{\rho}}\right)\delta\boldsymbol{\rho}\right]\mathrm{dt} \quad (3.87)$$

In Eqn. (3.87) it is possible to cancel $\delta\mathbf{y}$ by choosing $\boldsymbol{\mu}$ to be the solution of following adjoint ODE system

$$\dot{\boldsymbol{\mu}} = -\mathbf{f}_{\mathbf{y}}^{\mathrm{T}}\left(\boldsymbol{\mu} + g_{\dot{\mathbf{y}}}^{\mathrm{T}} + \boldsymbol{\lambda}^{*\mathrm{T}}_{\dot{\mathbf{y}}}g^{\mathrm{T}}_{\boldsymbol{\lambda}^*}\right) - g^{\mathrm{T}}_{\mathbf{y}} - \boldsymbol{\lambda}^{*\mathrm{T}}_{\mathbf{y}}g^{\mathrm{T}}_{\boldsymbol{\lambda}^*} \quad (3.88a)$$

$$\boldsymbol{\mu}_F = \left[w^{\mathrm{T}}_{\mathbf{y}} + \boldsymbol{\lambda}^{*\mathrm{T}}_{\mathbf{y}}w^{\mathrm{T}}_{\boldsymbol{\lambda}^*} + \mathbf{f}^{\mathrm{T}}_{\mathbf{y}}\left(w_{\dot{\mathbf{y}}} + w_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}^*_{\dot{\mathbf{y}}}\right)^{\mathrm{T}}\right]_F \quad (3.88b)$$

The adjoint system (3.88) is a first order linear ODE in $\boldsymbol{\mu}$. Since the initial conditions (3.88b)

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 Eqn. (3.87) the gradient of the cost function with respect to parameters can be obtained as

$$\begin{aligned}\nabla_{\boldsymbol{\rho}}\psi = &\left[w_{\boldsymbol{\rho}}^{\mathrm{T}} + \boldsymbol{\lambda}_{\boldsymbol{\rho}}^{*\mathrm{T}} w_{\boldsymbol{\lambda}^*}^{\mathrm{T}} + \mathbf{f}_{\boldsymbol{\rho}}^{\mathrm{T}}\left(w_{\dot{\mathbf{y}}} + w_{\boldsymbol{\lambda}^*}\boldsymbol{\lambda}_{\dot{\mathbf{y}}}^{*}\right)^{\mathrm{T}}\right]_F + \left[\frac{\partial \mathbf{y}^{\mathrm{T}}}{\partial \boldsymbol{\rho}}\boldsymbol{\mu}\right]_0 \\ &+ \int_{t_0}^{t_F}\left[\mathbf{f}_{\boldsymbol{\rho}}^{\mathrm{T}}\left(\boldsymbol{\mu} + g_{\dot{\mathbf{y}}}^{\mathrm{T}} + \boldsymbol{\lambda}_{\dot{\mathbf{y}}}^{*\mathrm{T}} g_{\boldsymbol{\lambda}^*}^{\mathrm{T}}\right) + g_{\boldsymbol{\rho}}^{\mathrm{T}} + \boldsymbol{\lambda}_{\boldsymbol{\rho}}^{*\mathrm{T}} g_{\boldsymbol{\lambda}^*}^{\mathrm{T}}\right]\mathrm{dt}\end{aligned} \tag{3.89}$$

where the identity $\delta\psi = \delta L$ was used. This holds if the EOM are satisfied, as can be seen from Eqn. (3.76).

In Eqns. (3.89) and (3.88) the derivatives of function g are known, since the objective function has a known expression. The derivatives of $\mathbf{f}$ are obtained using (3.73) as

$$\hat{\mathbf{M}}\frac{\partial \mathbf{f}}{\partial \mathbf{y}} + \hat{\mathbf{M}}_{\mathbf{y}}\mathbf{f} = \frac{\partial \hat{\mathbf{Q}}}{\partial \mathbf{y}} \quad \Rightarrow \quad \mathbf{f}_{\mathbf{y}} = \hat{\mathbf{M}}^{-1}\left(\hat{\mathbf{Q}}_{\mathbf{y}} - \hat{\mathbf{M}}_{\mathbf{y}}\mathbf{f}\right) \tag{3.90a}$$

$$\hat{\mathbf{M}}\frac{\partial \mathbf{f}}{\partial \boldsymbol{\rho}} + \hat{\mathbf{M}}_{\boldsymbol{\rho}}\mathbf{f} = \frac{\partial \hat{\mathbf{Q}}}{\partial \boldsymbol{\rho}} \quad \Rightarrow \quad \mathbf{f}_{\boldsymbol{\rho}} = \hat{\mathbf{M}}^{-1}\left(\hat{\mathbf{Q}}_{\boldsymbol{\rho}} - \hat{\mathbf{M}}_{\boldsymbol{\rho}}\mathbf{f}\right) \tag{3.90b}$$

The derivatives $\mathbf{f}_{\mathbf{y}}$ and $\mathbf{f}_{\boldsymbol{\rho}}$ can be calculated in block form as

$$\begin{aligned}\mathbf{f}_{\mathbf{y}} = &\begin{bmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \bar{\mathbf{M}}^{-1}\end{bmatrix}\left(\begin{bmatrix}\mathbf{0} & \mathbf{I}\\ -\bar{\mathbf{K}} & -\bar{\mathbf{C}}\end{bmatrix} - \begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \bar{\mathbf{M}}_{\mathbf{q}}\dot{\mathbf{v}} & \mathbf{0}\end{bmatrix}\right) = \\ &\begin{bmatrix}\mathbf{0} & \mathbf{I}\\ -\bar{\mathbf{M}}^{-1}\left(\bar{\mathbf{K}} + \bar{\mathbf{M}}_{\mathbf{q}}\dot{\mathbf{v}}\right) & -\bar{\mathbf{M}}^{-1}\bar{\mathbf{C}}\end{bmatrix}\end{aligned} \tag{3.91a}$$

$$\mathbf{f}_{\boldsymbol{\rho}} = \begin{bmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \bar{\mathbf{M}}\end{bmatrix}^{-1}\left(\begin{bmatrix}\mathbf{0}\\ \bar{\mathbf{Q}}_{\boldsymbol{\rho}}\end{bmatrix} - \begin{bmatrix}\mathbf{0}\\ \bar{\mathbf{M}}_{\boldsymbol{\rho}}\dot{\mathbf{v}}\end{bmatrix}\right) = \begin{bmatrix}\mathbf{0}\\ \bar{\mathbf{M}}^{-1}\left(\bar{\mathbf{Q}}_{\boldsymbol{\rho}} - \bar{\mathbf{M}}_{\boldsymbol{\rho}}\dot{\mathbf{v}}\right)\end{bmatrix} \tag{3.91b}$$

In Eqns. (3.91a) and (3.91b) the terms $\bar{\mathbf{K}}$, $\bar{\mathbf{C}}$, $\bar{\mathbf{Q}}_{\boldsymbol{\rho}}$, $\bar{\mathbf{M}}_{\mathbf{q}}\ddot{\mathbf{q}}$, and $\bar{\mathbf{M}}_{\boldsymbol{\rho}}\ddot{\mathbf{q}}$ are given by the following

expressions:

$$\bar{\mathbf{K}} = -\frac{\partial \bar{\mathbf{Q}}}{\partial \mathbf{q}} = \mathbf{K} + \mathbf{\Phi}_{\mathbf{qq}}^{\mathrm{T}} \alpha \left(\dot{\mathbf{\Phi}}_{\mathbf{q}} \dot{\mathbf{q}} + \dot{\mathbf{\Phi}}_t + 2\xi\omega\dot{\mathbf{\Phi}} + \omega^2 \mathbf{\Phi} \right) +$$
$$\mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}} \alpha \left(\left(\dot{\mathbf{\Phi}}_{\mathbf{q}} \dot{\mathbf{q}} \right)_{\mathbf{q}} + \left(\dot{\mathbf{\Phi}}_t \right)_{\mathbf{q}} + 2\xi\omega \left(\mathbf{\Phi}_{\mathbf{qq}} \dot{\mathbf{q}} + \mathbf{\Phi}_{t\mathbf{q}} \right) + \omega^2 \mathbf{\Phi}_{\mathbf{q}} \right) \tag{3.92a}$$

$$\bar{\mathbf{C}} = -\frac{\partial \bar{\mathbf{Q}}}{\partial \dot{\mathbf{q}}} = \mathbf{C} + \mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}} \alpha \left(\mathbf{\Phi}_{\mathbf{qq}} \dot{\mathbf{q}} + \dot{\mathbf{\Phi}}_{\mathbf{q}} + \mathbf{\Phi}_{tq} + 2\xi\omega\mathbf{\Phi}_{\mathbf{q}} \right) \tag{3.92b}$$

$$\bar{\mathbf{Q}}_{\boldsymbol{\rho}} = \frac{\partial \bar{\mathbf{Q}}}{\partial \boldsymbol{\rho}} = \mathbf{Q}_{\boldsymbol{\rho}} - \mathbf{\Phi}_{\mathbf{q}\boldsymbol{\rho}}^{\mathrm{T}} \alpha \left(\dot{\mathbf{\Phi}}_{\mathbf{q}} \dot{\mathbf{q}} + \dot{\mathbf{\Phi}}_t + 2\xi\omega\dot{\mathbf{\Phi}} + \omega^2 \mathbf{\Phi} \right) -$$
$$\mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}} \alpha \left(\left(\dot{\mathbf{\Phi}}_{\mathbf{q}} \dot{\mathbf{q}} \right)_{\boldsymbol{\rho}} + \dot{\mathbf{\Phi}}_{t\boldsymbol{\rho}} + 2\xi\omega\dot{\mathbf{\Phi}}_{\boldsymbol{\rho}} + \omega^2 \mathbf{\Phi}_{\boldsymbol{\rho}} \right) \tag{3.92c}$$

$$\bar{\mathbf{M}}_{\mathbf{q}} \ddot{\mathbf{q}} = \mathbf{M}_{\mathbf{q}} \ddot{\mathbf{q}} + \mathbf{\Phi}_{\mathbf{qq}}^{\mathrm{T}} \left(\alpha \mathbf{\Phi}_{\mathbf{q}} \ddot{\mathbf{q}} \right) + \mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}} \alpha \left(\mathbf{\Phi}_{\mathbf{qq}} \ddot{\mathbf{q}} \right) \tag{3.92d}$$

$$\bar{\mathbf{M}}_{\boldsymbol{\rho}} \ddot{\mathbf{q}} = \mathbf{M}_{\boldsymbol{\rho}} \ddot{\mathbf{q}} + \mathbf{\Phi}_{\mathbf{q}\boldsymbol{\rho}}^{\mathrm{T}} \left(\alpha \mathbf{\Phi}_{\mathbf{q}} \ddot{\mathbf{q}} \right) + \mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}} \alpha \left(\mathbf{\Phi}_{\mathbf{q}\boldsymbol{\rho}} \ddot{\mathbf{q}} \right) \tag{3.92e}$$

In Eqns. (3.92a) and (3.92b), $\mathbf{K} = -\mathbf{Q}_{\mathbf{q}}$ and $\mathbf{C} = -\mathbf{Q}_{\dot{\mathbf{q}}}$ respectively. For Eqns. (3.92d) and (3.92e), the following magnitudes are tensor-vector products that have to be calculated as explained in the nomenclature

$$\mathbf{M}_{\mathbf{q}} \ddot{\mathbf{q}} \equiv \mathbf{M}_{\mathbf{q}} \otimes \ddot{\mathbf{q}} \tag{3.93a}$$

$$\mathbf{M}_{\boldsymbol{\rho}} \ddot{\mathbf{q}} \equiv \mathbf{M}_{\boldsymbol{\rho}} \otimes \ddot{\mathbf{q}} \tag{3.93b}$$

$$\mathbf{\Phi}_{\mathbf{qq}}^{\mathrm{T}} \left(\alpha \mathbf{\Phi}_{\mathbf{q}} \ddot{\mathbf{q}} \right) \equiv \mathbf{\Phi}_{\mathbf{qq}}^{\mathrm{T}} \otimes \left(\alpha \mathbf{\Phi}_{\mathbf{q}} \ddot{\mathbf{q}} \right) \tag{3.93c}$$

$$\mathbf{\Phi}_{\mathbf{q}\boldsymbol{\rho}}^{\mathrm{T}} \left(\alpha \mathbf{\Phi}_{\mathbf{q}} \ddot{\mathbf{q}} \right) \equiv \mathbf{\Phi}_{\mathbf{q}\boldsymbol{\rho}}^{\mathrm{T}} \otimes \left(\alpha \mathbf{\Phi}_{\mathbf{q}} \ddot{\mathbf{q}} \right) \tag{3.93d}$$

$$\mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}} \alpha \left(\mathbf{\Phi}_{\mathbf{qq}} \ddot{\mathbf{q}} \right) \equiv \mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}} \alpha \left(\mathbf{\Phi}_{\mathbf{qq}} \otimes \ddot{\mathbf{q}} \right) \tag{3.93e}$$

$$\mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}} \alpha \left(\mathbf{\Phi}_{\mathbf{q}\boldsymbol{\rho}} \ddot{\mathbf{q}} \right) \equiv \mathbf{\Phi}_{\mathbf{q}}^{\mathrm{T}} \alpha \left(\mathbf{\Phi}_{\mathbf{q}\boldsymbol{\rho}} \otimes \ddot{\mathbf{q}} \right) \tag{3.93f}$$

To obtain expression (3.92a), the kinematic relation (2.18) was employed, and for expression (3.92b) the relations $\left(\dot{\mathbf{\Phi}}_{\mathbf{q}} \right)_{\dot{\mathbf{q}}} = \mathbf{\Phi}_{\mathbf{qq}}$, $\left(\dot{\mathbf{\Phi}}_t \right)_{\dot{\mathbf{q}}} = \mathbf{\Phi}_{t\mathbf{q}}$, were used. The last two relations can be checked by considering the following differentials

$$\delta\mathbf{\Phi}_{\mathbf{q}} = \mathbf{\Phi}_{\mathbf{qq}} \delta\mathbf{q} \Rightarrow \frac{d}{dt}\delta\mathbf{\Phi}_{\mathbf{q}} = \dot{\mathbf{\Phi}}_{\mathbf{qq}} \delta\mathbf{q} + \mathbf{\Phi}_{\mathbf{qq}} \delta\dot{\mathbf{q}} = \delta\dot{\mathbf{\Phi}}_{\mathbf{q}} = \dot{\mathbf{\Phi}}_{\mathbf{qq}} \delta\mathbf{q} + \dot{\mathbf{\Phi}}_{\mathbf{q}\dot{\mathbf{q}}} \delta\dot{\mathbf{q}} \Rightarrow \dot{\mathbf{\Phi}}_{\mathbf{q}\dot{\mathbf{q}}} = \mathbf{\Phi}_{\mathbf{qq}} \tag{3.94a}$$

$$\delta\mathbf{\Phi}_t = \mathbf{\Phi}_{t\mathbf{q}} \delta\mathbf{q} \Rightarrow \frac{d}{dt}\delta\mathbf{\Phi}_t = \dot{\mathbf{\Phi}}_{t\mathbf{q}} \delta\mathbf{q} + \mathbf{\Phi}_{t\mathbf{q}} \delta\dot{\mathbf{q}} = \delta\dot{\mathbf{\Phi}}_t = \dot{\mathbf{\Phi}}_{t\mathbf{q}} \delta\mathbf{q} + \dot{\mathbf{\Phi}}_{t\dot{\mathbf{q}}} \delta\dot{\mathbf{q}} \Rightarrow \dot{\mathbf{\Phi}}_{t\dot{\mathbf{q}}} = \mathbf{\Phi}_{t\mathbf{q}} \tag{3.94b}$$

Equations of motion written as a second-order ODE system

Considering the EOM (2.17a), the Lagrangian in this case has the following expression

$$\mathcal{L}(\boldsymbol{\rho}) = w(\mathbf{q}_F, \dot{\mathbf{q}}_F, \ddot{\mathbf{q}}_F, \boldsymbol{\rho}_F, \boldsymbol{\lambda}_F) + \int_{t_0}^{t_F} g(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}}, \boldsymbol{\lambda}^*, \boldsymbol{\rho})\,\mathrm{dt} - \int_{t_0}^{t_F} \boldsymbol{\mu}^{\mathrm{T}} \left(\bar{\mathbf{M}}(\mathbf{q}, \boldsymbol{\rho})\,\ddot{\mathbf{q}} - \bar{\mathbf{Q}}(t, \mathbf{q}, \dot{\mathbf{q}}, \boldsymbol{\rho})\right)\mathrm{dt} \tag{3.95}$$

Applying variational calculus

$$\begin{aligned}\delta\mathcal{L} = &\left[w_{\mathbf{q}}\delta\mathbf{q} + w_{\dot{\mathbf{q}}}\delta\dot{\mathbf{q}} + w_{\ddot{\mathbf{q}}}\delta\ddot{\mathbf{q}} + w_{\boldsymbol{\lambda}^*}\delta\boldsymbol{\lambda}^* + w_{\boldsymbol{\rho}}\delta\boldsymbol{\rho}\right]_{t_F} + \int_{t_0}^{t_F} \left(g_{\mathbf{q}} - \boldsymbol{\mu}^{\mathrm{T}}\left(\bar{\mathbf{M}}_{\mathbf{q}}\ddot{\mathbf{q}} + \bar{\mathbf{K}}\right)\right)\delta\mathbf{q}\mathrm{dt} \\ &+ \int_{t_0}^{t_F} \left(g_{\dot{\mathbf{q}}} - \boldsymbol{\mu}^{\mathrm{T}}\bar{\mathbf{C}}\right)\delta\dot{\mathbf{q}}\mathrm{dt} + \int_{t_0}^{t_F} \left(g_{\ddot{\mathbf{q}}} - \boldsymbol{\mu}^{\mathrm{T}}\bar{\mathbf{M}}\right)\delta\ddot{\mathbf{q}}\mathrm{dt} + \int_{t_0}^{t_F} g_{\boldsymbol{\lambda}^*}\delta\boldsymbol{\lambda}^*\mathrm{dt} + \int_{t_0}^{t_F} \left(g_{\boldsymbol{\rho}} - \boldsymbol{\mu}^{\mathrm{T}}\left(\bar{\mathbf{M}}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} - \bar{\mathbf{Q}}_{\boldsymbol{\rho}}\right)\right)\delta\boldsymbol{\rho}\mathrm{dt}\end{aligned} \tag{3.96}$$

The variation $\delta\boldsymbol{\lambda}^*$ can be removed by expressing it in terms of the variations $\delta\mathbf{q}$, $\delta\dot{\mathbf{q}}$ and $\delta\ddot{\mathbf{q}}$. From Eqn. (2.20)

$$\delta\boldsymbol{\lambda}^* = \boldsymbol{\alpha}\left(\delta\ddot{\boldsymbol{\Phi}} + 2\xi\omega\delta\dot{\boldsymbol{\Phi}} + \omega^2\delta\boldsymbol{\Phi}\right) \tag{3.97}$$

where

$$\begin{aligned}\delta\ddot{\boldsymbol{\Phi}} = &\boldsymbol{\Phi}_{\mathbf{q}}\delta\ddot{\mathbf{q}} + \left(\boldsymbol{\Phi}_{\mathbf{qq}}\dot{\mathbf{q}} + \dot{\boldsymbol{\Phi}}_{\mathbf{q}} + \boldsymbol{\Phi}_{t\mathbf{q}}\right)\delta\dot{\mathbf{q}} + \left(\boldsymbol{\Phi}_{\mathbf{qq}}\ddot{\mathbf{q}} + \left(\dot{\boldsymbol{\Phi}}_{\mathbf{q}}\right)_{\mathbf{q}}\dot{\mathbf{q}} + \left(\dot{\boldsymbol{\Phi}}_{t}\right)_{\mathbf{q}}\right)\delta\mathbf{q} \\ &+ \left(\boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}\ddot{\mathbf{q}} + \left(\dot{\boldsymbol{\Phi}}_{\mathbf{q}}\right)_{\boldsymbol{\rho}}\dot{\mathbf{q}} + \left(\dot{\boldsymbol{\Phi}}_{t}\right)_{\boldsymbol{\rho}}\right)\delta\boldsymbol{\rho}\end{aligned} \tag{3.98}$$

$$\delta\dot{\boldsymbol{\Phi}} = \boldsymbol{\Phi}_{\mathbf{q}}\delta\dot{\mathbf{q}} + \left(\boldsymbol{\Phi}_{\mathbf{qq}}\dot{\mathbf{q}} + \boldsymbol{\Phi}_{t\mathbf{q}}\right)\delta\mathbf{q} + \left(\boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}\dot{\mathbf{q}} + \boldsymbol{\Phi}_{t\boldsymbol{\rho}}\right)\delta\boldsymbol{\rho} \tag{3.99}$$

$$\delta\boldsymbol{\Phi} = \boldsymbol{\Phi}_{\mathbf{q}}\delta\mathbf{q} + \boldsymbol{\Phi}_{\boldsymbol{\rho}}\delta\boldsymbol{\rho} \tag{3.100}$$

Grouping together the terms associated to $\delta\ddot{\mathbf{q}}$, $\delta\dot{\mathbf{q}}$, $\delta\mathbf{q}$, $\delta\boldsymbol{\rho}$, Eqn. (3.97) becomes

$$\delta\boldsymbol{\lambda}^* = \boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\delta\ddot{\mathbf{q}} + \boldsymbol{\lambda}^*_{\dot{\mathbf{q}}}\delta\dot{\mathbf{q}} + \boldsymbol{\lambda}^*_{\mathbf{q}}\delta\mathbf{q} + \boldsymbol{\lambda}^*_{\boldsymbol{\rho}}\delta\boldsymbol{\rho} \tag{3.101}$$

Identifying the common terms in (3.97) and (3.101)

$$\boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}} = \boldsymbol{\alpha}\boldsymbol{\Phi}_{\mathbf{q}}, \tag{3.102a}$$

$$\boldsymbol{\lambda}^*_{\dot{\mathbf{q}}} = \alpha\left[\boldsymbol{\Phi}_{\mathbf{qq}}\dot{\mathbf{q}} + \dot{\boldsymbol{\Phi}}_{\mathbf{q}} + \boldsymbol{\Phi}_{t\mathbf{q}} + 2\xi\omega\boldsymbol{\Phi}_{\mathbf{q}}\right] \tag{3.102b}$$

$$\boldsymbol{\lambda}^*_{\mathbf{q}} = \alpha\left[\boldsymbol{\Phi}_{\mathbf{qq}}\ddot{\mathbf{q}} + \left(\dot{\boldsymbol{\Phi}}_{\mathbf{q}}\right)_{\mathbf{q}}\dot{\mathbf{q}} + \left(\dot{\boldsymbol{\Phi}}_{t}\right)_{\mathbf{q}} + 2\xi\omega\left(\boldsymbol{\Phi}_{\mathbf{qq}}\dot{\mathbf{q}} + \boldsymbol{\Phi}_{t\mathbf{q}}\right) + \omega^2\boldsymbol{\Phi}_{\mathbf{q}}\right] \tag{3.102c}$$

$$\boldsymbol{\lambda}^*_{\boldsymbol{\rho}} = \alpha\left[\boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}\ddot{\mathbf{q}} + \left(\dot{\boldsymbol{\Phi}}_{\mathbf{q}}\right)_{\boldsymbol{\rho}}\dot{\mathbf{q}} + \left(\dot{\boldsymbol{\Phi}}_{t}\right)_{\boldsymbol{\rho}} + 2\xi\omega\left(\boldsymbol{\Phi}_{\mathbf{q}\boldsymbol{\rho}}\dot{\mathbf{q}} + \boldsymbol{\Phi}_{t\boldsymbol{\rho}}\right) + \omega^2\boldsymbol{\Phi}_{\boldsymbol{\rho}}\right] \tag{3.102d}$$

The variation $\delta\ddot{\mathbf{q}}$ can be removed too, by expressing it in terms of the variations $\delta\mathbf{q}$, $\delta\dot{\mathbf{q}}$ and $\delta\ddot{\mathbf{q}}$. Making use of equations (2.16) and (3.101)

$$\mathbf{M}\delta\ddot{\mathbf{q}} + \mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}}\delta\mathbf{q} + \mathbf{M}_{\boldsymbol{\rho}}\ddot{\mathbf{q}}\delta\boldsymbol{\rho} + \boldsymbol{\Phi}^{\mathrm{T}}_{\mathbf{q}}\delta\boldsymbol{\lambda}^* + \boldsymbol{\Phi}^{\mathrm{T}}_{\mathbf{qq}}\boldsymbol{\lambda}^*\delta\mathbf{q} + \boldsymbol{\Phi}^{\mathrm{T}}_{\mathbf{q}\boldsymbol{\rho}}\boldsymbol{\lambda}^*\delta\boldsymbol{\rho} = \mathbf{Q}_{\mathbf{q}}\delta\mathbf{q} + \mathbf{Q}_{\dot{\mathbf{q}}}\delta\dot{\mathbf{q}} + \mathbf{Q}_{\boldsymbol{\rho}}\delta\boldsymbol{\rho} \tag{3.103}$$

$$\bar{\mathbf{M}}\delta\ddot{\mathbf{q}} = -\left(\mathbf{K} + \mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}^{\mathrm{T}}_{\mathbf{q}}\boldsymbol{\lambda}^*_{\mathbf{q}} + \boldsymbol{\Phi}^{\mathrm{T}}_{\mathbf{qq}}\boldsymbol{\lambda}^*\right)\delta\mathbf{q} - \left(\mathbf{C} + \boldsymbol{\Phi}^{\mathrm{T}}_{\mathbf{q}}\boldsymbol{\lambda}^*_{\dot{\mathbf{q}}}\right)\delta\dot{\mathbf{q}} + \left(\mathbf{Q}_{\boldsymbol{\rho}} - \mathbf{M}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} - \boldsymbol{\Phi}^{\mathrm{T}}_{\mathbf{q}}\boldsymbol{\lambda}^*_{\boldsymbol{\rho}} - \boldsymbol{\Phi}^{\mathrm{T}}_{\mathbf{q}\boldsymbol{\rho}}\boldsymbol{\lambda}^*\right)\delta\boldsymbol{\rho} \tag{3.104}$$

Then

$$\delta\ddot{\mathbf{q}} = \ddot{\mathbf{q}}_{\mathbf{q}}\delta\mathbf{q} + \ddot{\mathbf{q}}_{\dot{\mathbf{q}}}\delta\dot{\mathbf{q}} + \ddot{\mathbf{q}}_{\boldsymbol{\rho}}\delta\boldsymbol{\rho} \tag{3.105}$$

where

$$\ddot{\mathbf{q}}_{\mathbf{q}} = -\bar{\mathbf{M}}^{-1}\left(\mathbf{K} + \mathbf{M}_{\mathbf{q}}\ddot{\mathbf{q}} + \boldsymbol{\Phi}^{\mathrm{T}}_{\mathbf{q}}\boldsymbol{\lambda}^*_{\mathbf{q}} + \boldsymbol{\Phi}^{\mathrm{T}}_{\mathbf{qq}}\boldsymbol{\lambda}^*\right) \tag{3.106}$$

$$\ddot{\mathbf{q}}_{\dot{\mathbf{q}}} = -\bar{\mathbf{M}}^{-1}\left(\mathbf{C} + \boldsymbol{\Phi}^{\mathrm{T}}_{\mathbf{q}}\boldsymbol{\lambda}^*_{\dot{\mathbf{q}}}\right) \tag{3.107}$$

$$\ddot{\mathbf{q}}_{\boldsymbol{\rho}} = \bar{\mathbf{M}}^{-1}\left(\mathbf{Q}_{\boldsymbol{\rho}} - \mathbf{M}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} - \boldsymbol{\Phi}^{\mathrm{T}}_{\mathbf{q}}\boldsymbol{\lambda}^*_{\boldsymbol{\rho}} - \boldsymbol{\Phi}^{\mathrm{T}}_{\mathbf{q}\boldsymbol{\rho}}\boldsymbol{\lambda}^*\right) \tag{3.108}$$

Replacing equation (3.105) back in equation (3.101) yields

$$\delta\boldsymbol{\lambda}^* = \left(\boldsymbol{\lambda}^*_{\dot{\mathbf{q}}} + \boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}}\right)\delta\dot{\mathbf{q}} + \left(\boldsymbol{\lambda}^*_{\mathbf{q}} + \boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\mathbf{q}}\right)\delta\mathbf{q} + \left(\boldsymbol{\lambda}^*_{\boldsymbol{\rho}} + \boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}}\right)\delta\boldsymbol{\rho} \tag{3.109}$$

After replacing equations (3.105) and (3.109) in (3.96), the variations $\delta\ddot{\mathbf{q}}$ and $\delta\boldsymbol{\lambda}^*$ disappear

$$\begin{aligned}
\delta\mathcal{L} = &\left[\left(w_{\mathbf{q}} + w_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\mathbf{q}} + w_{\boldsymbol{\lambda}^*}\left(\boldsymbol{\lambda}^*_{\mathbf{q}} + \boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\mathbf{q}}\right)\right)\delta\mathbf{q} + \left(w_{\dot{\mathbf{q}}} + w_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}} + w_{\boldsymbol{\lambda}^*}\left(\boldsymbol{\lambda}^*_{\dot{\mathbf{q}}} + \boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}}\right)\right)\delta\dot{\mathbf{q}}\right.\\
&\left.+\left(w_{\boldsymbol{\rho}}+w_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}}+w_{\boldsymbol{\lambda}^*}\left(\boldsymbol{\lambda}^*_{\boldsymbol{\rho}}+\boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}}\right)\right)\delta\boldsymbol{\rho}\right]_F + \int_{t_0}^{t_F}\left(g_{\mathbf{q}}+g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\mathbf{q}}+g_{\boldsymbol{\lambda}^*}\left(\boldsymbol{\lambda}^*_{\mathbf{q}}+\boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\mathbf{q}}\right)-\boldsymbol{\mu}^{\mathrm{T}}\left(\bar{\mathbf{M}}_{\mathbf{q}}\ddot{\mathbf{q}}+\bar{\mathbf{K}}\right)\right)\delta\mathbf{q}\mathrm{d}t\\
&+\int_{t_0}^{t_F}\left(g_{\dot{\mathbf{q}}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}} + g_{\boldsymbol{\lambda}^*}\left(\boldsymbol{\lambda}^*_{\dot{\mathbf{q}}} + \boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}}\right) - \boldsymbol{\mu}^{\mathrm{T}}\bar{\mathbf{C}}\right)\delta\dot{\mathbf{q}}\mathrm{d}t\\
&-\int_{t_0}^{t_F}\boldsymbol{\mu}^{\mathrm{T}}\bar{\mathbf{M}}\delta\ddot{\mathbf{q}}\mathrm{d}t + \int_{t_0}^{t_F}\left(g_{\boldsymbol{\rho}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}} + g_{\boldsymbol{\lambda}^*}\left(\boldsymbol{\lambda}^*_{\boldsymbol{\rho}} + \boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}}\right) - \boldsymbol{\mu}^{\mathrm{T}}\left(\bar{\mathbf{M}}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} - \bar{\mathbf{Q}}_{\boldsymbol{\rho}}\right)\right)\delta\boldsymbol{\rho}\mathrm{d}t
\end{aligned} \tag{3.110}$$

Integrating by parts the integrals involving $\delta\dot{\mathbf{q}}, \delta\ddot{\mathbf{q}}$:

$$\begin{aligned}
\delta\mathcal{L} = &\left[\left(w_{\mathbf{q}} + w_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\mathbf{q}} + w_{\boldsymbol{\lambda}^*}\left(\boldsymbol{\lambda}^*_{\mathbf{q}} + \boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\mathbf{q}}\right)\right)\delta\mathbf{q} + \left(w_{\dot{\mathbf{q}}} + w_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}} + w_{\boldsymbol{\lambda}^*}\left(\boldsymbol{\lambda}^*_{\dot{\mathbf{q}}} + \boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}}\right)\right)\delta\dot{\mathbf{q}}\right.\\
&\left.+\left(w_{\boldsymbol{\rho}} + w_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}} + w_{\boldsymbol{\lambda}^*}\left(\boldsymbol{\lambda}^*_{\boldsymbol{\rho}} + \boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}}\right)\right)\delta\boldsymbol{\rho}\right]_F +\\
&\int_{t_0}^{t_F}\left(g_{\mathbf{q}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\mathbf{q}} + g_{\boldsymbol{\lambda}^*}\left(\boldsymbol{\lambda}^*_{\mathbf{q}} + \boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\mathbf{q}}\right) - \boldsymbol{\mu}^{\mathrm{T}}\left(\bar{\mathbf{M}}_{\mathbf{q}}\ddot{\mathbf{q}} + \bar{\mathbf{K}}\right)\right)\delta\mathbf{q}\mathrm{d}t\\
&+\left(g_{\dot{\mathbf{q}}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}} + g_{\boldsymbol{\lambda}^*}\left(\boldsymbol{\lambda}^*_{\dot{\mathbf{q}}} + \boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}}\right) - \boldsymbol{\mu}^{\mathrm{T}}\bar{\mathbf{C}}\right)\delta\mathbf{q}\Big|_{t0}^{tF}\\
&-\int_{t_0}^{t_F}\left(\frac{\mathrm{d}\left(g_{\dot{\mathbf{q}}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}} + g_{\boldsymbol{\lambda}^*}\left(\boldsymbol{\lambda}^*_{\dot{\mathbf{q}}} + \boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}}\right)\right)}{\mathrm{d}t} - \dot{\boldsymbol{\mu}}^{\mathrm{T}}\bar{\mathbf{C}} - \boldsymbol{\mu}^{\mathrm{T}}\dot{\bar{\mathbf{C}}}\right)\delta\mathbf{q}\mathrm{d}t\\
&-\left(\boldsymbol{\mu}^{\mathrm{T}}\bar{\mathbf{M}}\right)\delta\dot{\mathbf{q}}\Big|_{t0}^{tF} + \left(\dot{\boldsymbol{\mu}}^{\mathrm{T}}\bar{\mathbf{M}} + \boldsymbol{\mu}^{\mathrm{T}}\dot{\bar{\mathbf{M}}}\right)\delta\mathbf{q}\Big|_{t0}^{tF} - \int_{t_0}^{t_F}\left(\ddot{\boldsymbol{\mu}}^{\mathrm{T}}\bar{\mathbf{M}} + 2\dot{\boldsymbol{\mu}}^{\mathrm{T}}\dot{\bar{\mathbf{M}}} + \boldsymbol{\mu}^{\mathrm{T}}\ddot{\bar{\mathbf{M}}}\right)\delta\mathbf{q}\mathrm{d}t\\
&+\int_{t_0}^{t_F}\left(g_{\boldsymbol{\rho}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}} + g_{\boldsymbol{\lambda}^*}\left(\boldsymbol{\lambda}^*_{\boldsymbol{\rho}} + \boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}}\right) - \boldsymbol{\mu}^{\mathrm{T}}\left(\bar{\mathbf{M}}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} - \bar{\mathbf{Q}}_{\boldsymbol{\rho}}\right)\right)\delta\boldsymbol{\rho}\mathrm{d}t
\end{aligned} \tag{3.111}$$

Canceling all the integral terms that involve $\delta\mathbf{q}$ and $\delta\dot{\mathbf{q}}$ leads to the following adjoint ODE

$$\begin{aligned}
&\bar{\mathbf{M}}^{\mathrm{T}}\ddot{\boldsymbol{\mu}} + \left(2\dot{\bar{\mathbf{M}}} - \bar{\mathbf{C}}\right)^{\mathrm{T}}\dot{\boldsymbol{\mu}} + \left(\bar{\mathbf{M}}_{\mathbf{q}}\ddot{\mathbf{q}} + \bar{\mathbf{K}} - \dot{\bar{\mathbf{C}}} + \ddot{\bar{\mathbf{M}}}\right)^{\mathrm{T}}\boldsymbol{\mu} = \left(g_{\mathbf{q}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\mathbf{q}} + g_{\boldsymbol{\lambda}^*}\left(\boldsymbol{\lambda}^*_{\mathbf{q}} + \boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\mathbf{q}}\right)\right)^{\mathrm{T}}\\
&-\frac{\mathrm{d}\left(g_{\dot{\mathbf{q}}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}} + g_{\boldsymbol{\lambda}^*}\left(\boldsymbol{\lambda}^*_{\dot{\mathbf{q}}} + \boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}}\right)\right)^{\mathrm{T}}}{\mathrm{d}t}
\end{aligned} \tag{3.112a}$$

$$\left[\bar{\mathbf{M}}^{\mathrm{T}}\dot{\boldsymbol{\mu}} + \left(\dot{\bar{\mathbf{M}}} - \bar{\mathbf{C}}\right)^{\mathrm{T}}\boldsymbol{\mu}\right]_F =$$

$$-\left[w_{\mathbf{q}} + w_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\mathbf{q}} + w_{\boldsymbol{\lambda}^*}\left(\boldsymbol{\lambda}^*_{\mathbf{q}} + \boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\mathbf{q}}\right) + g_{\dot{\mathbf{q}}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}} + g_{\boldsymbol{\lambda}^*}\left(\boldsymbol{\lambda}^*_{\dot{\mathbf{q}}} + \boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}}\right)\right]_F^{\mathrm{T}} \tag{3.112b}$$

$$\left[\bar{\mathbf{M}}^{\mathrm{T}}\boldsymbol{\mu}\right]_F = \left[w_{\dot{\mathbf{q}}} + w_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}} + w_{\boldsymbol{\lambda}^*}\left(\boldsymbol{\lambda}^*_{\dot{\mathbf{q}}} + \boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}}\right)\right]_F^{\mathrm{T}} \tag{3.112c}$$

Finally, the gradient can be obtained from the remaining terms in the variational equation (3.111)

$$\begin{aligned}\nabla_{\boldsymbol{\rho}}\psi = &\left[w_{\boldsymbol{\rho}} + w_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}} + w_{\boldsymbol{\lambda}^*}\left(\boldsymbol{\lambda}^*_{\boldsymbol{\rho}} + \boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}}\right)\right]_F - \\ &\left[\mathbf{q}^{\mathrm{T}}_{\boldsymbol{\rho}}\left(g_{\dot{\mathbf{q}}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}} + g_{\boldsymbol{\lambda}^*}\left(\boldsymbol{\lambda}^*_{\dot{\mathbf{q}}} + \boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\dot{\mathbf{q}}}\right) + \dot{\boldsymbol{\mu}}^{\mathrm{T}}\bar{\mathbf{M}} + \boldsymbol{\mu}^{\mathrm{T}}\left(\dot{\bar{\mathbf{M}}} - \bar{\mathbf{C}}\right)\right)^{\mathrm{T}}\right]_{t0} \\ &+\left[\dot{\mathbf{q}}^{\mathrm{T}}_{\boldsymbol{\rho}}\left(\boldsymbol{\mu}^{\mathrm{T}}\bar{\mathbf{M}}\right)^{\mathrm{T}}\right]_{t0} + \int_{t_0}^{t_F}\left(g_{\boldsymbol{\rho}} + g_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}} + g_{\boldsymbol{\lambda}^*}\left(\boldsymbol{\lambda}^*_{\boldsymbol{\rho}} + \boldsymbol{\lambda}^*_{\ddot{\mathbf{q}}}\ddot{\mathbf{q}}_{\boldsymbol{\rho}}\right) - \boldsymbol{\mu}^{\mathrm{T}}\left(\bar{\mathbf{M}}_{\boldsymbol{\rho}}\ddot{\mathbf{q}} - \bar{\mathbf{Q}}_{\boldsymbol{\rho}}\right)\right)^{\mathrm{T}}\mathrm{dt} \qquad (3.113)\end{aligned}$$

Observe that more derivatives than those already obtained for the direct differentiation method are necessary in (3.112) and (3.113). The additional terms not obtained before are $\dot{\bar{\mathbf{M}}}$, $\ddot{\bar{\mathbf{M}}}$ and $\dot{\bar{\mathbf{C}}}$. From equation (2.17b)

$$\dot{\bar{\mathbf{M}}} = \dot{\mathbf{M}} + \dot{\boldsymbol{\Phi}}^{\mathrm{T}}_{\mathbf{q}}\boldsymbol{\alpha}\boldsymbol{\Phi}_{\mathbf{q}} + \boldsymbol{\Phi}^{\mathrm{T}}_{\mathbf{q}}\boldsymbol{\alpha}\dot{\boldsymbol{\Phi}}_{\mathbf{q}} \qquad (3.114)$$

$$\ddot{\bar{\mathbf{M}}} = \ddot{\mathbf{M}} + \ddot{\boldsymbol{\Phi}}^{\mathrm{T}}_{\mathbf{q}}\boldsymbol{\alpha}\boldsymbol{\Phi}_{\mathbf{q}} + 2\dot{\boldsymbol{\Phi}}^{\mathrm{T}}_{\mathbf{q}}\boldsymbol{\alpha}\dot{\boldsymbol{\Phi}}_{\mathbf{q}} + \boldsymbol{\Phi}^{\mathrm{T}}_{\mathbf{q}}\boldsymbol{\alpha}\ddot{\boldsymbol{\Phi}}_{\mathbf{q}} \qquad (3.115)$$

and from equation (3.92b)

$$\dot{\bar{\mathbf{C}}} = \dot{\mathbf{C}} + \dot{\boldsymbol{\Phi}}^{\mathrm{T}}_{\mathbf{q}}\alpha\left(\boldsymbol{\Phi}_{\mathbf{qq}}\dot{\mathbf{q}} + \dot{\boldsymbol{\Phi}}_{\mathbf{q}} + \boldsymbol{\Phi}_{t\mathbf{q}} + 2\xi\omega\boldsymbol{\Phi}_{\mathbf{q}}\right) + \boldsymbol{\Phi}^{\mathrm{T}}_{\mathbf{q}}\alpha\left(\dot{\boldsymbol{\Phi}}_{\mathbf{qq}}\dot{\mathbf{q}} + \boldsymbol{\Phi}_{\mathbf{qq}}\ddot{\mathbf{q}} + \ddot{\boldsymbol{\Phi}}_{\mathbf{q}} + \dot{\boldsymbol{\Phi}}_{t\mathbf{q}} + 2\xi\omega\dot{\boldsymbol{\Phi}}_{\mathbf{q}}\right) \qquad (3.116)$$

3.2.4 Maggi's formulation

The adjoint variable method of this section is most useful when the number of parameters p is large. Approaches are developed herein for each of the following three formulations of the EOM:

1. The equations of motion are written as a first-order explicit ODE system.
2. The equations of motion are written as a first-order implicit ODE system, and
3. The equations of motion are written as a second-order implicit ODE system.

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

The system (2.25) can be transformed into a first-order implicit one, by introducing a new set of variables $\dot{\mathbf{z}} = \mathbf{v}$, and defining the new vector $\mathbf{y} = \begin{bmatrix} \mathbf{z}^{\mathrm{T}} & \mathbf{v}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$

$$\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{M}} \end{bmatrix} \begin{bmatrix} \dot{\mathbf{z}} \\ \dot{\mathbf{v}} \end{bmatrix} = \begin{bmatrix} \mathbf{v} \\ \bar{\mathbf{Q}} \end{bmatrix} \tag{3.117a}$$

$$\hat{\mathbf{M}}(\mathbf{y}, \boldsymbol{\rho})\, \dot{\mathbf{y}} = \hat{\mathbf{Q}}(t, \mathbf{y}, \boldsymbol{\rho}) \tag{3.117b}$$

Taking the inverse of the leading matrix in (3.117b), the system can be expressed as a first-order explicit one,

$$\dot{\mathbf{y}} = \quad \hat{\mathbf{M}}^{-1}(\mathbf{y}, \boldsymbol{\rho})\, \hat{\mathbf{Q}}(t, \mathbf{y}, \boldsymbol{\rho}) \quad = \mathbf{f}(t, \mathbf{y}, \boldsymbol{\rho}) \tag{3.118}$$

The cost function becomes

$$\psi = \int_{t_0}^{t_F} g(\mathbf{y}, \boldsymbol{\rho})\, \mathrm{dt} \tag{3.119}$$

As proposed in [117] we consider the following Lagrangian, given by the cost function constrained by the EOM,

$$L(\boldsymbol{\rho}) = \int_{t_0}^{t_F} g(\mathbf{y}, \boldsymbol{\rho})\, \mathrm{dt} - \int_{t_0}^{t_F} \boldsymbol{\mu}^{\mathrm{T}} \left(\dot{\mathbf{y}} - \mathbf{f}(t, \mathbf{y}, \boldsymbol{\rho})\right) \mathrm{dt} \tag{3.120}$$

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

$$\delta L = \int_{t_0}^{t_F} \left(\frac{\partial g}{\partial \mathbf{y}}\delta\mathbf{y} + \frac{\partial g}{\partial \boldsymbol{\rho}}\delta\boldsymbol{\rho}\right) \mathrm{dt} - \int_{t_0}^{t_F} \delta\boldsymbol{\mu}^{\mathrm{T}} \left(\dot{\mathbf{y}} - \mathbf{f}(t, \mathbf{y}, \boldsymbol{\rho})\right) \mathrm{dt} - \int_{t_0}^{t_F} \boldsymbol{\mu}^{\mathrm{T}} \left(\delta\dot{\mathbf{y}} - \frac{\partial \mathbf{f}}{\partial \mathbf{y}}\delta\mathbf{y} - \frac{\partial \mathbf{f}}{\partial \boldsymbol{\rho}}\delta\boldsymbol{\rho}\right) \mathrm{dt} \tag{3.121}$$

The central term involves the dot product of $\delta\boldsymbol{\mu}$ with the EOM; if they are fulfilled at each time step, this term vanishes. For the last term, integration by parts is applied

$$\int_{t_0}^{t_F} \boldsymbol{\mu}^{\mathrm{T}} \delta\dot{\mathbf{y}} \mathrm{dt} = \boldsymbol{\mu}^{\mathrm{T}} \delta\mathbf{y}\Big|_{t_0}^{t_F} - \int_{t_0}^{t_F} \dot{\boldsymbol{\mu}}^{\mathrm{T}} \delta\mathbf{y} \mathrm{dt} \tag{3.122}$$

Therefore

$$\delta L = \int_{t_0}^{t_F} \left(\frac{\partial g}{\partial \mathbf{y}} + \boldsymbol{\mu}^{\mathrm{T}} \frac{\partial \mathbf{f}}{\partial \mathbf{y}} + \dot{\boldsymbol{\mu}}^{\mathrm{T}} \right) \delta \mathbf{y} \mathrm{d}t + \int_{t_0}^{t_F} \left(\frac{\partial g}{\partial \boldsymbol{\rho}} + \boldsymbol{\mu}^{\mathrm{T}} \frac{\partial \mathbf{f}}{\partial \boldsymbol{\rho}} \right) \delta \boldsymbol{\rho} \mathrm{d}t - \boldsymbol{\mu}^{\mathrm{T}} \left(t_F \right) \delta \mathbf{y} \left(t_F \right) + \boldsymbol{\mu}^{\mathrm{T}} \left(t_0 \right) \delta \mathbf{y} \left(t_0 \right) \tag{3.123}$$

In equation (3.123) $\delta \mathbf{y} \left(t_0 \right)$ in the last term is known, and the $\delta \mathbf{y} \left(t_F \right)$ term can be cancelled by choosing $\boldsymbol{\mu} \left(t_F \right) = \mathbf{0}$. Moreover, to avoid calculating $\delta \mathbf{y}$, the first integral is canceled by choosing $\boldsymbol{\mu}$ to be the solution of following adjoint ODE system:

$$\dot{\boldsymbol{\mu}} = -\mathbf{f}_{\mathbf{y}}^{\mathrm{T}} \boldsymbol{\mu} - g_{\mathbf{y}}^{\mathrm{T}} \tag{3.124a}$$

$$\boldsymbol{\mu} \left(t_F \right) = \mathbf{0} \tag{3.124b}$$

From equation (3.123) the gradient of the cost function (3.119) with respect to parameters can be obtained as

$$\nabla_{\boldsymbol{\rho}} \psi = \mathbf{y}_{\boldsymbol{\rho}}^{\mathrm{T}} \left(t_0 \right) \boldsymbol{\mu} \left(t_0 \right) + \int_{t_0}^{t_F} \left(\mathbf{f}_{\boldsymbol{\rho}}^{\mathrm{T}} \boldsymbol{\mu} + g_{\boldsymbol{\rho}}^{\mathrm{T}} \right) \mathrm{d}t \tag{3.125}$$

In the previous result the identity $\delta \psi = \delta L$ was used, which holds if the EOM (3.118) are satisfied, as can be derived from (3.120).

In (3.125) and (3.124a) the derivatives of function g are known, since the objective function has a known expression. The derivatives of $\mathbf{f}$ are obtained from (3.117b):

$$\begin{aligned} &\hat{\mathbf{M}} \frac{\partial \mathbf{f}}{\partial \mathbf{y}} + \hat{\mathbf{M}}_{\mathbf{y}} \mathbf{f} = \frac{\partial \hat{\mathbf{Q}}}{\partial \mathbf{y}} \Rightarrow \mathbf{f}_{\mathbf{y}} = \hat{\mathbf{M}}^{-1} \left(\hat{\mathbf{Q}}_{\mathbf{y}} - \hat{\mathbf{M}}_{\mathbf{y}} \mathbf{f} \right) = \\ &\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{M}} \end{bmatrix}^{-1} \left(\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\bar{\mathbf{K}} & -\bar{\mathbf{C}} \end{bmatrix} - \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{M}}_{\mathbf{z}} \dot{\mathbf{v}} & \mathbf{0} \end{bmatrix} \right) = \\ &\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\bar{\mathbf{M}}^{-1} \left(\bar{\mathbf{K}} + \bar{\mathbf{M}}_{\mathbf{z}} \dot{\mathbf{v}} \right) & -\bar{\mathbf{M}}^{-1} \bar{\mathbf{C}} \end{bmatrix} \end{aligned} \tag{3.126a}$$

$$\begin{aligned} &\hat{\mathbf{M}} \frac{\partial \mathbf{f}}{\partial \boldsymbol{\rho}} + \hat{\mathbf{M}}_{\boldsymbol{\rho}} \mathbf{f} = \frac{\partial \hat{\mathbf{Q}}}{\partial \boldsymbol{\rho}} \Rightarrow \mathbf{f}_{\boldsymbol{\rho}} = \hat{\mathbf{M}}^{-1} \left(\hat{\mathbf{Q}}_{\boldsymbol{\rho}} - \hat{\mathbf{M}}_{\boldsymbol{\rho}} \mathbf{f} \right) = \\ &\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{M}} \end{bmatrix}^{-1} \left(\begin{bmatrix} \mathbf{0} \\ \bar{\mathbf{Q}}_{\boldsymbol{\rho}} \end{bmatrix} - \begin{bmatrix} \mathbf{0} \\ \bar{\mathbf{M}}_{\boldsymbol{\rho}} \dot{\mathbf{v}} \end{bmatrix} \right) = \begin{bmatrix} \mathbf{0} \\ \bar{\mathbf{M}}^{-1} \left(\bar{\mathbf{Q}}_{\boldsymbol{\rho}} - \bar{\mathbf{M}}_{\boldsymbol{\rho}} \dot{\mathbf{v}} \right) \end{bmatrix} \end{aligned} \tag{3.126b}$$

Taking into account the first equation of (3.117a), $\mathbf{v} = \dot{\mathbf{z}}$, the terms $\bar{\mathbf{K}}$, $\bar{\mathbf{C}}$ and $\bar{\mathbf{M}}_{\mathbf{z}}\dot{\mathbf{v}}$ in (3.126a) are given by Eq. (3.28a), (3.28b), and (3.29a) respectively. Similarly in (3.126b), the terms $\bar{\mathbf{Q}}_{\boldsymbol{\rho}}$ and $\bar{\mathbf{M}}_{\boldsymbol{\rho}}\dot{\mathbf{v}}$ are given by (3.28c) and (3.29b) respectively.

Equations of motion written as a first-order implicit ODE system

Another form of the adjoint ODE system (3.124a) and the gradient of the cost function (3.125) can be obtained using the EOM (3.117b) as constraints in the Lagrangian

$$L(\boldsymbol{\rho}) = \int_{t_0}^{t_F} g(\mathbf{y}, \boldsymbol{\rho})\,\mathrm{d}t - \int_{t_0}^{t_F} \boldsymbol{\mu}^{\mathrm{T}} \left(\hat{\mathbf{M}}(\mathbf{y}, \boldsymbol{\rho})\,\dot{\mathbf{y}} - \hat{\mathbf{Q}}(t, \mathbf{y}, \boldsymbol{\rho}) \right) \mathrm{d}t \tag{3.127}$$

Applying variational calculus

$$\begin{aligned} \delta L = & \int_{t_0}^{t_F} \left(\frac{\partial g}{\partial \mathbf{y}}\delta\mathbf{y} + \frac{\partial g}{\partial \boldsymbol{\rho}}\delta\boldsymbol{\rho} \right) \mathrm{d}t - \int_{t_0}^{t_F} \delta\boldsymbol{\mu}^{\mathrm{T}} \left(\hat{\mathbf{M}}(\mathbf{y}, \boldsymbol{\rho})\,\dot{\mathbf{y}} - \hat{\mathbf{Q}}(t, \mathbf{y}, \boldsymbol{\rho}) \right) \mathrm{d}t \\ & - \int_{t_0}^{t_F} \boldsymbol{\mu}^{\mathrm{T}} \left(\hat{\mathbf{M}}\delta\dot{\mathbf{y}} + \hat{\mathbf{M}}_{\mathbf{y}}\dot{\mathbf{y}}\delta\mathbf{y} + \hat{\mathbf{M}}_{\boldsymbol{\rho}}\dot{\mathbf{y}}\delta\boldsymbol{\rho} - \frac{\partial \hat{\mathbf{Q}}}{\partial \mathbf{y}}\delta\mathbf{y} - \frac{\partial \hat{\mathbf{Q}}}{\partial \boldsymbol{\rho}}\delta\boldsymbol{\rho} \right) \mathrm{d}t \end{aligned} \tag{3.128}$$

Again, the central term vanishes if the EOM are fulfilled at each instant. For the last term, integration by parts is applied:

$$\int_{t_0}^{t_F} \boldsymbol{\mu}^{\mathrm{T}}\hat{\mathbf{M}}\delta\dot{\mathbf{y}}\mathrm{d}t = \boldsymbol{\mu}^{\mathrm{T}}\hat{\mathbf{M}}\delta\mathbf{y}\Big|_{t_0}^{t_F} - \int_{t_0}^{t_F} \left(\dot{\boldsymbol{\mu}}^{\mathrm{T}}\hat{\mathbf{M}} + \boldsymbol{\mu}^{\mathrm{T}}\dot{\hat{\mathbf{M}}} \right) \delta\mathbf{y}\mathrm{d}t \tag{3.129}$$

Therefore,

$$\begin{aligned} \delta L = & \int_{t_0}^{t_F} \left(\frac{\partial g}{\partial \mathbf{y}} - \boldsymbol{\mu}^{\mathrm{T}} \left(\hat{\mathbf{M}}_{\mathbf{y}}\dot{\mathbf{y}} - \frac{\partial \hat{\mathbf{Q}}}{\partial \mathbf{y}} - \dot{\hat{\mathbf{M}}} \right) + \dot{\boldsymbol{\mu}}^{\mathrm{T}}\hat{\mathbf{M}} \right) \delta\mathbf{y}\mathrm{d}t \\ & + \int_{t_0}^{t_F} \left(\frac{\partial g}{\partial \boldsymbol{\rho}} - \boldsymbol{\mu}^{\mathrm{T}} \left(\hat{\mathbf{M}}_{\boldsymbol{\rho}}\dot{\mathbf{y}} - \frac{\partial \hat{\mathbf{Q}}}{\partial \boldsymbol{\rho}} \right) \right) \delta\boldsymbol{\rho}\mathrm{d}t - \boldsymbol{\mu}^{\mathrm{T}}\hat{\mathbf{M}}\delta\mathbf{y}\Big|_{t_0}^{t_F} \end{aligned} \tag{3.130}$$

The first integral and the last term at the final time are canceled if $\boldsymbol{\mu}$ is the solution of the

following adjoint ODE:

$$\hat{\mathbf{M}}\,\dot{\boldsymbol{\mu}} = \left(\hat{\mathbf{M}}_{\mathbf{y}}\dot{\mathbf{y}} - \hat{\mathbf{Q}}_{\mathbf{y}} - \dot{\hat{\mathbf{M}}}\right)^{\mathrm{T}} \boldsymbol{\mu} - g_{\mathbf{y}}^{\mathrm{T}} \tag{3.131a}$$

$$\boldsymbol{\mu}\,(t_F) = \mathbf{0} \tag{3.131b}$$

where the symmetry of $\hat{\mathbf{M}}$ was assumed, $g_{\mathbf{y}}$ is known and

$$\hat{\mathbf{M}}_{\mathbf{y}}\dot{\mathbf{y}} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{M}}_{\mathbf{z}}\dot{\mathbf{v}} & \mathbf{0} \end{bmatrix} \tag{3.132}$$

$$\tfrac{\partial \hat{\mathbf{Q}}}{\partial \mathbf{y}} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\bar{\mathbf{K}} & -\bar{\mathbf{C}} \end{bmatrix} \tag{3.133}$$

Observe that the previous derivatives (3.132) and (3.133) have the same terms already explained in the derivation of (3.126a).

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

$$\dot{\hat{\mathbf{M}}} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \dot{\bar{\mathbf{M}}} \end{bmatrix} \tag{3.134a}$$

$$\dot{\bar{\mathbf{M}}} = \sum_i \frac{\partial \bar{\mathbf{M}}}{\partial z_i}\dot{z}_i = \sum_i \left(\mathbf{R}_{z_i}^{\mathrm{T}}\mathbf{M}\mathbf{R} + \mathbf{R}^{\mathrm{T}}\mathbf{M}_{z_i}\mathbf{R} + \mathbf{R}^{\mathrm{T}}\mathbf{M}\mathbf{R}_{z_i}\right)\dot{z}_i \tag{3.134b}$$

where the terms of $\mathbf{R}_{z_i}$ are given by Eq. (3.37). From Eq. (3.130), after removing the terms made zero in Eq. (3.131), the gradient of the cost function with respect to parameters can be obtained as

$$\nabla_{\boldsymbol{\rho}}\psi = \mathbf{y}_{\boldsymbol{\rho}}^{\mathrm{T}}\,(t_0)\,\hat{\mathbf{M}}\,(t_0)\,\boldsymbol{\mu}\,(t_0) + \int_{t_0}^{t_F}\left(\left(\hat{\mathbf{Q}}_{\boldsymbol{\rho}} - \hat{\mathbf{M}}_{\boldsymbol{\rho}}\dot{\mathbf{y}}\right)^{\mathrm{T}}\boldsymbol{\mu} + g_{\boldsymbol{\rho}}^{\mathrm{T}}\right)\mathrm{dt} \tag{3.135}$$

where the symmetry of $\hat{\mathbf{M}}$ was used,

$$\hat{\mathbf{Q}}_{\boldsymbol{\rho}} = \begin{bmatrix} \mathbf{0} \\ \bar{\mathbf{Q}}_{\boldsymbol{\rho}} \end{bmatrix} \tag{3.136}$$

$$\hat{\mathbf{M}}_{\boldsymbol{\rho}}\dot{\mathbf{y}} = \begin{bmatrix} \mathbf{0} \\ \bar{\mathbf{M}}_{\boldsymbol{\rho}}\dot{\mathbf{v}} \end{bmatrix} \tag{3.137}$$

and $\bar{\mathbf{Q}}_{\boldsymbol{\rho}}$ and $\bar{\mathbf{M}}_{\boldsymbol{\rho}}\dot{\mathbf{v}}$ are given by (3.28c) and (3.29b) respectively.

Equations of motion written as a second-order implicit ODE system

The first-order adjoint ODE system (3.131) of size $2\,n$ can be replaced by a second-order adjoint system of size n. The EOM (2.25) are used in the construction of the Lagrangian

$$L\left(\boldsymbol{\rho}\right) = \int_{t_0}^{t_F} g\left(\mathbf{z}, \dot{\mathbf{z}}, \boldsymbol{\rho}\right) \mathrm{d}t - \int_{t_0}^{t_F} \boldsymbol{\mu}^{\mathrm{T}} \left(\bar{\mathbf{M}}\left(\mathbf{z}, \boldsymbol{\rho}\right) \ddot{\mathbf{z}} - \bar{\mathbf{Q}}\left(t, \mathbf{z}, \dot{\mathbf{z}}, \boldsymbol{\rho}\right)\right) \mathrm{d}t \tag{3.138}$$

Applying variational calculus

$$\begin{aligned} \delta L = &\int_{t_0}^{t_F} \left(\frac{\partial g}{\partial \mathbf{z}}\delta\mathbf{z} + \frac{\partial g}{\partial \dot{\mathbf{z}}}\delta\dot{\mathbf{z}} + \frac{\partial g}{\partial \boldsymbol{\rho}}\delta\boldsymbol{\rho}\right) \mathrm{d}t - \int_{t_0}^{t_F} \delta\boldsymbol{\mu}^{\mathrm{T}} \left(\bar{\mathbf{M}}\left(\mathbf{z}, \boldsymbol{\rho}\right) \ddot{\mathbf{z}} - \bar{\mathbf{Q}}\left(t, \mathbf{z}, \dot{\mathbf{z}}, \boldsymbol{\rho}\right)\right) \mathrm{d}t - \\ &\int_{t_0}^{t_F} \boldsymbol{\mu}^{\mathrm{T}} \left(\bar{\mathbf{M}}\delta\ddot{\mathbf{z}} + \bar{\mathbf{M}}_{\mathbf{z}}\ddot{\mathbf{z}}\delta\mathbf{z} + \bar{\mathbf{M}}_{\boldsymbol{\rho}}\ddot{\mathbf{z}}\delta\boldsymbol{\rho} - \frac{\partial \bar{\mathbf{Q}}}{\partial \mathbf{z}}\delta\mathbf{z} - \frac{\partial \bar{\mathbf{Q}}}{\partial \dot{\mathbf{z}}}\delta\dot{\mathbf{z}} - \frac{\partial \bar{\mathbf{Q}}}{\partial \boldsymbol{\rho}}\delta\boldsymbol{\rho}\right) \mathrm{d}t \end{aligned} \tag{3.139}$$

The central term vanishes if the EOM are fulfilled at each instant. Application of integration by parts gives

$$\int_{t_0}^{t_F} \boldsymbol{\mu}^{\mathrm{T}}\bar{\mathbf{M}}\delta\ddot{\mathbf{z}}\mathrm{d}t = \boldsymbol{\mu}^{\mathrm{T}}\bar{\mathbf{M}}\delta\dot{\mathbf{z}}\Big|_{t_0}^{t_F} - \int_{t_0}^{t_F} \left(\dot{\boldsymbol{\mu}}^{\mathrm{T}}\bar{\mathbf{M}} + \boldsymbol{\mu}^{\mathrm{T}}\dot{\bar{\mathbf{M}}}\right) \delta\dot{\mathbf{z}}\mathrm{d}t \tag{3.140}$$

$$\int_{t_0}^{t_F} \frac{\partial g}{\partial \dot{\mathbf{z}}}\delta\dot{\mathbf{z}}\mathrm{d}t = \frac{\partial g}{\partial \dot{\mathbf{z}}}\delta\mathbf{z}\Big|_{t_0}^{t_F} - \int_{t_0}^{t_F} \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial g}{\partial \dot{\mathbf{z}}}\delta\mathbf{z}\mathrm{d}t \tag{3.141}$$

$$\int_{t_0}^{t_F} \boldsymbol{\mu}^{\mathrm{T}}\frac{\partial \bar{\mathbf{Q}}}{\partial \dot{\mathbf{z}}}\delta\dot{\mathbf{z}}\mathrm{d}t = -\boldsymbol{\mu}^{\mathrm{T}}\bar{\mathbf{C}}\delta\mathbf{z}\Big|_{t_0}^{t_F} + \int_{t_0}^{t_F} \left(\boldsymbol{\mu}^{\mathrm{T}}\dot{\bar{\mathbf{C}}} + \dot{\boldsymbol{\mu}}^{\mathrm{T}}\bar{\mathbf{C}}\right) \delta\mathbf{z}\mathrm{d}t \tag{3.142}$$

where $\bar{\mathbf{C}}$ is described in (3.28b) and $\dot{\bar{\mathbf{C}}}$ later in this section. Applying integration by parts again to the integral on the right hand side of (3.140) leads to

$$\int_{t_0}^{t_F}\left(\dot{\boldsymbol{\mu}}^{\mathrm{T}}\bar{\mathbf{M}}+\boldsymbol{\mu}^{\mathrm{T}}\dot{\bar{\mathbf{M}}}\right)\delta\dot{\mathbf{z}}\mathrm{d}t=\left(\dot{\boldsymbol{\mu}}^{\mathrm{T}}\bar{\mathbf{M}}+\boldsymbol{\mu}^{\mathrm{T}}\dot{\bar{\mathbf{M}}}\right)\delta\mathbf{z}\Big|_{t_0}^{t_F}-\int_{t_0}^{t_F}\left(\ddot{\boldsymbol{\mu}}^{\mathrm{T}}\bar{\mathbf{M}}+2\dot{\boldsymbol{\mu}}^{\mathrm{T}}\dot{\bar{\mathbf{M}}}+\boldsymbol{\mu}^{\mathrm{T}}\ddot{\bar{\mathbf{M}}}\right)\delta\mathbf{z}\mathrm{d}t \tag{3.143}$$

From Eq. (3.143) and Eq. (3.140)

$$\int_{t_0}^{t_F}\boldsymbol{\mu}^{\mathrm{T}}\bar{\mathbf{M}}\delta\ddot{\mathbf{z}}\mathrm{d}t=\boldsymbol{\mu}^{\mathrm{T}}\bar{\mathbf{M}}\delta\dot{\mathbf{z}}\Big|_{t_0}^{t_F}-\left(\dot{\boldsymbol{\mu}}^{\mathrm{T}}\bar{\mathbf{M}}+\boldsymbol{\mu}^{\mathrm{T}}\dot{\bar{\mathbf{M}}}\right)\delta\mathbf{z}\Big|_{t_0}^{t_F}+\int_{t_0}^{t_F}\left(\ddot{\boldsymbol{\mu}}^{\mathrm{T}}\bar{\mathbf{M}}+2\dot{\boldsymbol{\mu}}^{\mathrm{T}}\dot{\bar{\mathbf{M}}}+\boldsymbol{\mu}^{\mathrm{T}}\ddot{\bar{\mathbf{M}}}\right)\delta\mathbf{z}\mathrm{d}t \tag{3.144}$$

Therefore,

$$\begin{aligned}\delta L=&\int_{t_0}^{t_F}\left(\frac{\partial g}{\partial\mathbf{z}}-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial g}{\partial\dot{\mathbf{z}}}-\boldsymbol{\mu}^{\mathrm{T}}\left(\bar{\mathbf{M}}_{\mathbf{z}}\ddot{\mathbf{z}}+\bar{\mathbf{K}}-\dot{\bar{\mathbf{C}}}+\ddot{\bar{\mathbf{M}}}\right)-\dot{\boldsymbol{\mu}}^{\mathrm{T}}\left(2\dot{\bar{\mathbf{M}}}-\bar{\mathbf{C}}\right)-\ddot{\boldsymbol{\mu}}^{\mathrm{T}}\bar{\mathbf{M}}\right)\delta\mathbf{z}\mathrm{d}t+\\&\left(\frac{\partial g}{\partial\dot{\mathbf{z}}}+\dot{\boldsymbol{\mu}}^{\mathrm{T}}\bar{\mathbf{M}}+\boldsymbol{\mu}^{\mathrm{T}}\left(\dot{\bar{\mathbf{M}}}-\bar{\mathbf{C}}\right)\right)\delta\mathbf{z}\Big|_{t_0}^{t_F}-\boldsymbol{\mu}^{\mathrm{T}}\bar{\mathbf{M}}\delta\dot{\mathbf{z}}\Big|_{t_0}^{t_F}+\int_{t_0}^{t_F}\left(\frac{\partial g}{\partial\boldsymbol{\rho}}-\boldsymbol{\mu}^{\mathrm{T}}\left(\bar{\mathbf{M}}_{\boldsymbol{\rho}}\ddot{\mathbf{z}}-\frac{\partial\bar{\mathbf{Q}}}{\partial\boldsymbol{\rho}}\right)\right)\delta\boldsymbol{\rho}\mathrm{d}t\end{aligned} \tag{3.145}$$

The first integral and the terms evaluated in $t=t_F$ are canceled in (3.145) if $\boldsymbol{\mu}$ is the solution of the following second-order adjoint ODE system

$$\bar{\mathbf{M}}\,\ddot{\boldsymbol{\mu}}+\left(2\dot{\bar{\mathbf{M}}}-\bar{\mathbf{C}}\right)^{\mathrm{T}}\dot{\boldsymbol{\mu}}+\left(\bar{\mathbf{M}}_{\mathbf{z}}\ddot{\mathbf{z}}+\bar{\mathbf{K}}-\dot{\bar{\mathbf{C}}}+\ddot{\bar{\mathbf{M}}}\right)^{\mathrm{T}}\boldsymbol{\mu}=g_{\mathbf{z}}^{\mathrm{T}}-\frac{\mathrm{d}}{\mathrm{d}t}g_{\dot{\mathbf{z}}}^{\mathrm{T}} \tag{3.146a}$$

$$\boldsymbol{\mu}\left(t_F\right)=\mathbf{0} \tag{3.146b}$$

$$\bar{\mathbf{M}}\left(t_F\right)\dot{\boldsymbol{\mu}}\left(t_F\right)=-g_{\dot{\mathbf{z}}}^{\mathrm{T}}\left(t_F\right) \tag{3.146c}$$

where the symmetry of the matrix $\bar{\mathbf{M}}$ was used. From Eq. (3.145) the gradient of the cost function with respect to parameters is

$$\nabla_{\boldsymbol{\rho}}\psi=-\mathbf{z}_{\boldsymbol{\rho}}^{\mathrm{T}}\left(t_0\right)\left(g_{\dot{\mathbf{z}}}+\dot{\boldsymbol{\mu}}^{\mathrm{T}}\bar{\mathbf{M}}+\boldsymbol{\mu}^{\mathrm{T}}\left(\dot{\bar{\mathbf{M}}}-\bar{\mathbf{C}}\right)\right)^{\mathrm{T}}\Big|_{t_0}+\dot{\mathbf{z}}_{\boldsymbol{\rho}}^{\mathrm{T}}\left(t_0\right)\bar{\mathbf{M}}\left(t_0\right)\boldsymbol{\mu}\left(t_0\right)+\int_{t_0}^{t_F}\left(\left(\bar{\mathbf{Q}}_{\boldsymbol{\rho}}-\bar{\mathbf{M}}_{\boldsymbol{\rho}}\ddot{\mathbf{z}}\right)^{\mathrm{T}}\boldsymbol{\mu}+g_{\boldsymbol{\rho}}^{\mathrm{T}}\right)\mathrm{d}t \tag{3.147}$$

Most of the terms appearing in Eq. (3.146a) and Eq. (3.147) were obtained in sections 2.5 and 3.1.4. Some additional terms are also needed. From Eq. (3.28b) and Eqn. (3.30b)

$$\dot{\mathbf{C}} = \dot{\mathbf{R}}^{\mathrm{T}} \left(\mathbf{C} + \mathbf{M}\mathbf{S}\mathbf{c}_{\dot{\mathbf{q}}} \right) \mathbf{R} + \mathbf{R}^{\mathrm{T}} \left(\mathbf{C} + \mathbf{M}\mathbf{S}\mathbf{c}_{\dot{\mathbf{q}}} \right) \dot{\mathbf{R}} + \mathbf{R}^{\mathrm{T}} \left(\dot{\mathbf{C}} + \dot{\mathbf{M}}\mathbf{S}\mathbf{c}_{\dot{\mathbf{q}}} + \mathbf{M}\dot{\mathbf{S}}\mathbf{c}_{\dot{\mathbf{q}}} + \mathbf{M}\mathbf{S}\dot{\mathbf{c}}_{\dot{\mathbf{q}}} \right) \mathbf{R} \tag{3.148}$$

From Eq. (3.35)

$$\dot{\mathbf{c}}_{\dot{\mathbf{q}}} = \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{c}_{\dot{\mathbf{q}}} = -\mathbf{\Phi}_{\mathbf{q}\mathbf{q}} \ddot{\mathbf{q}} - \frac{\mathrm{d}}{\mathrm{d}t} \left(\mathbf{\Phi}_{\mathbf{q}\mathbf{q}} \right) \dot{\mathbf{q}} - \ddot{\mathbf{\Phi}}_{\mathbf{q}} - \dot{\mathbf{\Phi}}_{t\mathbf{q}} \tag{3.149}$$

The terms $\dot{\mathbf{M}}$ and $\ddot{\mathbf{M}}$ can be calculated either from (3.134b), or by taking derivatives of (2.25b)

$$\dot{\mathbf{M}} = \dot{\mathbf{R}}^{\mathrm{T}} \mathbf{M} \mathbf{R} + \mathbf{R}^{\mathrm{T}} \dot{\mathbf{M}} \mathbf{R} + \mathbf{R}^{\mathrm{T}} \mathbf{M} \dot{\mathbf{R}} \tag{3.150a}$$

$$\ddot{\mathbf{M}} = \ddot{\mathbf{R}}^{\mathrm{T}} \mathbf{M} \mathbf{R} + \mathbf{R}^{\mathrm{T}} \mathbf{M} \ddot{\mathbf{R}} + 2\dot{\mathbf{R}}^{\mathrm{T}} \mathbf{M} \dot{\mathbf{R}} + \mathbf{R}^{\mathrm{T}} \ddot{\mathbf{M}} \mathbf{R} + 2 \left(\dot{\mathbf{R}}^{\mathrm{T}} \dot{\mathbf{M}} \mathbf{R} + \mathbf{R}^{\mathrm{T}} \dot{\mathbf{M}} \dot{\mathbf{R}} \right) \tag{3.150b}$$

where the symmetry of the mass matrix $\mathbf{M}$ was assumed. Taking derivatives in equation (2.27) leads to:

$$\begin{bmatrix} \mathbf{\Phi}_{\mathbf{q}} \\ \mathbf{B} \end{bmatrix} \begin{bmatrix} \dot{\mathbf{S}} & \dot{\mathbf{R}} \end{bmatrix} = - \begin{bmatrix} \dot{\mathbf{\Phi}}_{\mathbf{q}} \mathbf{S} & \dot{\mathbf{\Phi}}_{\mathbf{q}} \mathbf{R} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \Rightarrow \begin{cases} \dot{\mathbf{R}} = -\mathbf{S}\dot{\mathbf{\Phi}}_{\mathbf{q}}\mathbf{R} \\ \dot{\mathbf{S}} = -\mathbf{S}\dot{\mathbf{\Phi}}_{\mathbf{q}}\mathbf{S} \end{cases} \tag{3.151}$$

Differentiating once more in Eq. (3.151) yields

$$\ddot{\mathbf{R}} = -\mathbf{S} \left(\ddot{\mathbf{\Phi}}_{\mathbf{q}} \mathbf{R} + 2\dot{\mathbf{\Phi}}_{\mathbf{q}} \dot{\mathbf{R}} \right) \tag{3.152}$$

3.3 Validation of the computed sensitivities

The validation of the computed sensitivities is crucial because small errors in individual terms can result in completely wrong sensitivities and even if wrong sensitivities can usually solve optimization problems, they are still wrong. The strategies proposed and employed here to validate sensitivities are the following:

1. Compare the results of the direct differentiation method and the adjoint variable method: they should be equal within the truncation error.

2. Compare the results of different formulations of the EOM: Index-3, Index-1, the penalty and

Maggi's sensitivities were compared.

3. Compute the sensitivities using a third party code: FATODE [118] was used to double-check the results presented.

4. Use real finite difference method to approximate whole sensitivities or individual derivatives: it can be a very inaccurate or even completely useless strategy. Small truncation errors (small δ) cause important loss-of-significance errors.The first-order approximation of the derivatives with real perturbations reads

$$\frac{\mathrm{d}\psi}{\mathrm{d}\rho_k} = \frac{\psi\left(\boldsymbol{\rho} + \delta \mathbf{e}_k\right) - \psi\left(\boldsymbol{\rho}\right)}{\delta} \tag{3.153}$$

   The truncation error in this case is $\mathcal{O}(\delta)$ cf. (3.153) and the loss of significance errors are order $\mathcal{O}(\delta^{-1})$, where δ is the perturbation. This fact can make these derivatives highly inaccurate.

5. Use complex finite difference method to approximate whole sensitivities or individual derivatives: it is a much more reliable approach than the previous one, but more complex to implement. Since there is not subtraction, there are not loss-of-significance errors in the imaginary part. The first-order approximation of derivatives with complex perturbations is the following

$$\frac{\mathrm{d}\psi}{\mathrm{d}\rho_k} = \frac{\Im\left(\psi\left(\boldsymbol{\rho} + i\delta \mathbf{e}_k\right)\right)}{\delta} \tag{3.154}$$

   where i is the imaginary unit and $\Im$ is the imaginary part of a complex number. The approach is considerably more accurate than the previous one, because there are no subtractions in the imaginary parts and therefore the perturbations can be chosen arbitrarily small without loss-of-significance errors appearing in the calculation of the approximation. The practical difficulty to apply complex finite difference method is that not all codes can be changed easily to accommodate complex arithmetic. Special attention should be paid to the third party functions (if any) involved in the code (*transpose* functions, *norm* functions, numerical integrator chosen, etc).

Chapter 4

Numerical optimization

There are numerous optimization methods in the literature. Since sensitivity analysis is performed prior to optimization in this study, most gradient-based optimization methods with constraints should be good choices. L-BFGS-B, which is a popular gradient-based optimization package based on quasi-Newton method, is used in this work to perform optimization.

This study only uses L-BFGS-B to perform optimization and no new optimization algorithm is developed here, thus the motivation of this chapter is to briefly review the theory of gradient-based optimization methods and the algorithm of L-BFGS-B. The global optimization methods and multi-objective methods will not be reviewed since this study doesn't include any results from them. For more theoretical details on these methods, the reader is referred to Venkataraman 2009 [119], Bazaraa 2013 [120], Fletcher 2013 [121], and Wright 1999 [122].

4.1 Optimality conditions for unconstrained case

The local optimization method relies on the first-order or high-order derivatives of the objective function, which guarantees the convergence to a local minimum close to the initial guess. In this section, the unconstrained optimization problem will be presented, which can be written as follows:

$$\begin{aligned} \text{Minimize} \quad & \Psi(\mathbf{q}) \\ & \mathbf{q} \in \mathbb{R}^n \end{aligned} \tag{4.1}$$

In order to perform unconstrained optimization, the first-order and second-order conditions will be reviewed first since they are the most relevant optimality conditions.

The first-order necessary condition is given as follows:

Let $\Psi : D \subset \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable on D, then $\nabla\Psi(\mathbf{q}^*) = \mathbf{0}$ if $\mathbf{q}^* \in D$ is a local minimum. This condition shows us the slope at $\mathbf{q}^*$ is $\mathbf{0}$ if $\mathbf{q}^*$ is a local minimum.

The second-order necessary condition is given as follows:

Let $\Psi : D \subset \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable on D, then $\nabla\Psi(\mathbf{q}^*) = \mathbf{0}$ and $\nabla^2\Psi(\mathbf{q}^*)$ is positive semidefinite if $\mathbf{q}^* \in D$ is a local minimum.

A zero slope doesn't guarantee a local minimum, it could be also a local maximum. Thus, the second-order sufficient condition is given as follows:

Let $\Psi : D \subset \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable on D, if $\nabla\Psi(\mathbf{q}^*) = \mathbf{0}$ and $\nabla^2\Psi(\mathbf{q}^*)$ is positive definite, then $\mathbf{q}^* \in D$ is a local minimum.

Actually, solving the optimality conditions for unconstrained case is equivalent to finding solutions of algebraic nonlinear system.

4.2 Optimality conditions for constrained case

The general from of a constrained optimization problem can be written as follows:

$$\begin{aligned} \text{Minimize} \quad & \Psi(\mathbf{q}) \\ \text{Subject to} \quad & \mathbf{C}_{\text{eq}}(\mathbf{q}) = \mathbf{0} \\ & \mathbf{C}_{\text{ineq}}(\mathbf{q}) \leq \mathbf{0} \\ & \mathbf{q} \in \mathbb{R}^n \end{aligned} \tag{4.2}$$

where $\mathbf{C}_{\text{eq}}(\mathbf{q})$ are the equality constraints and $\mathbf{C}_{\text{ineq}}(\mathbf{q})$ are the inequality constraints. In order to solve this constrained optimization problem, Karush in 1939, Kuhn and Tucker in 1951 presented

the flowing well-know conditions, which is also called KKT condition:

$$\nabla\Psi(\mathbf{q}^*) + \nabla\mathbf{C}_{\text{eq}}(\mathbf{q}^*)^{\text{T}}\boldsymbol{\mu} + \nabla\mathbf{C}_{\text{ineq}}(\mathbf{q}^*)^{\text{T}}\boldsymbol{\nu} = \mathbf{0} \tag{4.3}$$

$$\boldsymbol{\mu}^{\text{T}}\nabla\mathbf{C}_{\text{eq}}(\mathbf{q}^*) = 0 \tag{4.4}$$

$$\boldsymbol{\mu} \geq \mathbf{0} \tag{4.5}$$

where $\boldsymbol{\mu} \in \mathbb{R}^{\mathbf{C}_{\text{ineq}}}$ and $\boldsymbol{\nu} \in \mathbb{R}^{\mathbf{C}_{\text{eq}}(\mathbf{q}}$.

KKT conditions are not necessary or sufficient for a local minimum. But if some conditions hold, KKT conditions could become necessary or sufficient for a local minimum. First let c_i and g_i be the i_{th} element of $\mathbf{C}_{\text{ineq}}$ and $\mathbf{C}_{\text{eq}}$. $I(\bar{x}) = \{i : c_i(\bar{x}) = 0\}$, we have some necessary and sufficient conditions as follows:

- Necessary condition
  - Let $\mathbf{q}^*$ be a local minimum. If $g_i\ \forall i$ are affine, and $c_j\ \forall j \in I(q^*)$ are (pseudo)concave, then $\mathbf{q}^*$ is a KKT point.
  - Let $\mathbf{q}^*$ be a local minimum. If $\nabla g_i(\mathbf{q}^*)\ \forall i$ and $\nabla c_j(\mathbf{q}^*)\ \forall j \in I(q^*)$ are linear independent, then $\mathbf{q}^*$ is a KKT point.
  - Let $\mathbf{q}^*$ be a local minimum. If $g_i(\mathbf{q}^*) \forall i$ are locally affine, $c_j(\mathbf{q}^*)\ \forall j \in I(q^*)$ are locally convex, and $\nabla g_i(\mathbf{q}^*)\ \forall i$ are linear independent, and $\exists\hat{\mathbf{q}}$ such that $c_j(\hat{\mathbf{q}}) < 0\ \forall j \in I(q^*)$, then $\mathbf{q}^*$ is a KKT point.
- Sufficient condition
  - Let $\mathbf{q}^*$ be a KKT point. If Ψ is locally (pseudo)convex, $g_i\ \forall i$ are affine, and $c_j\ \forall j \in I(q^*)$ are locally (pseudo)convex, then $\mathbf{q}^*$ is a local minimum.

4.3 The L-BFGS-B algorithm

Introduction

Since L-BFGS-B is the only optimization package used in this study, its algorithm will be briefly introduced here. For more details, the reader is referred to [123].

The problem is described as follows:

$$\begin{aligned} &\text{Minimize } \Psi(\mathbf{q}) \\ &\text{Subject to } \mathrm{l_i} \leq \mathrm{q_i} \leq \mathrm{u_i} \ \forall \mathrm{i} \end{aligned} \tag{4.6}$$

where Ψ is the cost function, $\mathrm{l_i}$ and $\mathrm{u_i}$ represent the lower bound and the upper bounds for each parameter $\mathrm{q_i}$. This algorithm requires the first derivative of the cost function that is provided by MBSVT while the second derivative of the cost function is not required. Thus this algorithm can be applied when the Hessian matrix is not available. L-BFGS-B uses a limited memory BFGS update to replace the Hessian matrix.

Outline of the algorithm

Assume $\mathbf{q_k}$ is the current state where $\mathrm{k} = 0$ for the initial guess, in order to minimize the cost function $\Psi(\mathbf{q}_k)$, a quadratic model is formed as follows:

$$\Psi(\mathbf{q}) = \Psi(\mathbf{q}_k) + \mathbf{G}_k^T(\mathbf{q} - \mathbf{q}_k) + \frac{1}{2}(\mathbf{q} - \mathbf{q}_k)^T \mathbf{B}_k(\mathbf{q} - \mathbf{q}_k) \tag{4.7}$$

where $\Psi(\mathbf{q}_k)$ is the value of cost function, $\mathbf{G}_k$ is the value of gradient, and $\mathbf{B}_k$ is the limited memory BFGS matrix that is used to replace the Hessian matrix. Then with the bounds from 4.6, $\Psi(\mathbf{q})$ can be approximately minimized. To do this, the piece-wise linear path is described first as follows:

$$\mathbf{q}(t) = P(\mathbf{q}_k - t\mathbf{G}_k, \mathbf{l}, \mathbf{u}) \tag{4.8}$$

where

$$P_i(\mathbf{q},\mathbf{l},\mathbf{u}) = \begin{cases} l_i & if \quad q_i < l_i \\ q_i & if \quad l_i < q_i < u_i \\ u_i & if \quad q_i > v_i \end{cases} \tag{4.9}$$

With equation 4.8 and 4.9, Cauchy point $\mathbf{q}^c$ is computed. Cauchy point is defined as the first local minimum of $\Psi(\mathbf{q})$. The variables whose value at the $\mathbf{q}^c$ at lower or upper bound are fixed, which comprise the active set $A(\mathbf{q}^c)$. After obtaining the Cauchy point and active set, the following quadratic problem is used to calculate an approximate solution $\bar{\mathbf{q}}_{k+1}$ as follows:

$$\begin{aligned} &\min \{\Psi(\mathbf{q}) \,:\, q_i = q_i^c \;\forall\, i \in A(\mathbf{q}^c)\} \\ &\text{subject to } l_i < q_i < u_i \;\forall\, i \notin A(\mathbf{q}^c) \end{aligned} \tag{4.10}$$

After obtaining an approximate solution $\bar{\mathbf{q}}_{k+1}$, the new state $\mathbf{q}_{k+1}$ can be obtained by satisfying the following two conditions:

$$\begin{aligned} &\Psi(\mathbf{q}_{k+1}) = \Psi(\mathbf{q}_k) + \alpha\lambda_k \mathbf{G}_k^T \mathbf{D}_k \\ &\left|\mathbf{G}_{k+1}^T \mathbf{D}_k\right| \leq \beta \left|\mathbf{G}_k^T \mathbf{D}_k\right| \end{aligned} \tag{4.11}$$

where $\mathbf{D}_k = \bar{\mathbf{q}}_{k+1} - \mathbf{q}_k$ determines the direction and step length of line search, λ_k is the step length of the line search, α and β are 10^{-4} and 0.9 respectively. After obtaining the new state $\mathbf{q}_{k+1}$, If it converges, then the process is ended, otherwise a new iteration starts.

For more details about the computation of the Cauchy point $\mathbf{q}^c$, the computation of the limited memory BFGS matrix $\mathbf{B}_k$, and the minimization of the quadratic problem, the reader is referred to [123].

Chapter 5

Numerical experiments

In this chapter, all of the new sensitivity methods developed from chapter 3 are applied to two case studies, five-bar mechanism and a full vehicle. The resulting sensitivities are then applied to perform dynamical optimization. For five-bar mechanism, since it's a unconstrained optimization problem, the first-order necessary condition from section 4.1 is used to solve the optimization problem. For the road vehicle problem, the ride response and handling response are optimized by using L-BFGS-B. In addition, the dynamic analysis of a passive dynamic robot with knees is presented and several different point and surface contact models are employed to model the contact dynamics for the passive dynamic robot.

5.1 Five-bar mechanism optimization

The first case study chosen to test the formulations proposed in this study is the five-bar mechanism with 2 *DOF* shown in Fig.5.1.

The five bars are constrained by five revolute joints located in points A, 1, 2, 3 and B. The masses of the bars are $m_1 = 1\ kg$, $m_2 = 1.5\ kg$, $m_3 = 1.5\ kg$, $m_4 = 1\ kg$ and the polar moments of inertia are calculated under the assumption of a uniform distribution of mass. The mechanism is subjected to the action of gravity and two elastic forces coming from the springs. The stiffness coefficients of the springs are $k_1 = k_2 = 100\ N/m$ and their natural lengths are initially chosen $L_{01} = \sqrt{2^2 + 1^2}\ m$ and $L_{02} = \sqrt{2^2 + 0.5^2}\ m$, coincident with the initial configuration shown in Fig.5.1.

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

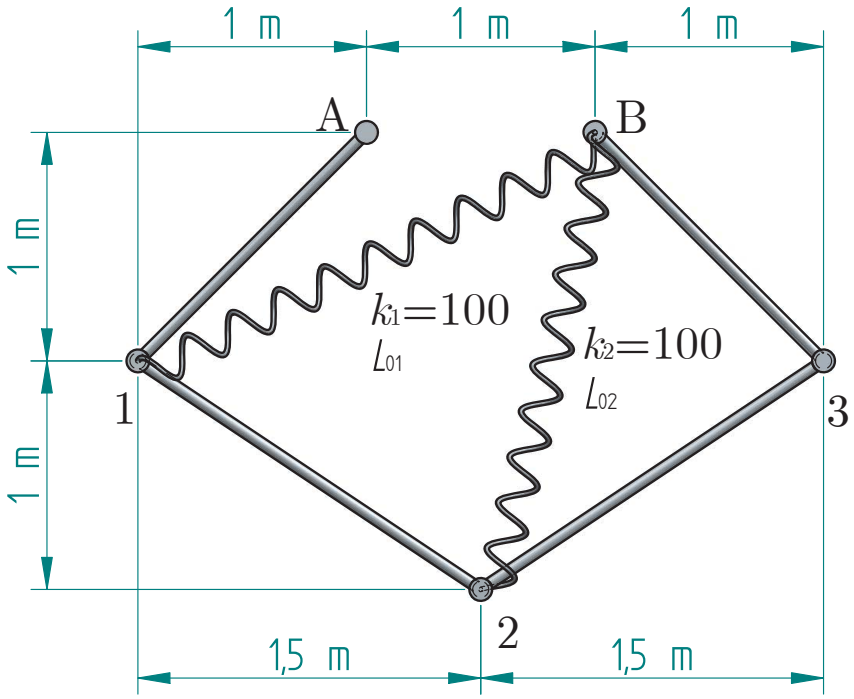


Figure 5.1: The five-bar mechanism

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

$$\psi = \int_{t_0}^{t_F} (\mathbf{r}_2 - \mathbf{r}_{20})^{\mathrm{T}} (\mathbf{r}_2 - \mathbf{r}_{20}) \, \mathrm{dt} \tag{5.1}$$

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

The condition to obtain the minimum is the following:

$$\nabla_{\boldsymbol{\rho}} \psi = \mathbf{0} \tag{5.2}$$

The gradient (5.2) was obtained by the following approaches with the EOM proposed in the paper:

1. Direct sensitivity:Index-3 DAE formulation.
2. Direct sensitivity:Index-1 DAE formulation.

3. Direct sensitivity:The penalty formulation.

4. Direct sensitivity:Maggi's formulation.

5. Adjoint sensitivity.:Index-3 DAE formulation.

6. Adjoint sensitivity.:Index-1 DAE formulation.

7. Adjoint sensitivity.:The penalty formulation.

8. Adjoint sensitivity.:Maggi's formulation.

9. Numerical sensitivity with real perturbations.

10. Numerical sensitivity with complex perturbations.

The response of the system is shown in Fig.5.2 for a 5 seconds simulation.

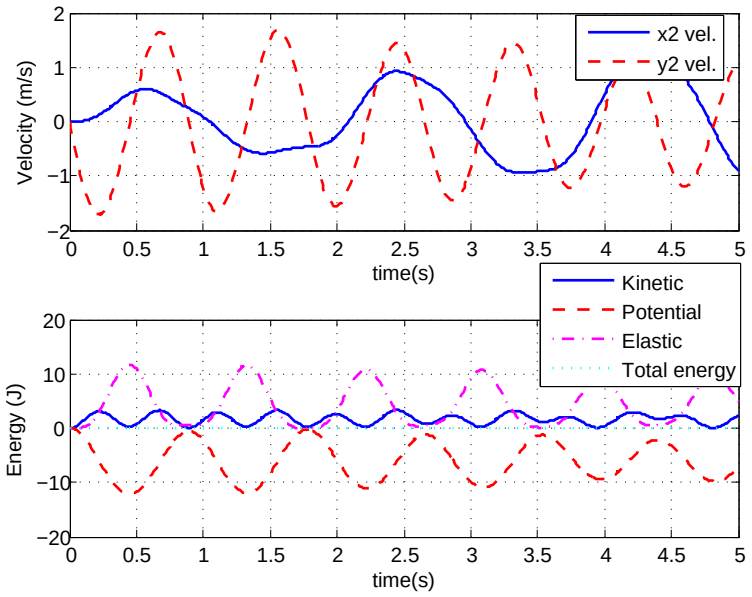


Figure 5.2: Mechanism response: top) velocity of point 2; bottom) energy of the system

The upper plot represents the horizontal and vertical velocities of the point 2 while the lower one represents the energy taking as reference for the potential energy the initial configuration of the system.

The results for the sensitivities with the mentioned methods are presented in table 5.1.

Table 5.1: Results for the five-bar mechanism.

Approach	Parameters	$\mathrm{d}\psi/\mathrm{d}L_{01}$	$\mathrm{d}\psi/\mathrm{d}L_{02}$
1: Direct index-3	$h = 10^{-2}s$	-4.2381	3.2170
2: Direct index-1	$h = 10^{-2}s$	-4.2383	3.2169
3: Direct penalty	$h = 10^{-2}s$	-4.2305	3.2154
4: Direct maggi's	$h = 10^{-2}s$	-4.2300	3.2112
5: Adjoint index-3	$h = 10^{-2}s$	-4.2287	3.2090
6: Adjoint index-1	$h = 10^{-2}s$	-4.2294	3.2094
7: Adjoint penalty	$h = 10^{-2}s$	-4.2293	3.2137
8: Adjoint maggi's	$h = 10^{-2}s$	-4.2294	3.2093
9: FATODE	$Tol = 10^{-3}$	-4.2257	3.2077
10: Num. diff. real with penalty	$\delta = 10^{-7}m$	-9.7390	-4.0344
11: Num. diff. complex	$\delta/i = 10^{-7}m$	-4.2288	3.2116

As can be seen in table5.1, all the approaches, except the numerical sensitivities with real perturbations, offer similar results, which guarantees that the schemes proposed are correct. The numerical sensitivities with real perturbations are not reliable if accurate results for the sensitivities are important for the application to tackle. Given the simplicity of the system proposed, definitive conclusions in terms of efficiency cannot be stated.

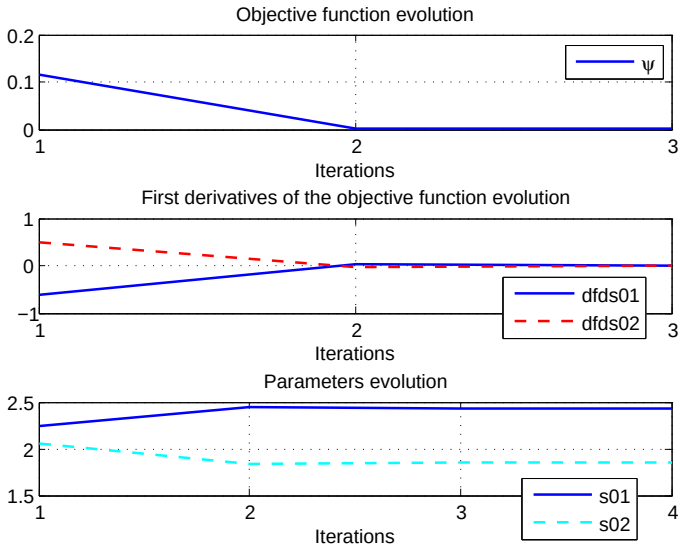


Figure 5.3: Objective function, gradient and parameters evaluation.

The computed sensitivities can be employed for the optimization proposed. All the methods perform similar to solve the optimization problem. In this case the simulation time was reduced to $1s$ and the results for the objective function, derivatives and parameters are presented in figure5.3 for the *Adjoint penalty* approach. The plots for the direct differentiation method coincide with the ones presented and they are not presented for clarity.

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

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

5.2 Road vehicle optimization

5.2.1 Vehicle model

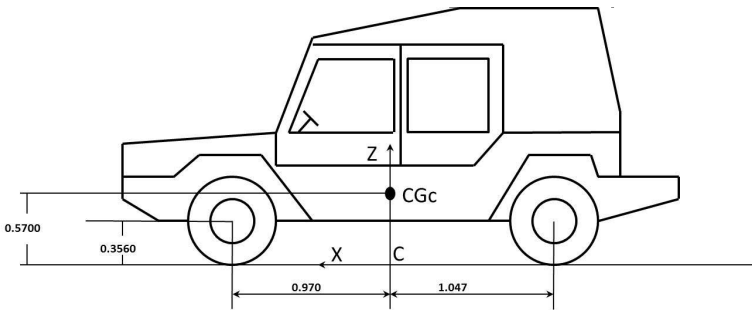


Figure 5.4: The Bombardier Iltis vehicle (Adapted from [1])

The Iltis vehicle was proposed as a benchmark problem by the European automobile industry to check multibody dynamics codes. The vehicle model is extensively described in [1], therefore only a summary of the most important parameters of the model is given here. The vehicle is represented in Fig. 5.4 and a diagram of the model is given in Fig. 5.5 showing that the model 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.

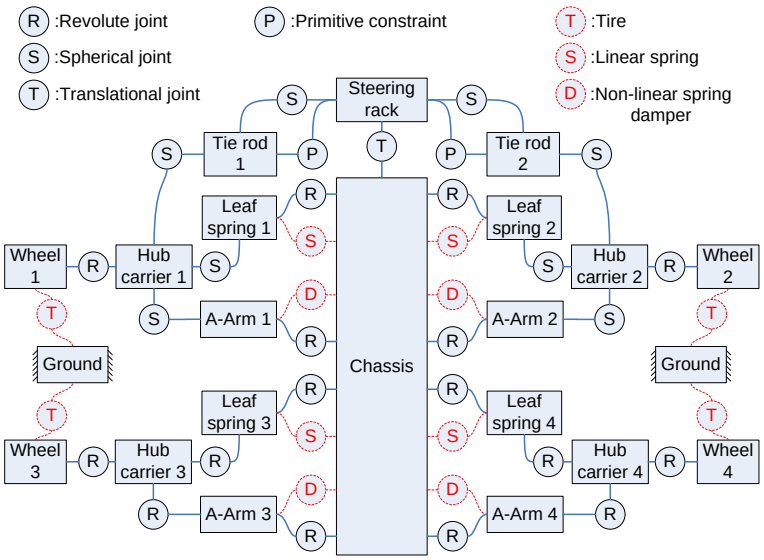


Figure 5.5: Multibody model diagram

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 5.2. As indicated in [1], 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 *CG* reference frames and all products of inertia are considered to be zero. Centers of mass locations are given in table 5.3, expressed in the reference frame *C*, indicated in Fig. 5.4.

The topology of the suspensions is explained in diagram Fig. 5.5 and the geometry of the left front suspensions is shown in Fig. 5.6. 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 5.2: Mass and principal moments of inertia (Adapted from [1])

Body	Mass [kg]	I_{xx} [kg m^2]	I_{yy} [kg m^2]	I_{zz} [kg m^2]
Chassis	1260	130	1620	1670
Wheel/hub/brake assembly	57.35	1.2402	1.908	1.2402
A-arm	6.0	0.052099	0.023235	0.068864

Table 5.3: Positions of centers of mass (origin C. Fig. 5.4) (Adapted from [1])

Body	Coordinates of CG [m]		
	x	y	z
Chassis	0	0	0.57
Right front wheel with hub and brake assembly	0.97	-0.615	0.356
Left rear wheel with hub and brake assembly	-1.047	0.615	0.356
Right front A-arm	0.97	-0.4155	0.2655
Left rear A-arm	-1.047	0.4155	0.2655

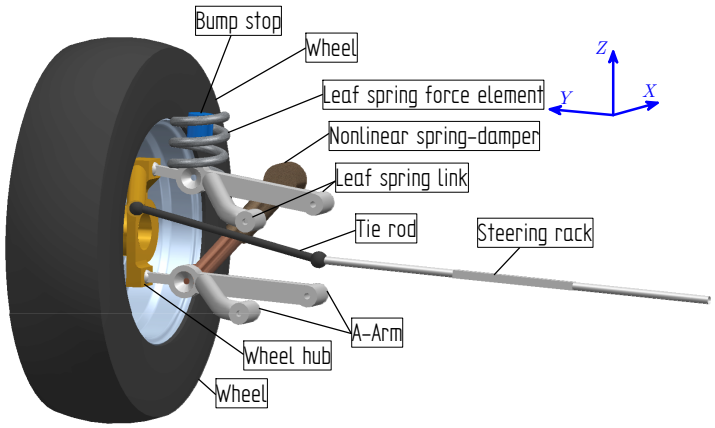


Figure 5.6: Left front suspension system

The key point positions for the left front suspension are given in table 5.4. The corresponding points for the left and rear suspensions can be easily obtained since all the suspensions are identical,

except for the fact that the tie rods are not present in the rear ones, since there is not steering in the back.

Table 5.4: Positions of joints (left front suspension, origin C. Fig. 5.4) (Adapted from [1])

Point description	x [m]	y [m]	z [m]
Wheel center	0.97	0.615	0.356
A-arm to hub carrier	0.97	0.572	0.229
A-arm to chassis	0.97	0.259	0.302
Leaf spring connection to hub carrier	0.97	0.488	0.531
Leaf spring connection to chassis	0.97	0.1585	0.600
Damper connection to A-arm	1.045	0.500	0.241
Damper connection to chassis	1.045	0.297	0.632
Tie rod connection to hub carrier	0.83	0.448	0.531
Tie rod connection to chassis	0.83	0.07	0.600
Steering rack connection to chassis	0.83	0.00	0.600

Suspension forces

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 5.5.

Table 5.5: Suspension forces in the nominal configuration (Adapted from [1])

Leaf spring force	2728.9 N
Non-linear Spring-Damper force	128.0 N
Bump stop force	0.0 N

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

$$F_L = -k_L \left(L - (1 + 2728.9/35906\ N/m)\right) \tag{5.3}$$

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

$$F_B = \quad -10^7\,(s - 0.93) \quad ;\ s < 0.93\ m \tag{5.4a}$$

$$F_B = \qquad 0 \qquad\quad ;\ s \geq 0.93\ m \tag{5.4b}$$

The elastic and damping force of the nonlinear spring damper is given by the following expression

$$F_s = -4.0092 \cdot 10^6 + ks - 6.7061 \cdot 10^7 s^2 + 5.2796 \cdot 10^7 s^3 \tag{5.5a}$$

$$F_d = cv + 33955.72v^2 - 59832.25v^3 - 395651.0\,v^4;\ -0.2 < v < 0.21\ m/s \tag{5.5b}$$

$$F_d = -416.4200 + 1844.3\,v;\ v < -0.2\ m/s \tag{5.5c}$$

$$F_d = 1919.1638 + 1634.727v;\ v > 0.21\ m/s \tag{5.5d}$$

where s is the distance between the extreme points of the nonlinear spring-damper. c and k are the dominant suspension damping coefficient and stiffness, where $c = 9945.627\ N\ s/m$, $k = 2.8397 \cdot 10^7\ N/m$.

The tire model

The tire is composed of normal, longitudinal, and lateral models. The normal model is a linear spring-damper element, and the longitudinal and lateral models are linearized models with saturation. [1] the normal model is

$$\mathbf{F}_n = -k_n\,(r - R)\,\mathbf{n};\ r < R \tag{5.6}$$

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 5.6.

Table 5.6: Tire forces in the nominal configuration (Adapted from [1])

Front tyre load	3829.6 N
Rear tyre load	3593.6 N

The longitudinal and lateral models implemented in this study are described in [124].

$$\mathbf{F}_t = F_x \mathbf{b} + F_y \left(\mathbf{n} \times \mathbf{b}\right) \tag{5.7a}$$

$$F_x = \begin{cases} \dfrac{\mu_x \left|\mathbf{F}_{rad}\right|}{\kappa_c}\kappa; \ \kappa \leq \kappa_c \\ \mu_x \left|\mathbf{F}_{rad}\right|; \ \kappa > \kappa_c \end{cases} \tag{5.7b}$$

$$F_y = \begin{cases} \dfrac{\mu_y \left|\mathbf{F}_{rad}\right|}{\alpha_c}\alpha; \ \alpha \leq \alpha_c \\ \mu_y \left|\mathbf{F}_{rad}\right|; \ \alpha > \alpha_c \end{cases} \tag{5.7c}$$

where $\mathbf{u}$ is the unit vector coincident with the wheel rotation axis, $\mathbf{b} = \left(\mathbf{u} \times \mathbf{n}\right)/\left|\mathbf{u} \times \mathbf{n}\right|$ is the longitudinal vector, κ is the longitudinal slip, α is the slip angle, and κ_c, α_c are the critical slip factors for the longitudinal and lateral models, which are parameters of the tire model.

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

$$\kappa = \frac{-\mathbf{b}^T \mathbf{v}_{slip}}{\mathbf{b}^T \mathbf{v}_c} = \frac{-\mathbf{b}^T (\mathbf{v}_c - \mathbf{v}_r)}{\mathbf{b}^T \mathbf{v}_c} = \frac{-\mathbf{b}^T \left(\mathbf{v}_c - \omega \times r\mathbf{n}\right)}{\mathbf{b}^T \mathbf{v}_c} \tag{5.8a}$$

$$\alpha = -\arcsin\left(\mathbf{n}^T \left(\mathbf{b} \times \frac{\mathbf{v}_c - \left(\mathbf{n}^{\mathrm{T}} \mathbf{v}_c\right)\mathbf{n}}{\left|\mathbf{v}_c - \left(\mathbf{n}^{\mathrm{T}} \mathbf{v}_c\right)\mathbf{n}\right|}\right)\right) \tag{5.8b}$$

where $\mathbf{v}_c$ is the velocity of the center of the wheel ω is the angular velocity of the wheel, and r the effective radius defined before.

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

$$\left(\frac{F_x^{sat}}{\mu_x}\right)^2 + \left(\frac{F_y^{sat}}{\mu_y}\right)^2 \leq \left|\mathbf{F}_{rad}\right|^2 \tag{5.9}$$

where μ_x and μ_y stand for the longitudinal and lateral friction coefficients and are parameters of the tire model.

If the components evaluated from Eqn. (5.7b) and Eqn. (5.7c) are not inside the ellipse Eqn. (5.9), the saturation of the forces take place and the previously calculated forces Eqn. (5.7b) and Eqn. (5.7c) doesn't hold. In this case they have to be replaced by the following

$$F_x^{sat} = \frac{|\mathbf{F}_n|}{\sqrt{\left(\frac{F_x}{\mu_x}\right)^2 + \left(\frac{F_y}{\mu_y}\right)^2}} F_x = \frac{|\mathbf{F}_n|}{f_{roz}} F_x \tag{5.10a}$$

$$F_y^{sat} = \frac{|\mathbf{F}_n|}{\sqrt{\left(\frac{F_x}{\mu_x}\right)^2 + \left(\frac{F_y}{\mu_y}\right)^2}} F_y = \frac{|\mathbf{F}_n|}{f_{roz}} F_y \tag{5.10b}$$

5.2.2 Vehicle ride

The ride response, which is closely related to the comfort analysis, is defined as the ability of a vehicle to filter vibration from different vibration sources, such as road roughness, speed bumps, and on-board sources.

Vibration evaluation

Sitting passengers feel vibration from the seat, the floor, and the seat back. However, the contact with seat is the dominant factor. When the whole system is undergoing motion, the passengers are undergoing whole-body vibration. On the other hand, when one or more parts are in contact with the vibration sources, the passengers are undergoing local vibration. In this study, only whole-body vibration is considered. The vibration evaluation is involved with many different concepts, such as sources, magnitude, types, frequency, duration, and direction. Although many attempts have been made, it's very difficult to unify the methods to evaluate the vibration severity. In this study, only the tactile vibrations are evaluated, the acoustic vibrations are not accounted for. According to ISO 2631-1, the whole-body vibration frequency range from 0.5Hz to 80Hz is considered for health, comfort, and perception, and 0.1Hz to 0.5Hz is considered for motion sickness. Since motion sickness will not be discussed here and higher frequency will be filtered by the seat cushion, only frequency range from 0.5Hz to 80Hz will be considered. Vibration

frequency is not the only factor that affects the human body discomfort. Vibration magnitude, duration, and direction also affect the human body discomfort. When frequency is low, the human body is approximate a rigid body and the discomfort is approximately proportional to the weighted acceleration. In order to simplify the evaluation, human body is assumed to be rigid in this study so that acceleration becomes the dominant magnitude.

According to ISO 2631-1, there are two basic evaluation methods, the root-mean-square(RMS) method and the four power vibration dose value(VDV) method.

The RMS method measures the root-mean-square value of the weighted acceleration as follows:

$$\text{RMS} = \sqrt{\frac{1}{T}\int_{t_0}^{t_F} a_w^2(t)\text{dt}} \tag{5.11}$$

where a_w is the weighted acceleration and T is the duration of the measurement.

Comparing with the RMS method, The VDV method is more sensitive to the peaks:

$$\text{VDV} = \sqrt[4]{\int_{t_0}^{t_F} a_w^4(t)\text{dt}} \tag{5.12}$$

Standard tests

In this study, four-post test and the speed bumps test are presented as two standard tests for the research of vehicle ride.

Four-post test

Four-post test is designed for the testing of vehicles, which is used here for the research of vehicle ride. The test system consists of four vertical posts on top of which the tires are placed. These four posts provide random vertical displacements according to a uniformly-distributed random profile.

This study focuses mainly on the affect of suspension design on the comfort analysis, the tire characteristics is not considered here. Thus, the posts are directly connected to the wheel axis instead of the tire surface. The random profile based on white noise is shown in figure 5.7. Between

two profile points, the vertical displacement of the wheel center is interpolated linearly.

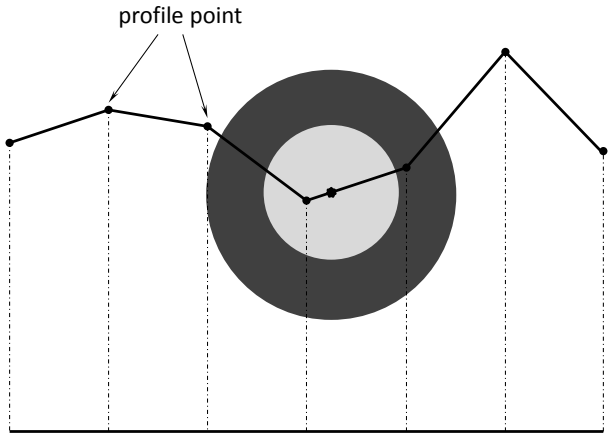


Figure 5.7: The random profile of the four-post test (Adapted from [2])

Speed bumps test

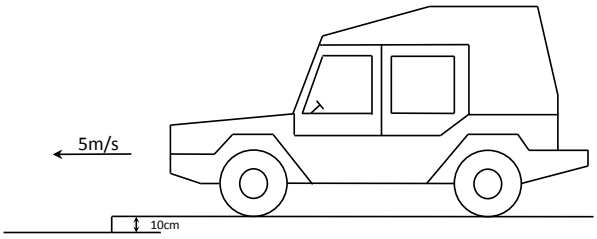


Figure 5.8: The modified speed bumps test

The speed bumps test is a very common test for vehicle ride comfort analysis. Generally, 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 where a step is used to replace the bumps here. Furthermore, since the initial vertical position doesn't correspond to the static equilibrium in the vertical direction, the vehicle system should be released after equilibrium is reached.

5.2.3 Ride optimization

Ride optimization based on four-post test with white noise

As discussed in 5.2.2, four-post test is employed as a case study for the research of vehicle ride. The posts are directly connected to the wheel axis, the actuators on the posts generate white noise for $2s$. By slightly modifying the VDV method for, the objective function becomes the integral of the fourth power of the chassis CG vertical acceleration.

$$\mathbf{\Psi} = \int_{t_0}^{t_F} \ddot{z}_{\text{chassis}}^4 \, dt \tag{5.13}$$

Four different optimization experiments are carried out. The design parameters chosen for these optimization experiments are the rear and front stiffness of the leaf spring $[k_{L1}, k_{L2}]$ from (5.3), the dominant damping of the rear and front suspension $[c_1, c_2]$ from (5.5), and the dominant stiffness of the rear and front suspension $[k_1, k_2]$ from (5.5).

The first optimization experiment is the optimization of the objective function with respect to all six parameters; the second optimization experiment is the optimization of the objective function with respect to $[k_{L1}, k_{L2}]$; the third optimization experiment is the optimization of the objective function with respect to $[c_1, c_2]$; the fourth optimization experiment is the optimization of the objective function with respect to $[k_1, k_2]$. The following constraints on the parameter are imposed to the optimization problem:

$$\begin{aligned} &0 \leq k_{L1}, k_{L2} \leq \infty \ N/m \\ &28300000 \leq k_1, k_2 \leq 28500000 \ N/m \\ &0 \leq c_1, c_2 \leq \infty \ N \ s/m \end{aligned} \tag{5.14a}$$

The initial guess of these parameters are the default values in [1]. The parameters evolution for different experiments are given in Fig. 5.9, Fig. 5.11, Fig. 5.13, and Fig. 5.15 where it is shown that each parameter converges after several iterations. In Fig. 5.10, Fig. 5.12, Fig. 5.14, and Fig. 5.16 the

dynamic responses of the original and of the optimized systems are shown. They clearly illustrate that all the dynamic responses are improved.

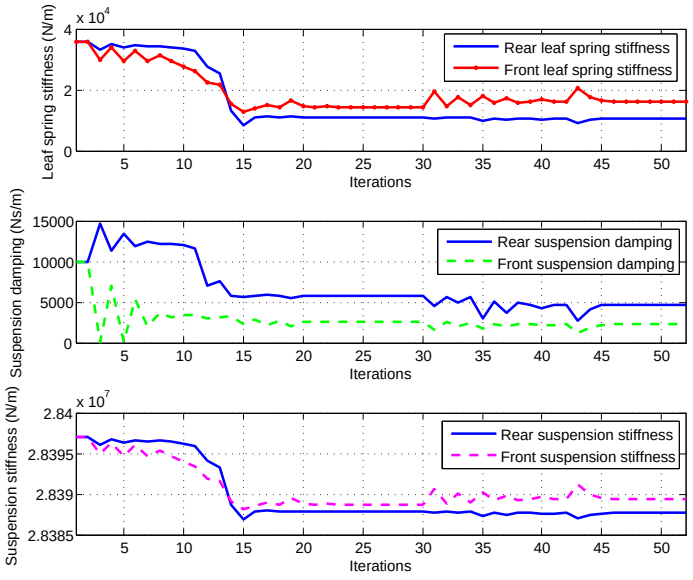


Figure 5.9: The parameters evolution of optimization with 6 parameters for four-post test with white noise

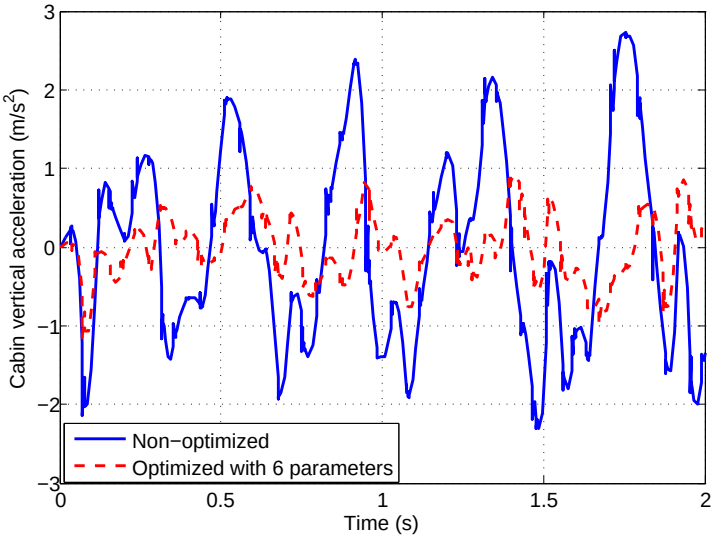


Figure 5.10: Dynamic response of chassis vertical acceleration: non-optimized vs. optimized with 6 parameters for four-post test with white noise

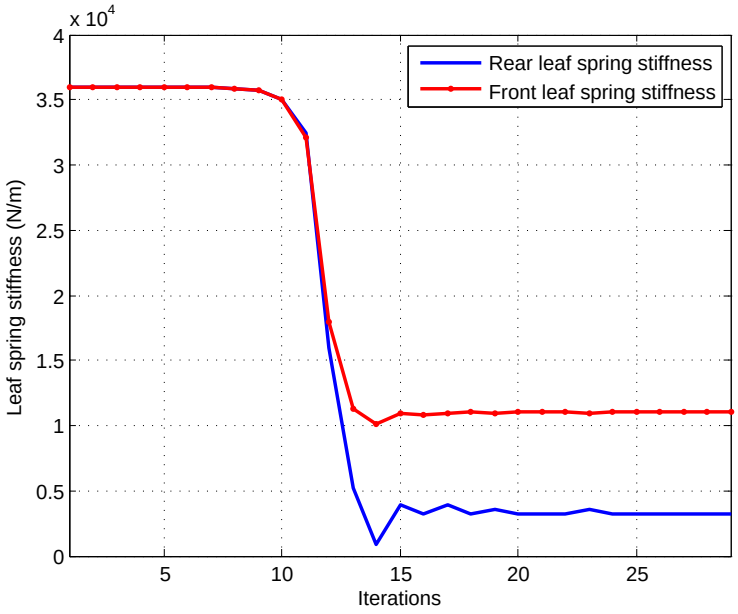


Figure 5.11: The parameters evolution of optimization with $[k_{L1}, k_{L2}]$ for four-post test with white noise

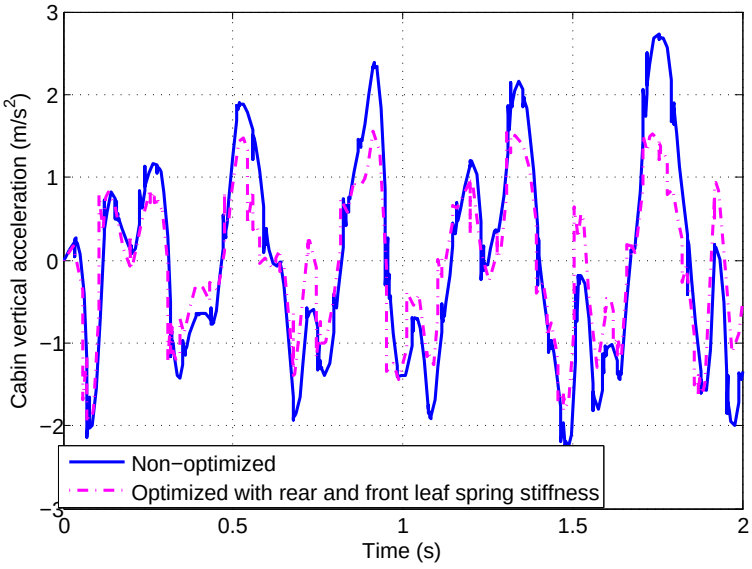


Figure 5.12: Dynamic response of chassis vertical acceleration: non-optimized vs. optimized with $[k_{L1}, k_{L2}]$ for four-post test with white noise

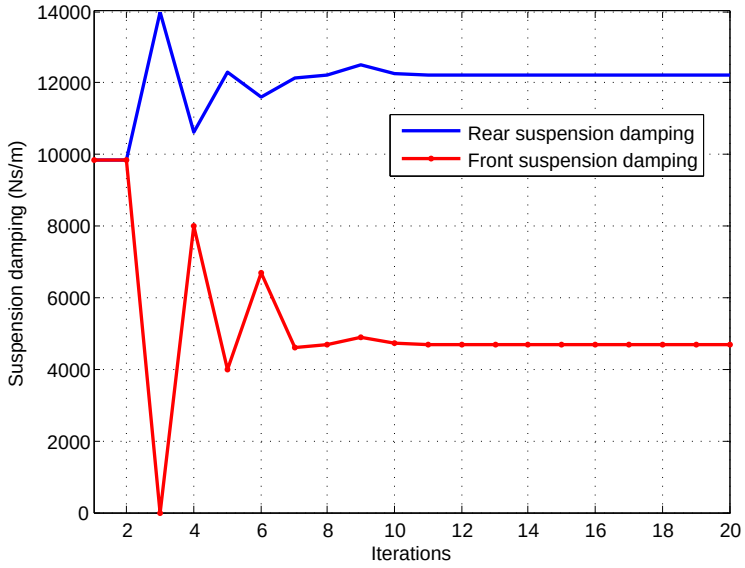


Figure 5.13: The parameters evolution of optimization with $[c_1, c_2]$ for four-post test with white noise

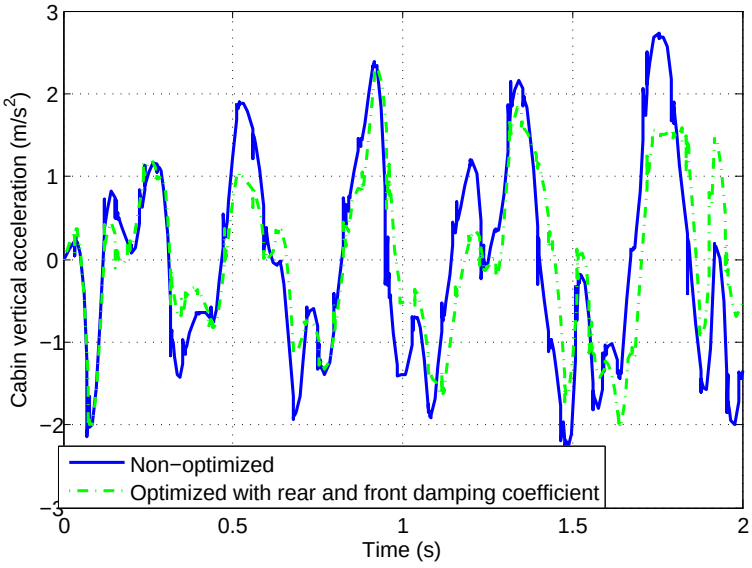


Figure 5.14: Dynamic response of chassis vertical acceleration: non-optimized vs. optimized with $[c_1, c_2]$ for four-post test with white noise

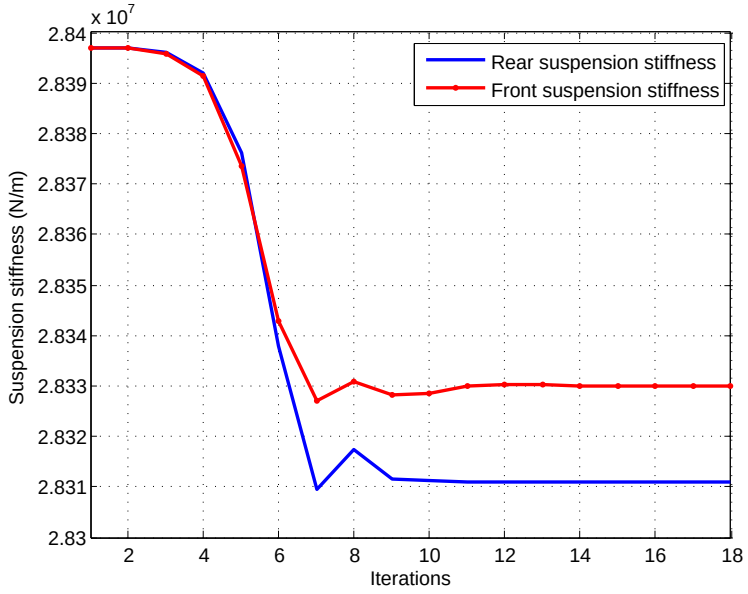


Figure 5.15: The parameters evolution of optimization with $[k_1, k_2]$ for four-post test with white noise

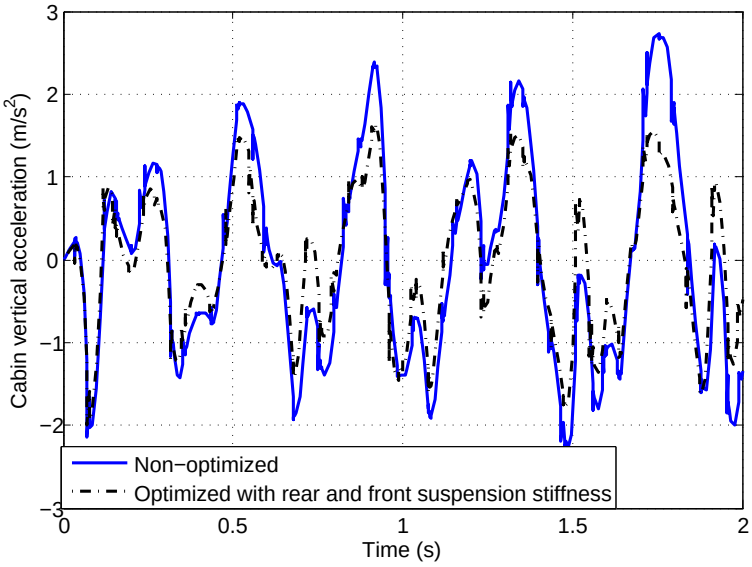


Figure 5.16: Dynamic response of chassis vertical acceleration: non-optimized vs. optimized with $[k_1, k_2]$ for four-post test with white noise

Table 5.7: Objective values before and after optimization (Four-post test with white noise)

	initial Ψ	optimized Ψ
Optimization with 6 parameters	9.85	0.15
Optimization with $[k_{L1}, k_{L2}]$	9.85	2.26
Optimization with $[c_1, c_2]$	9.85	4.2
Optimization with $[k_1, k_2]$	9.85	2.29

Table 5.7 shows the objective value before and after optimization. The values of the optimal parameters for different combinations are shown in table 5.8. It is easily seen that the responses are improved to the largest extent when the optimization is performed for all six parameters. Based on the severity of the effect the suspension parameters were shown to have on the ride comfort, the following ranking can be identified:

$$[k_1, k_2] \approx [k_{L1}, k_{L2}] > [c_1, c_2] \tag{5.15}$$

Table 5.8: Optimized parameters with four-post test with white noise

Optimizations	k_{L1} [N/m]	k_{L2} [N/m]	c_1 [N s/m]	c_2 [N s/m]	k_1 [N/m]	k_2 [N/m]
Original value	35906	35906	9946	9946	28397000	28397000
Optimized with 6 parameters	10611	16255	4699	2239	28387651	28389379
Optimized with k_{L1} and k_{L2}	3252	11014	9946	9946	28397000	28397000
Optimized with c_1 and c_2	35906	35906	12203	4694	28397000	28397000
Optimized with k_1 and k_2	35906	35906	9946	9946	28310905	28330112

Ride optimization based on four-post test with red noise

In the previous section, four-post test with white noise is employed. However, generally the road profile of rear wheels is highly correlated with the road profile of front wheels. Furthermore, white noise has a constant power spectral density, which is not typically generated from any kind of road profile. Therefore, four-post test with correlated red noise is employed to perform vehicle

ride optimization in this section. The actuators on the posts generate correlated red noise for more than 8 s. The vibration provided to the left and respectively to the right side of the vehicle are independent. The vibrations between the front and the rear of the vehicle are correlated with a time delay as follows:

$$\Delta T = L/v \tag{5.16}$$

Same design parameters, cost function, and constraints with equation (5.14a) are chosen to implement optimization experiments.

Like the parameters evolution in the last section, the parameters evolution for different experiments are given in Fig. 5.17, Fig. 5.19, Fig. 5.21, and Fig. 5.23 where it is shown that each parameter converges after several iterations. In Fig. 5.18, Fig. 5.20, Fig. 5.22, and Fig. 5.24 the dynamic responses of the original and of the optimized systems are shown. They clearly illustrate that all the dynamic responses are improved.

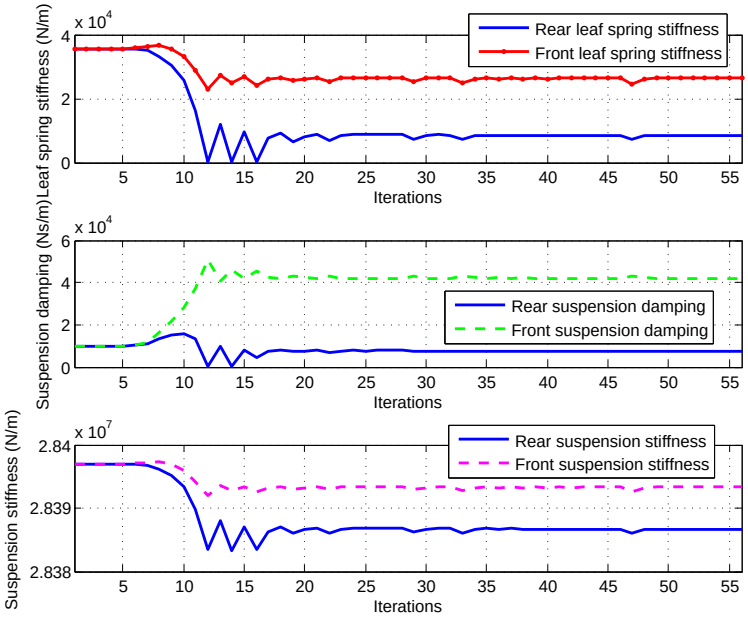


Figure 5.17: The parameters evolution of optimization with 6 parameters for four-post test with red noise

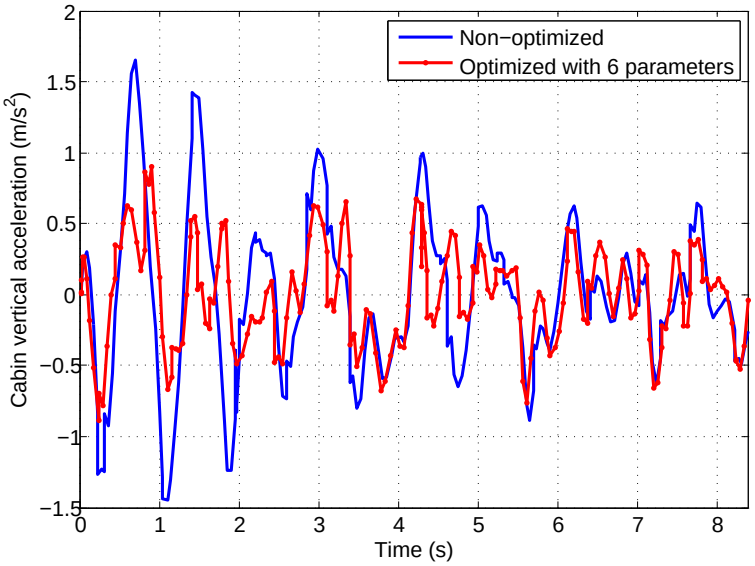


Figure 5.18: Dynamic response of chassis vertical acceleration: non-optimized vs. optimized with 6 parameters for four-post test with red noise

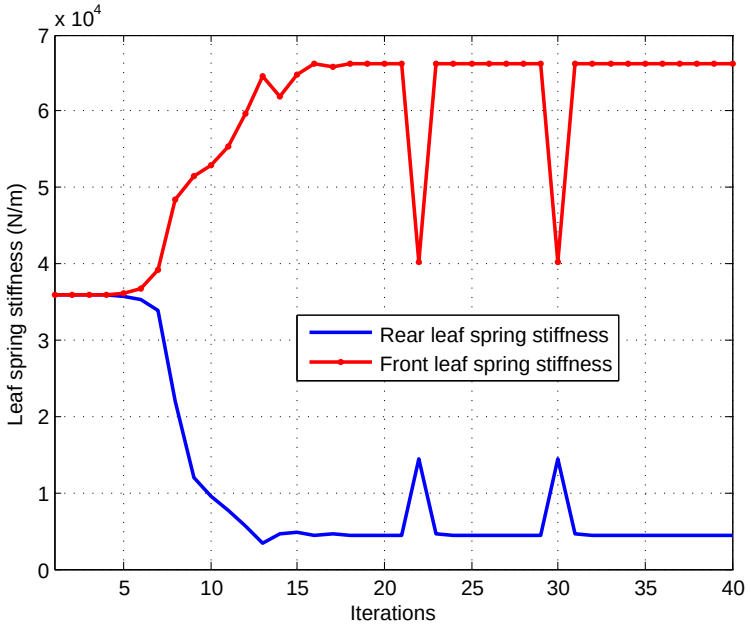


Figure 5.19: The parameters evolution of optimization with $[k_{L1}, k_{L2}]$ for four-post test with red noise

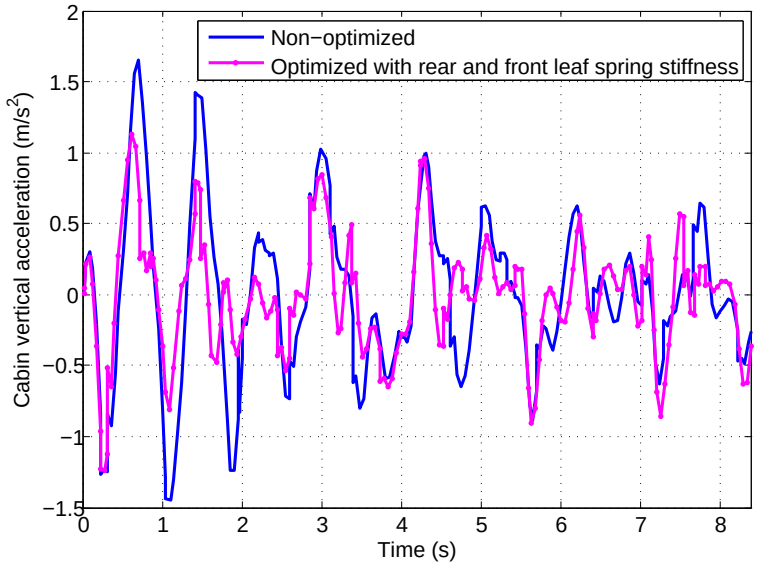


Figure 5.20: Dynamic response of chassis vertical acceleration: non-optimized vs. optimized with $[k_{L1}, k_{L2}]$ for four-post test with red noise

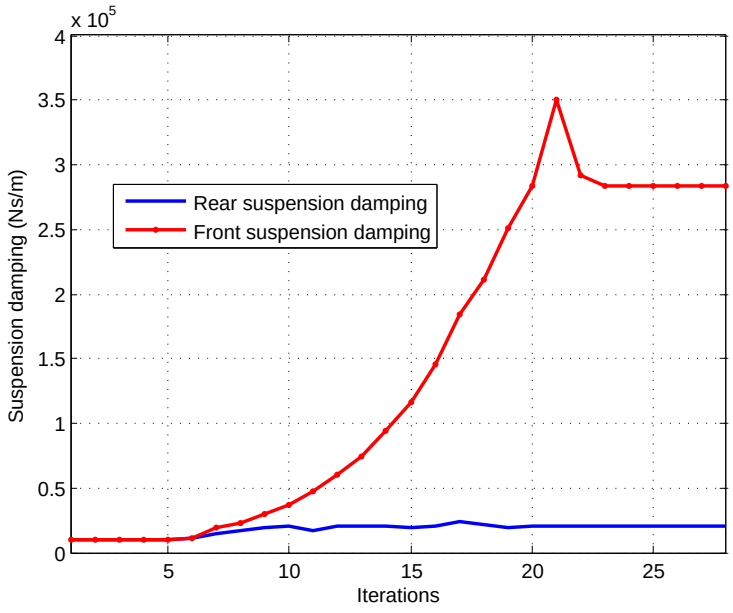


Figure 5.21: The parameters evolution of optimization with $[c_1, c_2]$ for four-post test with red noise

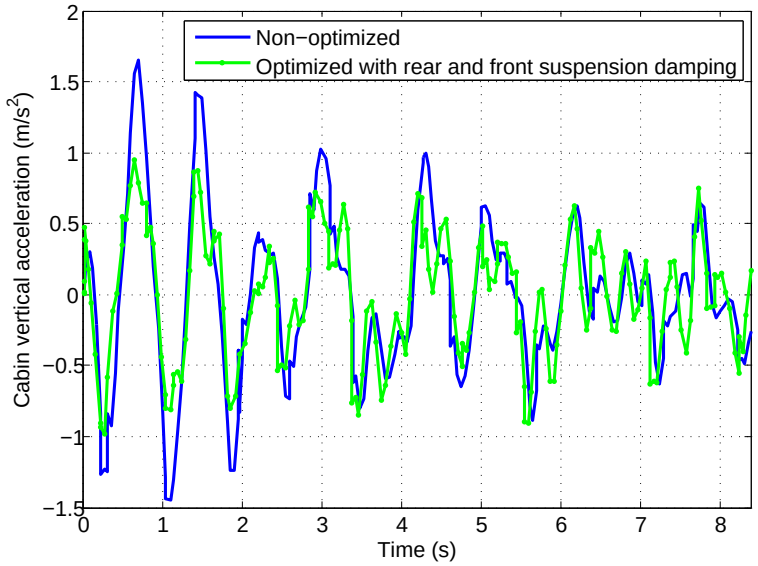


Figure 5.22: Dynamic response of chassis vertical acceleration: non-optimized vs. optimized with $[c_1, c_2]$ for four-post test with red noise

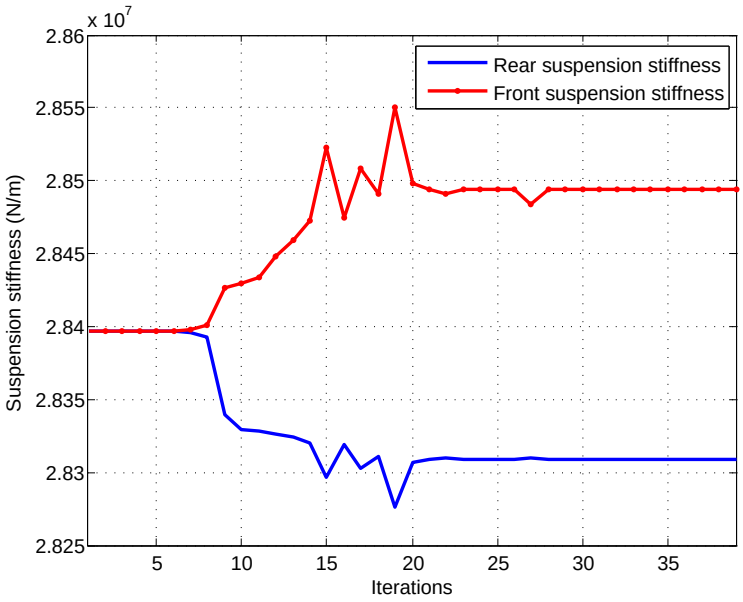


Figure 5.23: The parameters evolution of optimization with $[k_1, k_2]$ for four-post test with red noise

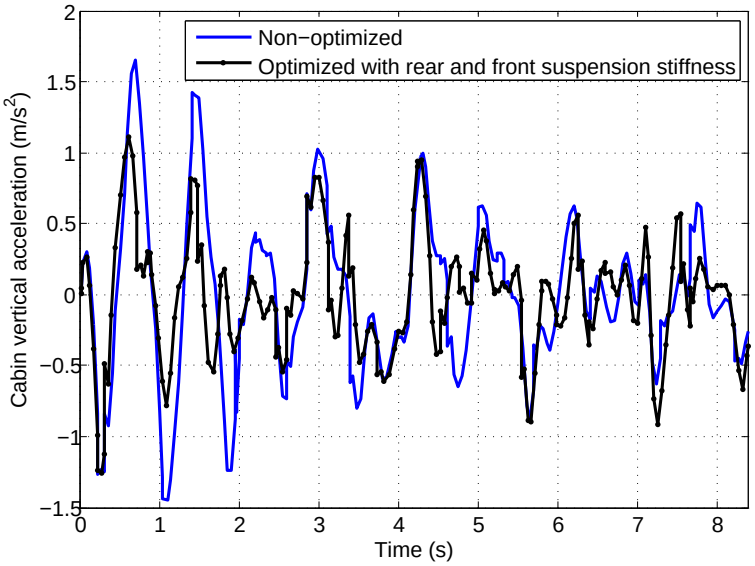


Figure 5.24: Dynamic response of chassis vertical acceleration: non-optimized vs. optimized with $[k_1, k_2]$ for four-post test with red noise

Table 5.9: Objective values before and after optimization (Four-post test with red noise)

	initial Ψ	optimized Ψ
Optimization with 6 parameters	3.11	0.34
Optimization with $[k_{L1}, k_{L2}]$	3.11	0.86
Optimization with $[c_1, c_2]$	3.11	0.66
Optimization with $[k_1, k_2]$	3.11	0.87

Table 5.9 shows the objective value before and after optimization. The values of the optimal parameters for different combinations are shown in table 5.10. It is easily seen that the responses are improved to the largest extent when the optimization is performed for all six parameters. Based on the severity of the effect the suspension parameters were shown to have on the ride comfort, the following ranking can be identified:

$$[c_1, c_2] > [k_1, k_2] \approx [k_{L1}, k_{L2}] \tag{5.17}$$

Table 5.10: Optimized parameters with four-post test with red noise

Optimizations	k_{L1} [N/m]	k_{L2} [N/m]	c_1 [N s/m]	c_2 [N s/m]	k_1 [N/m]	k_2 [N/m]
Original value	35906	35906	9946	9946	28397000	28397000
Optimized with 6 parameters	8663	26572	7684	42141	28386728	28393277
Optimized with k_{L1} and k_{L2}	4491	66284	9946	9946	28397000	28397000
Optimized with c_1 and c_2	35906	35906	20075	283898	28397000	28397000
Optimized with k_1 and k_2	35906	35906	9946	9946	28309081	28494516

Ride optimization based on the speed bumps test

In the speed bumps test, 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. Again, same design parameters, same constraints with equation (5.14a), and same cost function are chosen to implement optimization experiments.

Parameters evolution for different experiments are given in Fig. 5.25, Fig. 5.27, Fig. 5.29, and Fig. 5.31. Again, it is shown that each parameter converges after several iterations. In Fig. 5.26, Fig. 5.28, Fig. 5.30, and Fig. 5.32 the dynamic responses of the original and of the optimized systems are shown. They clearly illustrate that all the dynamic responses are improved.

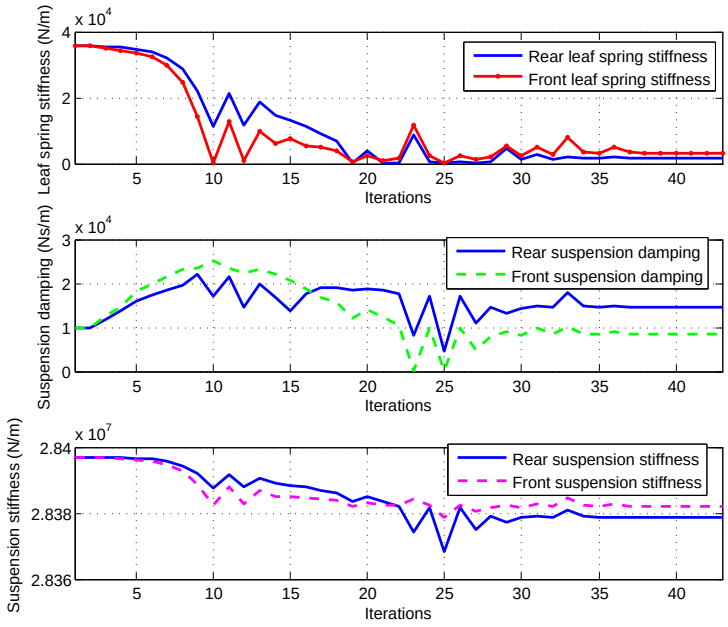


Figure 5.25: The parameters evolution of optimization with 6 parameters for the speed bumps test

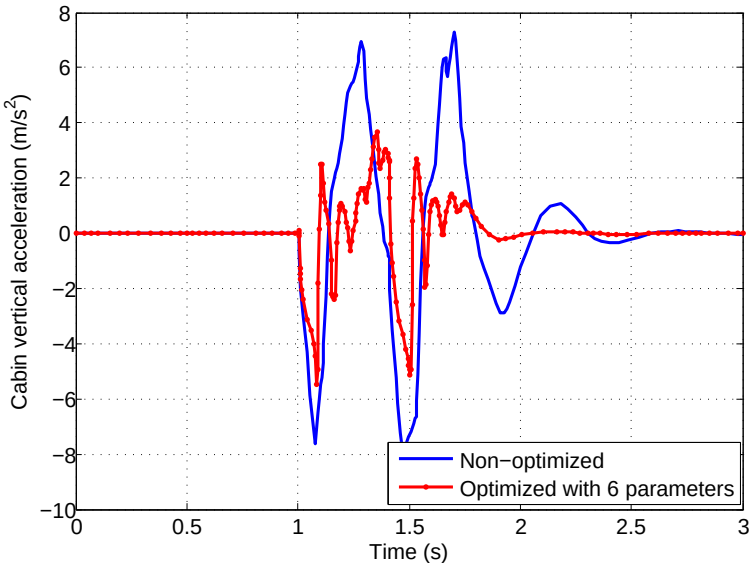


Figure 5.26: Dynamic response of chassis vertical acceleration: non-optimized vs. optimized with 6 parameters for the speed bumps test

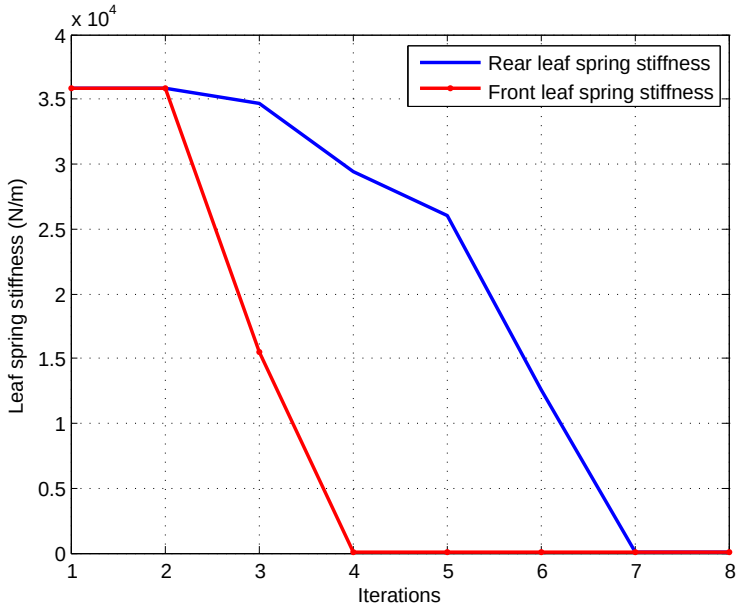


Figure 5.27: The parameters evolution of optimization with $[k_{L1}, k_{L2}]$ for the speed bumps test

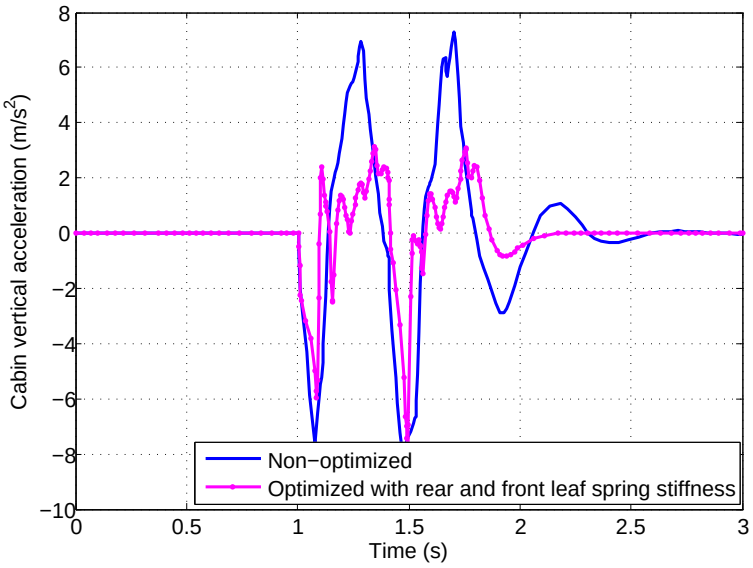


Figure 5.28: Dynamic response of chassis vertical acceleration: non-optimized vs. optimized with $[k_{L1}, k_{L2}]$ for the speed bumps test

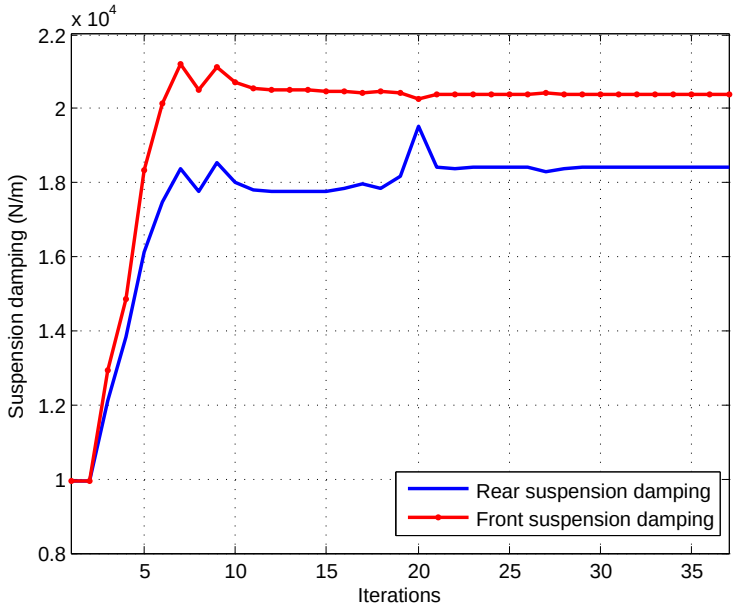


Figure 5.29: The parameters evolution of optimization with $[c_1, c_2]$ for the speed bumps test

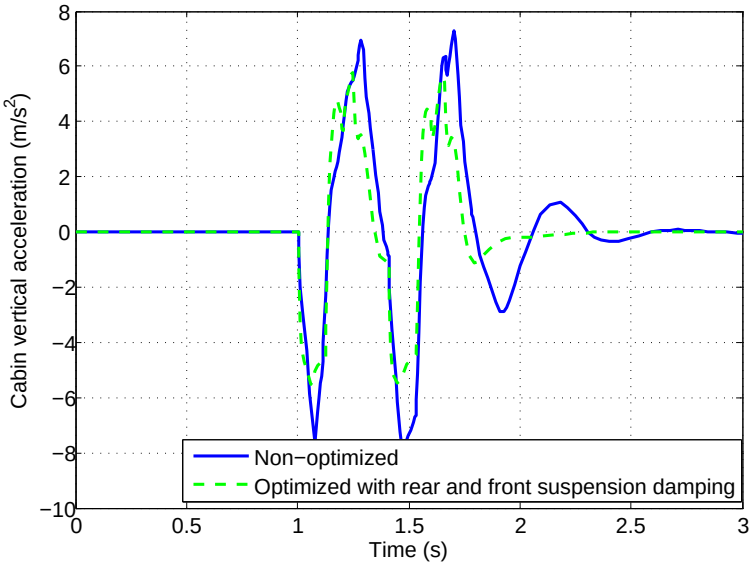


Figure 5.30: Dynamic response of chassis vertical acceleration: non-optimized vs. optimized with $[c_1, c_2]$ for the speed bumps test

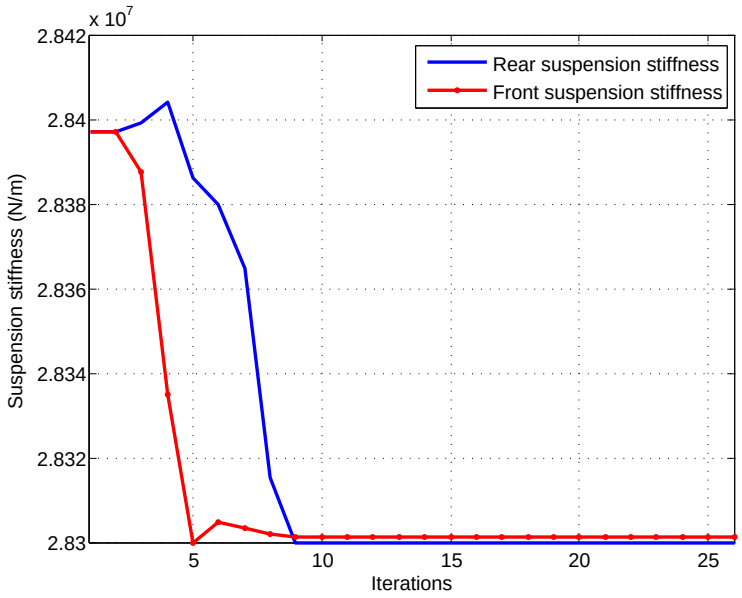


Figure 5.31: The parameters evolution of optimization with $[k_1, k_2]$ for the speed bumps test

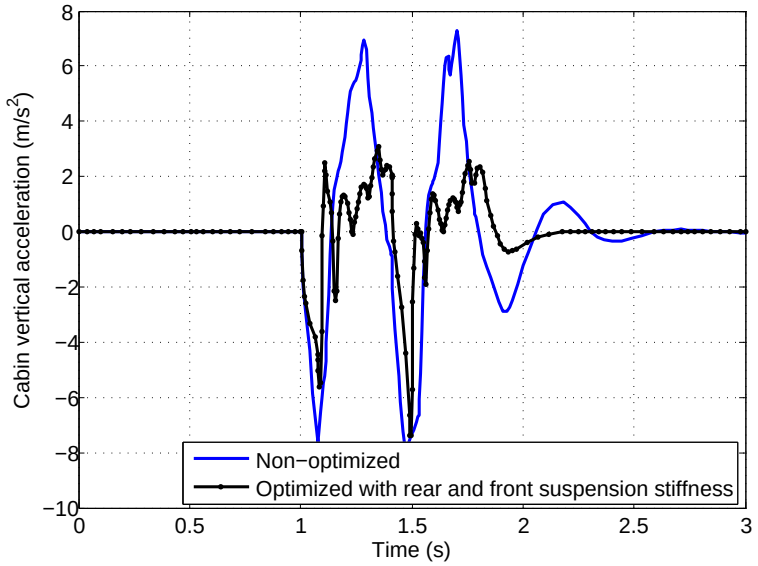


Figure 5.32: Dynamic response of chassis vertical acceleration: non-optimized vs. optimized with $[k_1, k_2]$ for the speed bumps test

Table 5.11: Objective values before and after optimization (the speed bumps test)

	initial Ψ	optimized Ψ
Optimization with 6 parameters	678	48
Optimization with $[k_{L1}, k_{L2}]$	678	96
Optimization with $[c_1, c_2]$	678	257
Optimization with $[k_1, k_2]$	678	81

Table 5.11 shows the objective value before and after optimization. The values of the optimal parameters for different combinations are shown in table 5.12. It is easily seen that the responses are improved to the largest extent when the optimization is performed for all six parameters. Based on the severity of the effect the suspension parameters were shown to have on the ride comfort, the following ranking can be identified:

$$[k_1, k_2] > [k_{L1}, k_{L2}] > [c_1, c_2] \tag{5.18}$$

Table 5.12: Optimized parameters with the speed bumps test

Optimizations	k_{L1} [N/m]	k_{L2} [N/m]	c_1 [N s/m]	c_2 [N s/m]	k_1 [N/m]	k_2 [N/m]
Original value	35906	35906	9946	9946	28397000	28397000
Optimized with 6 parameters	1497	3052	14531	8418	28378782	28382080
Optimized with k_{L1} and k_{L2}	0	0	9946	9946	28397000	28397000
Optimized with c_1 and c_2	35906	35906	18399	20402	28397000	28397000
Optimized with k_1 and k_2	35906	35906	9946	9946	28300000	28301252

5.2.4 Vehicle handling

Handling analysis is always the counterpoint to ride analysis. Usually, stiffer suspension systems increase the road-hold ability during corning and turning, but also increase vibration discomfort. In general, vehicle handling of a vehicle can be defined as the ability to perform transverse to their direction of motion during turning and cornering, with precision and promptness.

There are many maneuvers in which handling response can be test. Some good examples are double lane-change (ISO 3888-1), obstacle avoidance (ISO 3888-2), lateral transient response test methods (ISO 7401), and steady-state circular driving behavior (ISO 4138). This study employs double lane-change(DLC) as a case study for the research of handling response.

Double lane-change maneuver

According to ISO 3888-1, the double lane-change maneuver consists of turning first to the adjacent lane and returning to the initial lane without exceeding lane boundaries. The basic idea of this maneuver is to test the lateral stability of vehicles. The suspension design is critical for the lateral stability. For this reason, a basic requirement for suspension design is to maintain a good lane-change performance. Otherwise, the vehicle can lose grip or even roll over in some extreme maneuvers.

Assume the vehicle width is 1.8 m, the track dimensions for a double lane-change are presented as follows according to ISO 3888-1:

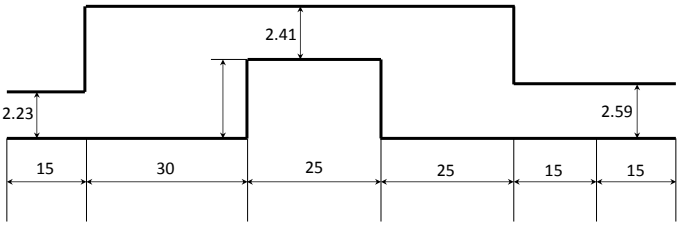


Figure 5.33: Double lane-change track (in meters)

5.2.5 Handling optimization

As discussed in 5.2.4, this study employs a double lane-change maneuver as a case study for the research of vehicle handling. The test speed is 50 km/h. The objective function is the integral of the square of the chassis CG roll rate.

$$\mathbf{\Psi} = \int_{t_0}^{t_F} \dot{\theta}^2_{\text{chassis}}\, dt \tag{5.19}$$

Four different optimization experiments are carried out. The design parameters chosen are the same as for the four-post test. The following constraints on the parameter are imposed to the optimization problem:

$$
\begin{aligned}
&0 \leq k_{L1}, k_{L2} \leq 200000 \ N/m \\
&0 \leq k_1, k_2 \leq 29000000 \ N/m \\
&0 \leq c_1, c_2 \leq 100000 \ N \ s/m
\end{aligned}
\tag{5.20a}
$$

The parameters evolutions for different experiments are given in Fig. 5.34, Fig. 5.36, Fig. 5.38, and Fig. 5.40. The dynamic responses of original and optimized systems are shown in Fig. 5.35, Fig. 5.37, Fig. 5.39, and Fig. 5.41. These figures clearly illustrate that all the dynamic responses are improved.

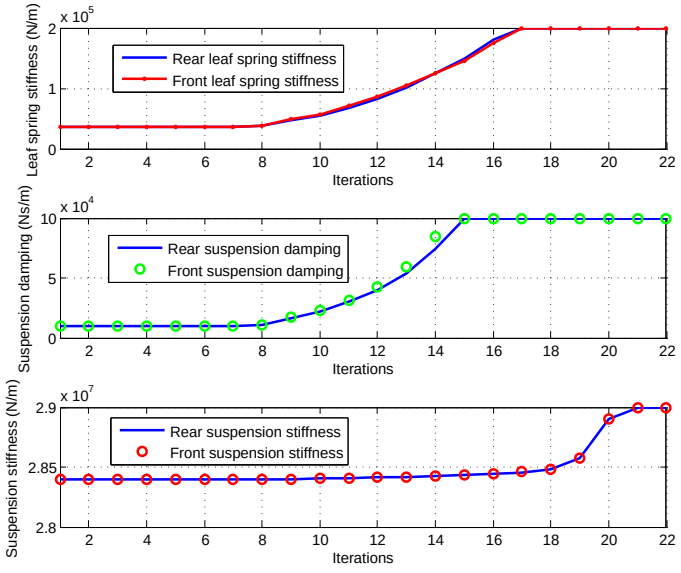


Figure 5.34: The parameters evolution of optimization with 6 parameters for the double lane-change maneuver

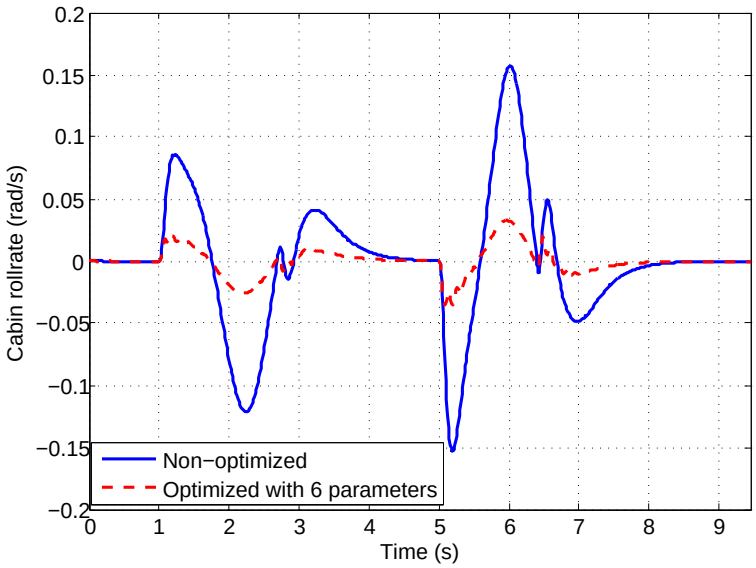


Figure 5.35: Dynamic response of chassis roll rate: non-optimized vs. optimized with 6 parameters for the double lane-change maneuver

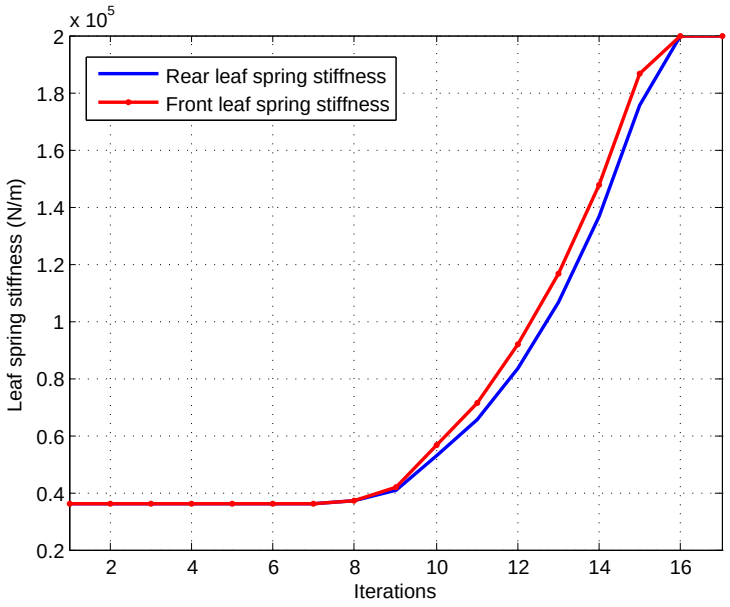


Figure 5.36: The parameters evolution of optimization with $[k_{L1}, k_{L2}]$ for the double lane-change maneuver

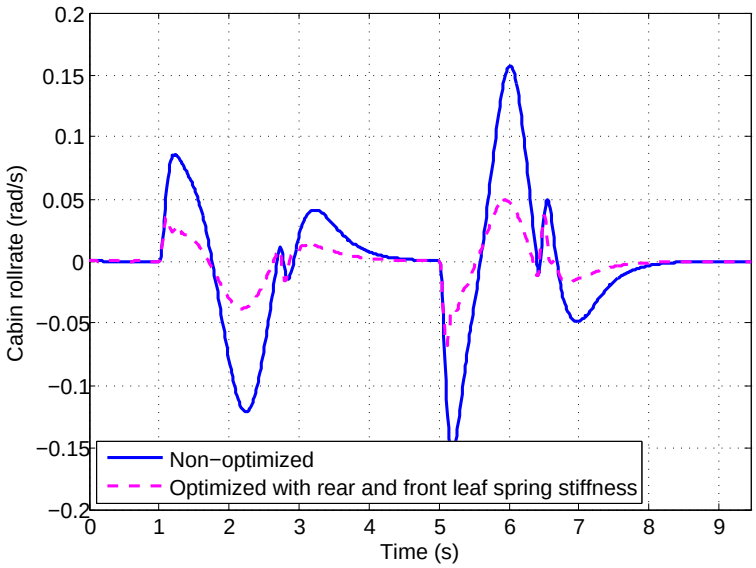


Figure 5.37: Dynamic response of chassis roll rate: non-optimized vs. optimized with $[k_{L1}, k_{L2}]$ for the double lane-change maneuver

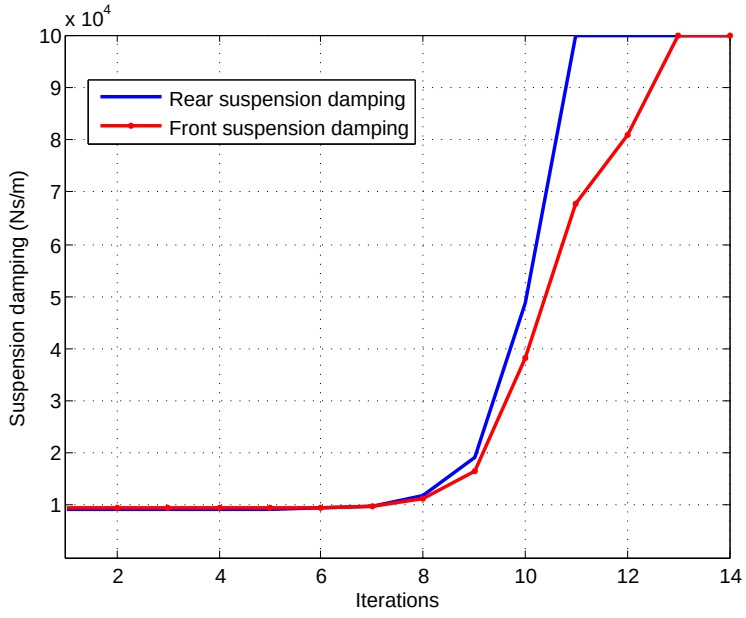


Figure 5.38: The parameters evolution of optimization with $[c_1, c_2]$ for the double lane-change maneuver

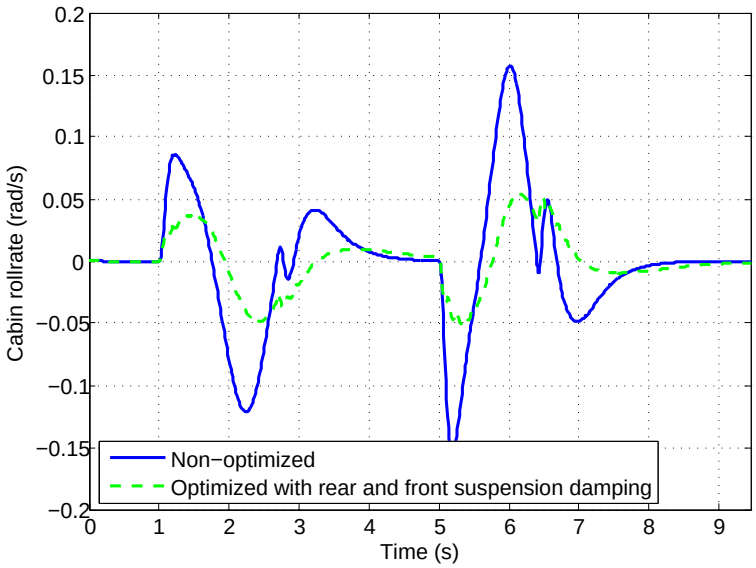


Figure 5.39: Dynamic response of chassis roll rate: non-optimized vs. optimized with $[c_1, c_2]$ for the double lane-change maneuver

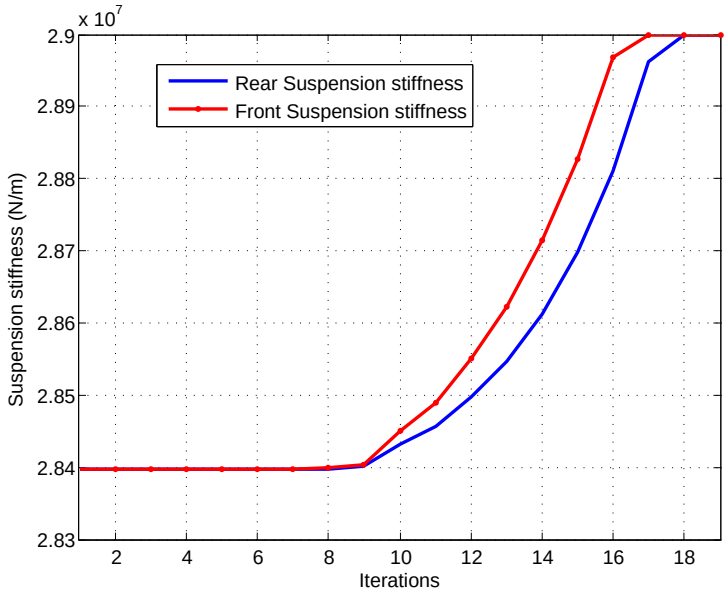


Figure 5.40: The parameters evolution of optimization with $[k_1, k_2]$ for the double lane-change maneuver

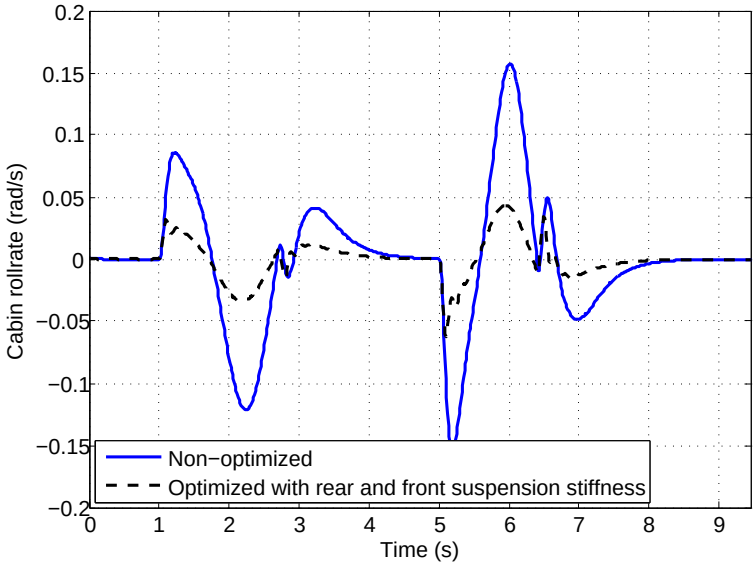


Figure 5.41: Dynamic response of chassis roll rate: non-optimized vs. optimized with $[k_1, k_2]$ for the double lane-change maneuver

Table 5.13: Objective values before and after optimization (the double lane-change maneuver)

	initial Ψ	optimized Ψ
Optimization with 6 parameters	0.0257	0.0012
Optimization with $[k_{L1}, k_{L2}]$	0.0257	0.0028
Optimization with $[c_1, c_2]$	0.0257	0.0048
Optimization with $[k_1, k_2]$	0.0257	0.0022

Table 5.14: Optimized parameters with the double lane-change maneuver

Optimizations	k_{L1} [N/m]	k_{L2} [N/m]	c_1 [N s/m]	c_2 [N s/m]	k_1 [N/m]	k_2 [N/m]
Original value	35906	35906	9946	9946	28397000	28397000
Optimized with 6 parameters	200000	200000	100000	100000	29000000	29000000
Optimized with k_{L1} and k_{L2}	200000	200000	9946	9946	28397000	28397000
Optimized with c_1 and c_2	35906	35906	100000	100000	28397000	28397000
Optimized with k_1 and k_2	35906	35906	9946	9946	29000000	29000000

Theoretically, the stiffer the suspension system is, the better are the handling characteristics. The results of the optimization experiments implemented in this section perfectly validate this assumption. From 5.14 it is clearly seen that all the parameters converge to their upper bounds. Furthermore, table 5.13 shows that the optimization for all six parameters looks better than the other optimization experiments that involved fewer design parameters.

5.3 Dynamic analysis of a passive dynamic robot

An interesting area of application for the multibody dynamics techniques is robotics. Thus, in this study, the dynamic simulation of a passive dynamic robot with feet and knees is implemented. A passive dynamic robot is able to perform a stable walking motion without any motor and controller, which is a pretty good case to study human walking behavior. In the last 24 years a lot of work has already been done for the research of passive dynamic robot. The reader is referred to [125–139] for details.

Unlike the traditional methods used for modeling of robotic systems in the last 24 years, different coordinate system and formulation are employed for the dynamic analysis for a passive dynamic robot in this study. Generally, relative coordinates and Lagrange formulation are used for the dynamic analysis for robotic systems. There are several drawbacks associated to the use of this approach. First, it is difficult to assess if all the DOF associated with these coordinates will be valid during the entire simulation; the system may become unstable when it goes through a singular or a bifurcation position. Second, it is more difficult to model and to write the equations with relative coordinates than with reference point coordinates. Thus, in this study, reference point coordinates and the penalty formulation are employed to perform dynamic analysis for the passive dynamic robot, significantly simplifying the modeling stage and making the robotic system more stable. In addition, the passive dynamic robot is also used to test and validate all the point contact and surface contact models developed in MBSVT.

5.3.1 Contact-impact force models

The treatment of impact and contact forces is always a challenging issue in multibody systems with instantaneous impacts or contacts of a duration between different bodies. Generally, there are two families of methods to solve the impact problems: the discontinuous methods and the continuous methods [140]. The discontinuous method assume the impact occurs instantaneously [141, 142] while the continuous method relates the force and the deformation when impact occurs [140, 143].

In this study, the continuous method and the penalty formulation are used to model the impact and contact forces for a passive dynamic robot. Several different contact and friction models, which can be used to model point contact and surface contact, are developed and included in MBSVT 6. These models include a static friction model, Ambrósio dry friction model, Kelvin-Voigt viscous-elastic model, and a simplified tire model, which will be briefly introduced in the coming section. For other impact and contact models the reader is referred to [3, 4]

Point contact models

- Kelvin-Voigt viscous-elastic model

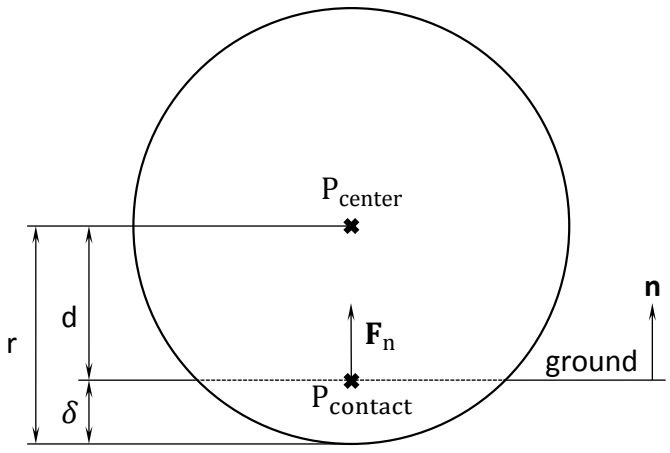


Figure 5.42: Normal contact between sphere and plane (Adapted from [3])

$$\mathbf{F}_n = \begin{cases} K\delta^n + D\dot{\delta}; & \dot{\delta} > 0 \quad (loading\ phase) \\ K\delta^n + cD\dot{\delta}; & \dot{\delta} < 0 \quad (unloading\ phase) \end{cases} \tag{5.21}$$

In this study, Kelvin-Voigt viscous-elastic model is employed to model the normal contact between sphere and plane, this model assumes that both the spring and the damper are linear. As shown in figure 5.42 and equation 5.21, r is the radius of the sphere, d is distance between the CG of the sphere and the plane, δ is the relative penetration depth, $\mathbf{n}$ is the direction of the force that is perpendicular to the plane, $\mathbf{F}_n$ is the normal force, K is the stiffness, D is the damping, c is the restitution coefficient, $\dot{\delta}$ is the relative normal velocity of the colliding bodies. The normal impact is compose of two phases: loading phase and unloading phase. In the loading phase two contact bodies are compressing each other while in the unloading phase two contact bodies are separating from each other. As shown in equation 5.21, the energy loss from impact is included by multiplying the damping force with a restitution coefficient c, where $0 < c < 1$.

- Ambrósio dry friction model

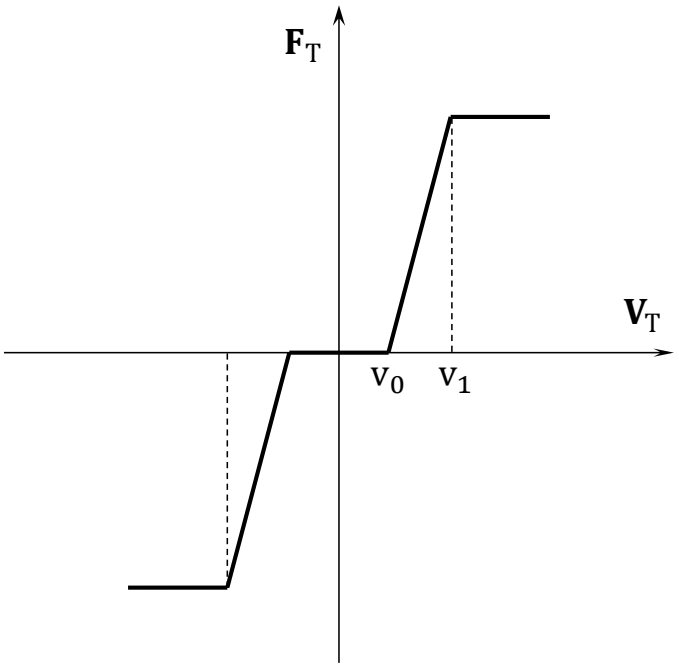


Figure 5.43: Ambrósio friction force (Adapted from [4])

Ambrósio friction model, which is obtained by making some modifications to Coulomb's friction law, is used to model the dry friction in this study. Ambrósio friction force is expressed as:

$$\mathbf{F}_T = -\mu c_d f_n \frac{\mathbf{V}_T}{v_T} \tag{5.22}$$

where μ is the dry friction coefficient, $\mathbf{F}_n$ is the normal contact force, $\mathbf{V}_T$ is the relative tangential velocity between contact bodies and c_d is a correction coefficient, which can be expressed as:

$$c_d = \begin{cases} 0 & if \quad v_T < v_0 \\ \frac{v_T - v_0}{v_1 - v_0} & if \quad v_0 < v_T < v_1 \\ 1 & if \quad v_T > v_1 \end{cases} \tag{5.23}$$

- A static friction model

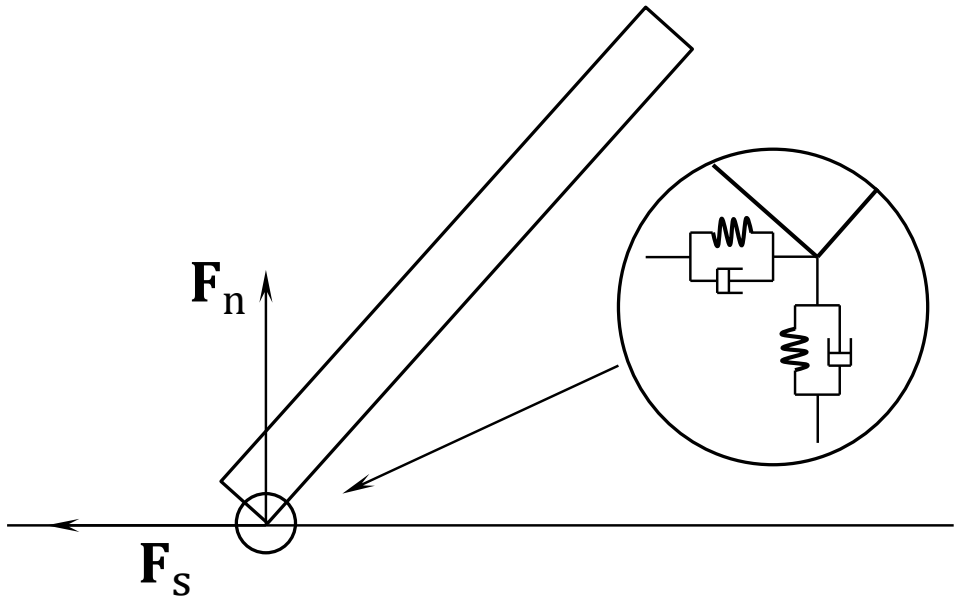


Figure 5.44: A static friction model

As shown in figure 5.44, the normal contact force and the tangential force are both modeled as spring-damper systems. As mentioned before, the normal contact force is Kelvin-Voigt viscous-elastic model and the force direction is perpendicular to the contact plane. The tangential force, which is also the static friction force here, is parallel to the contact plane.

With a large stiffness and damping, the functionality of the contact would be very similar with a revolute joint.

- A simplified tire model

  A simplified tire model is used in this study that is specifically described in section 5.2.1.

Surface contact model

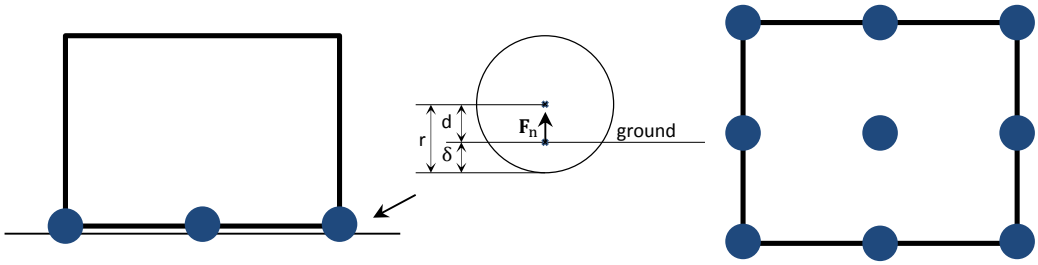


Figure 5.45: Surface contact model

Generally, most contacts occur between surfaces. In this study, the surface contact is modeled by adding multiple point contact to the contact surface. Figure 5.45 shows the front view and bottom view of a typical surface contact, where each blue circle is a point contact. Multiple point contact can be also used to model complex surface contact models, in this case different point contact might have different radius of the sphere, stiffness, and damping. In this study, multiple point contact is used to model the surface contact between the feet and the slope for a passive dynamic robot, which will be described in section 5.3.3.

5.3.2 Dynamic analysis of a passive dynamic robot with point contact

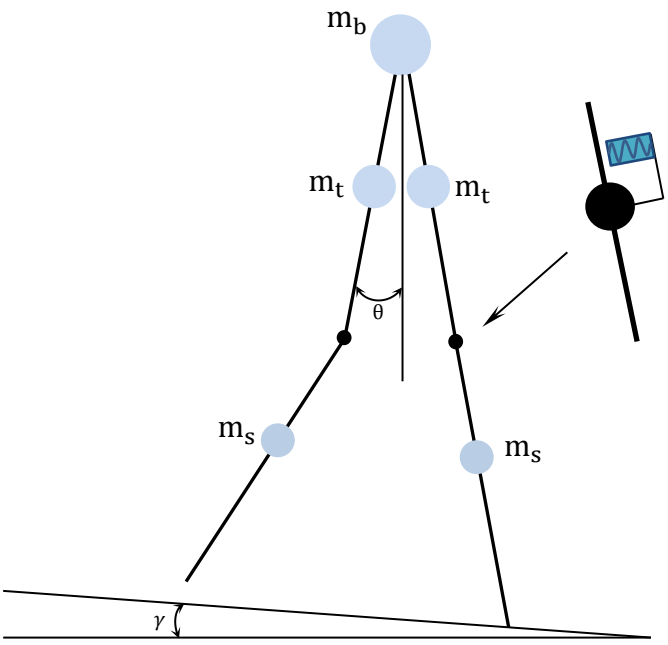


Figure 5.46: passive dynamic robot with point contact (Adapted from [5])

From Hsu's thesis [5], the dynamic analysis of a passive dynamic robot is specifically described. In this study, the same model and initial conditions from [5] is employed to perform dynamic analysis of a passive dynamic robot by using the penalty formulation and reference point coordinates system. As shown in figure 5.46, the kneed robot is placed on slope with a fixed angle γ from the horizontal plane. A point contact is added between the robot feet and the slope. A bump stop is installed on the knee to prevent it from being over rotated. The mass of the body, the thigh, and the shank are expressed as m_b, m_t, and m_s respectively. θ is the relative angle between the upper link of leg 1 and the vertical line, which will be used later to show us the phase portrait of the gait. For other physical parameters and initial conditions, the reader is referred to [5].

Dynamic analysis of a passive dynamic robot with static friction force

The first numerical experiment is to add static friction forces between robot feet and the slope. With a large stiffness and damping, the functionality of the contact would be very similar with a revolute joint. Initially the robot is placed on the slope. With the affect of gravity, the robot is walking along the slope automatically for 5 seconds, as shown in figure 5.47:

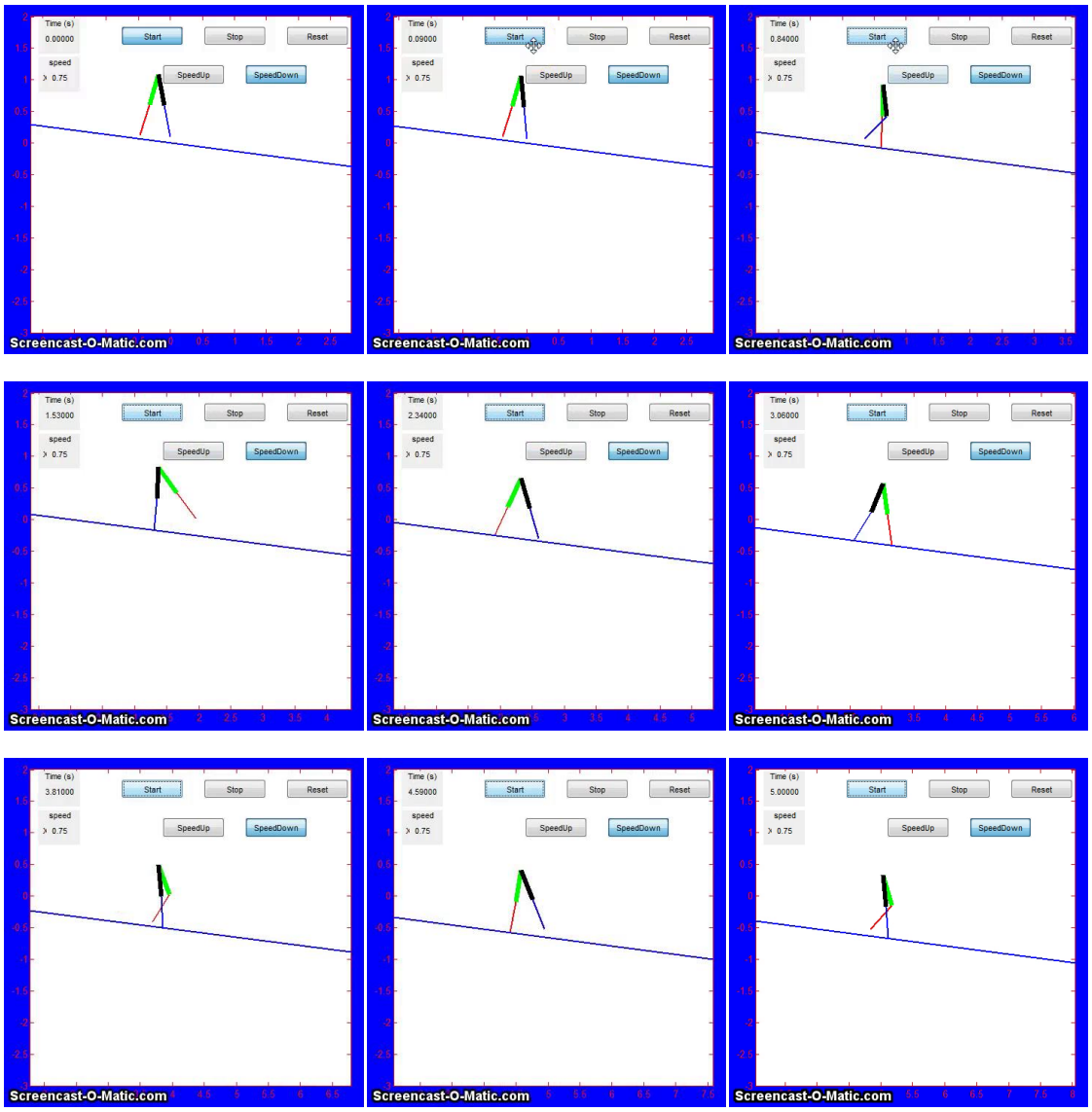


Figure 5.47: Screenshots from dynamic simulation for the first experiment

The phase portrait of the periodic gait is shown in figure 5.48 as follows:

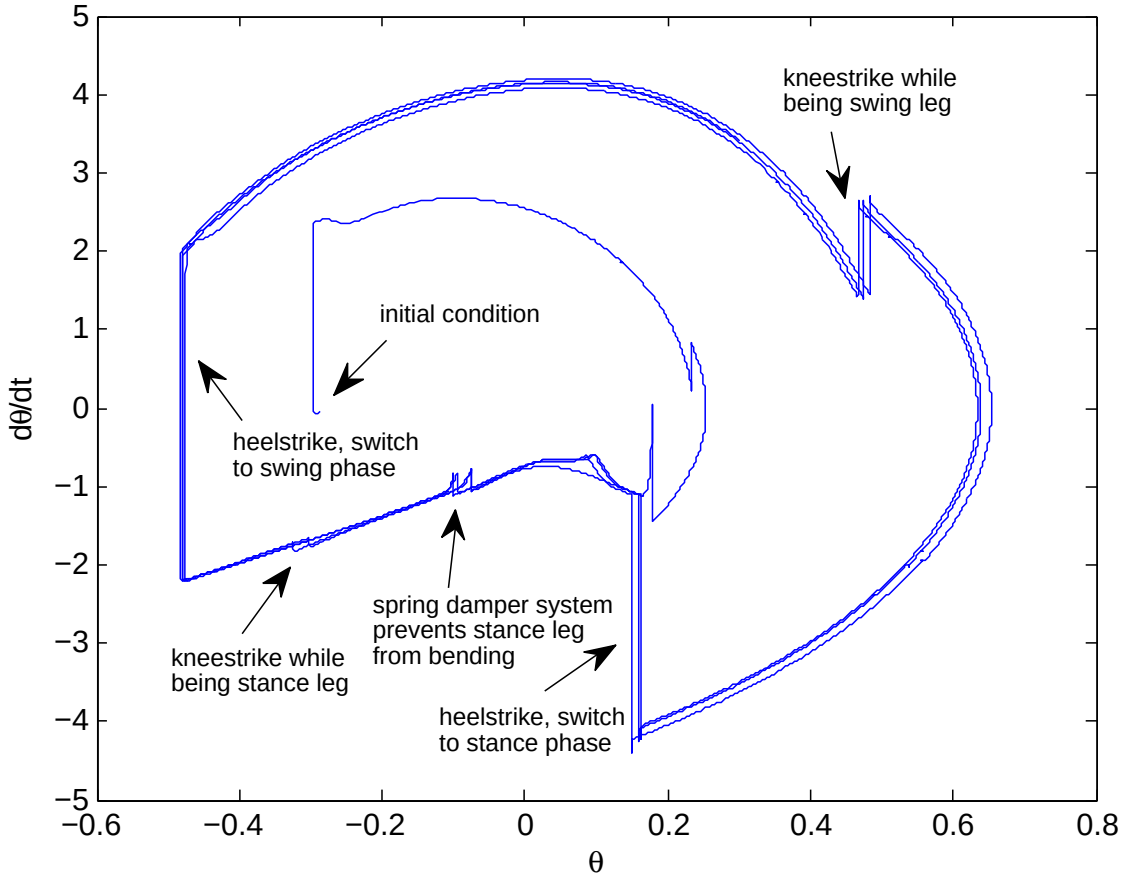


Figure 5.48: Phase portrait for the upper link of the first leg for the first experiment

From the phase portrait presented in 5.48 it's clear to see that, with the given initial conditions, the robot simulation converges to the stable limit cycle within a few steps. At the beginning, the stance leg that is leg 2 is placed on the slope and the swing leg that is leg 1 is released with an initial angular velocity. At a certain moment the knee-strike occurs when the swing leg straightens out, which leads to an instantaneous angular velocity change for both legs. Then the swing leg swings as a single link before heel-strike to the slope. After the heel-strike, the knee is locked by a spring-damper system that connects the CG of the thigh and shack and the swing leg switches to the stance leg. After becoming the stance leg, the spring-damper system that generally has high stiffness and damping will prevent the knee from bending. Then, the knee-strike from the swing

leg leads to an small instantaneous angular velocity change for both legs again. Finally, after the heel-strike of the swing leg, the stance leg switches to the swing leg and start a new cycle again.

Dynamic analysis of a passive dynamic robot with dry friction force

In the first numerical experiment static friction forces are used and the robot converges to the limit cycle within a few steps. In the second numerical experiment static friction forces are replaced by Ambrósio friction force with $v_0 = 1^{-8}$, $v_1 = 1^{-5}$, and $\mu = 0.9$ and the phase portrait is shown as follows:

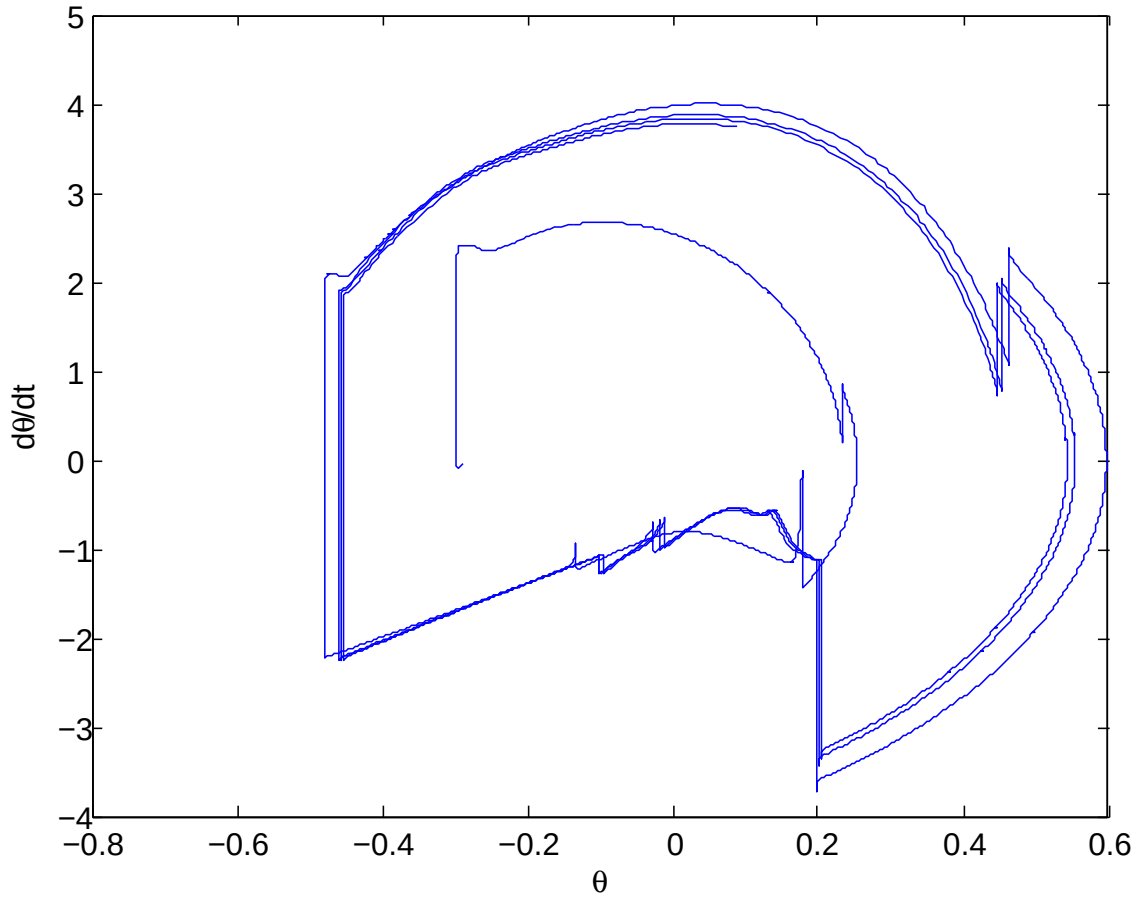


Figure 5.49: Phase portrait for the upper link of the first leg for the second experiment

From the phase portrait 5.49 it's clear to see the robot simulation converges to the stable limit cycle

within a few steps just like the first experiment. This means with a large dry friction coefficient $\mu = 0.9$, the robot is still very stable.

In the third experiment, the same contact forces are used except the dry friction coefficient is changed from 0.9 to 0.2. The robot becomes unstable and falls over after one step, as shown in figure 5.50.

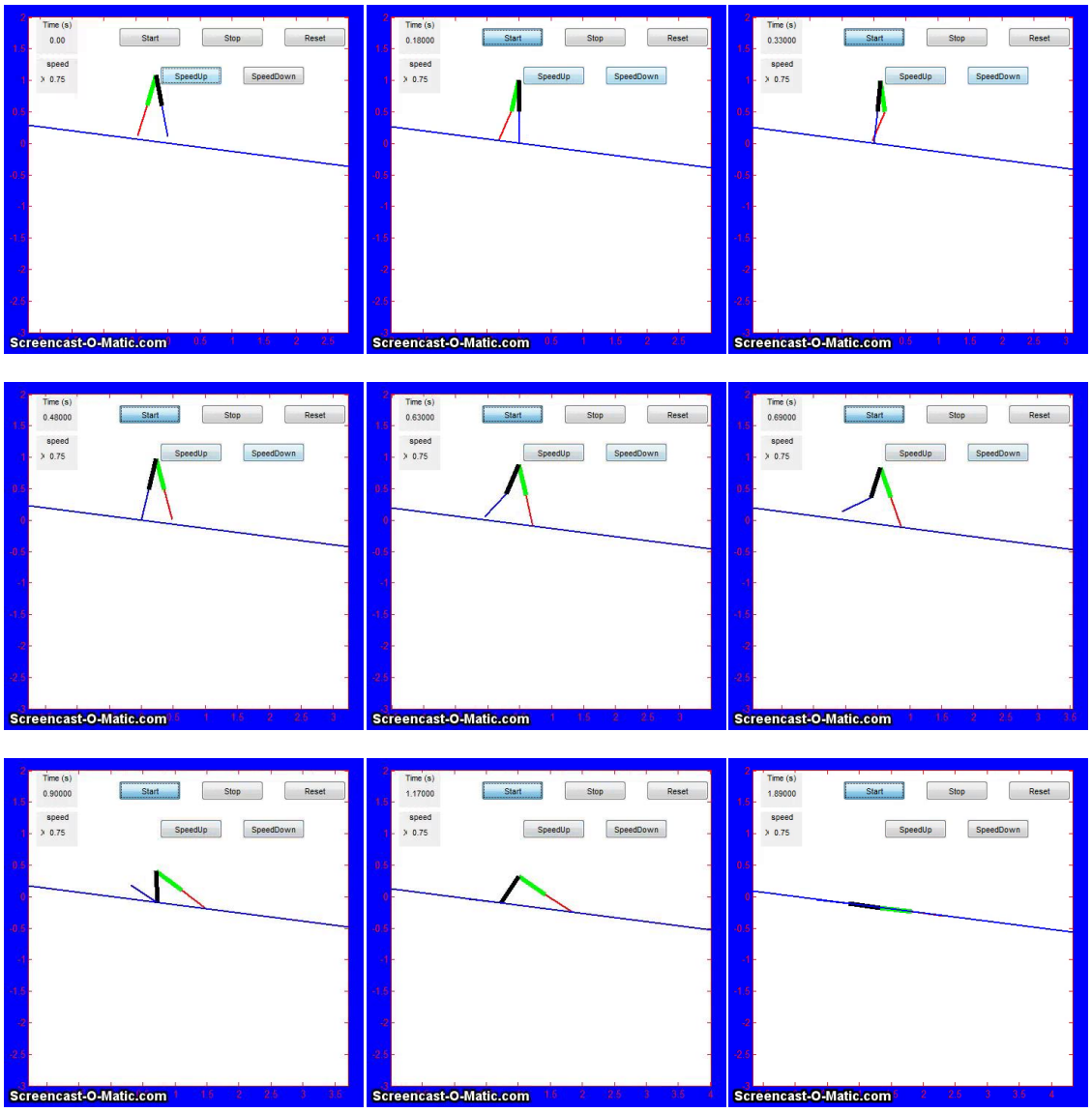


Figure 5.50: Screenshots from dynamic simulation

5.3.3 Dynamic analysis of a passive dynamic robot with surface contact

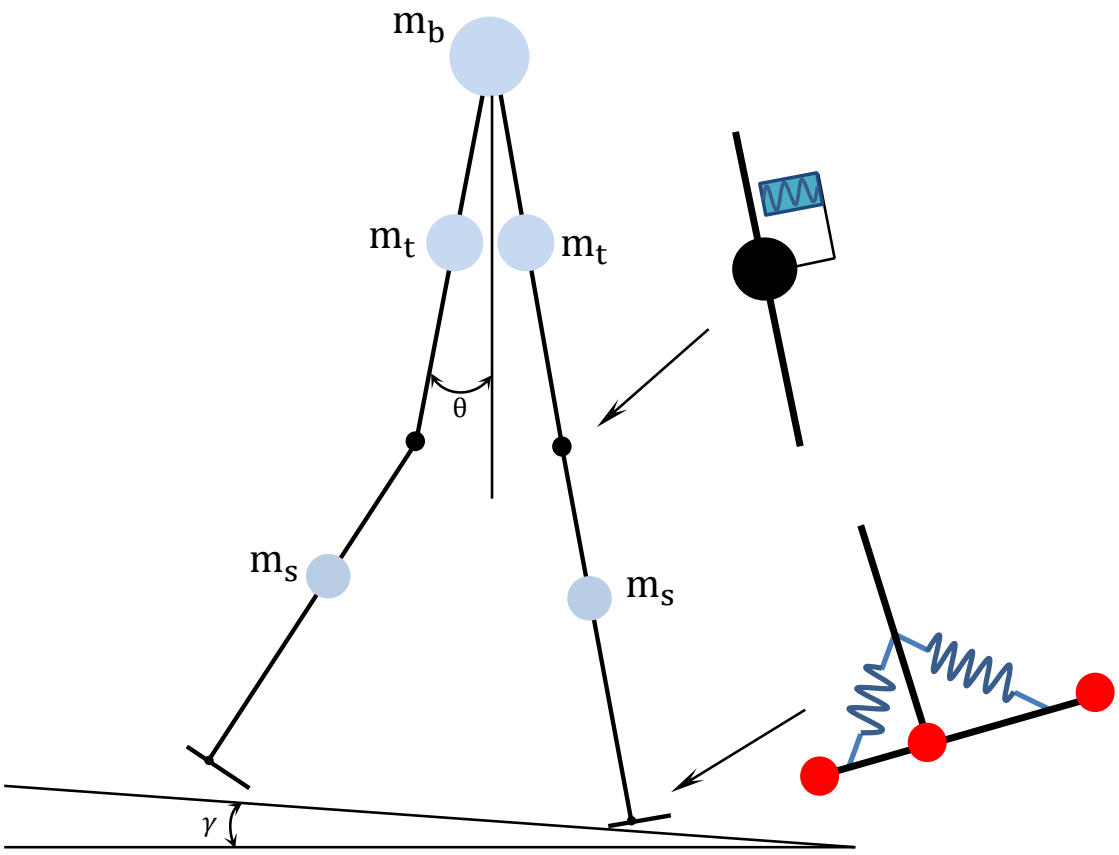


Figure 5.51: passive dynamic robot with surface contact (Adapted from [5])

From Hsu's thesis [5], the dynamic analysis of kneed robot with point contact is implemented, which is also implemented by using different approach in the previous section. In this section, the dynamic analysis of kneed robot with surface contact will be performed. As shown in figure 5.51, three contact points are added to the robot foot, one on the heel, one on the arch, and one on the toe. Two spring-damper systems are placed between the foot and the shank in order to prevent the foot from rotating during the swing stage.

Dynamic analysis of a passive dynamic robot with static friction force

In the fourth numerical experiment for the passive dynamic robot, static friction forces are employed for all the contact points. Same initial conditions and physical parameters are used for the simulation. The screen shots from walking simulation is shown in figure 5.52 as follows:

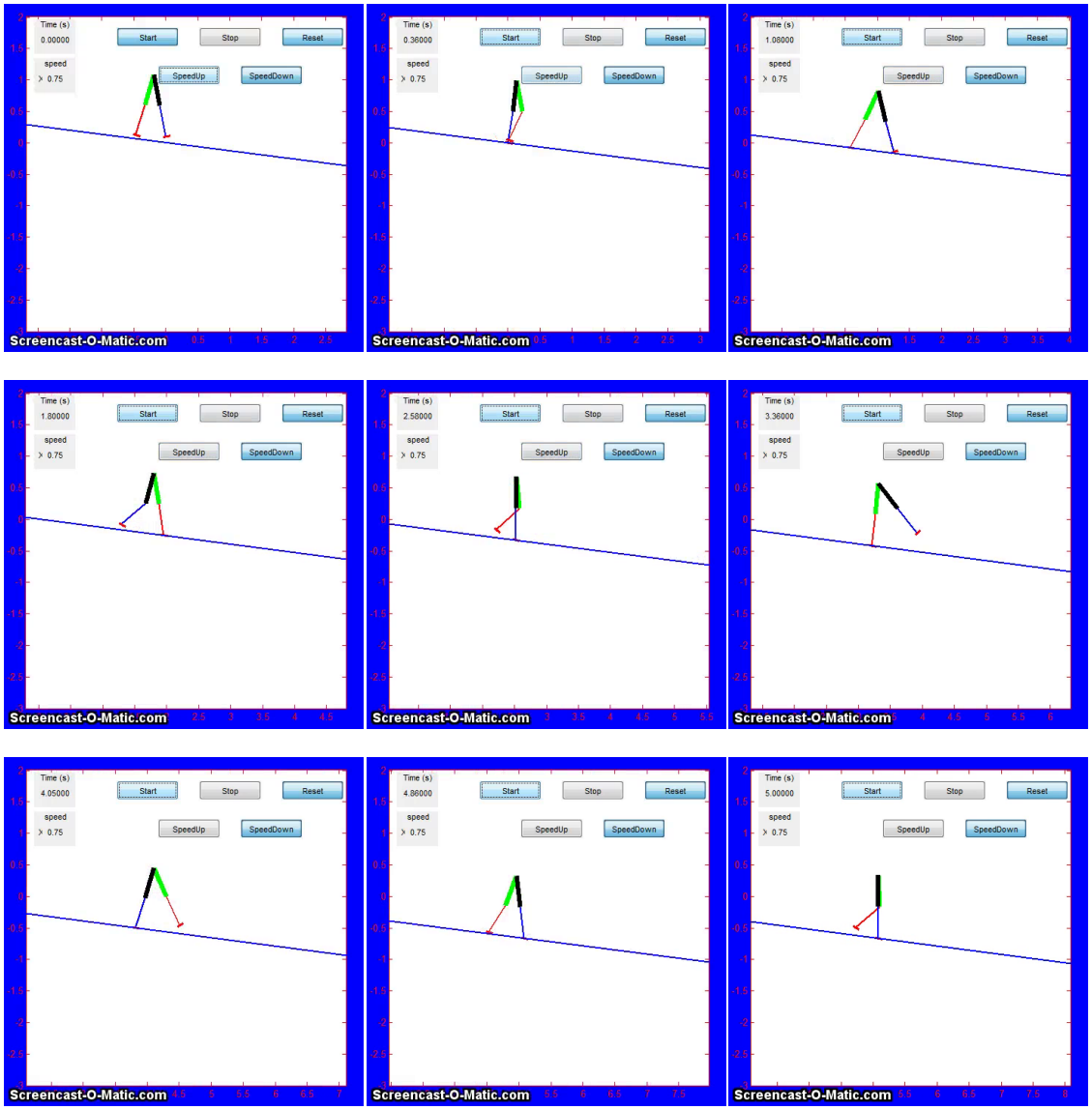
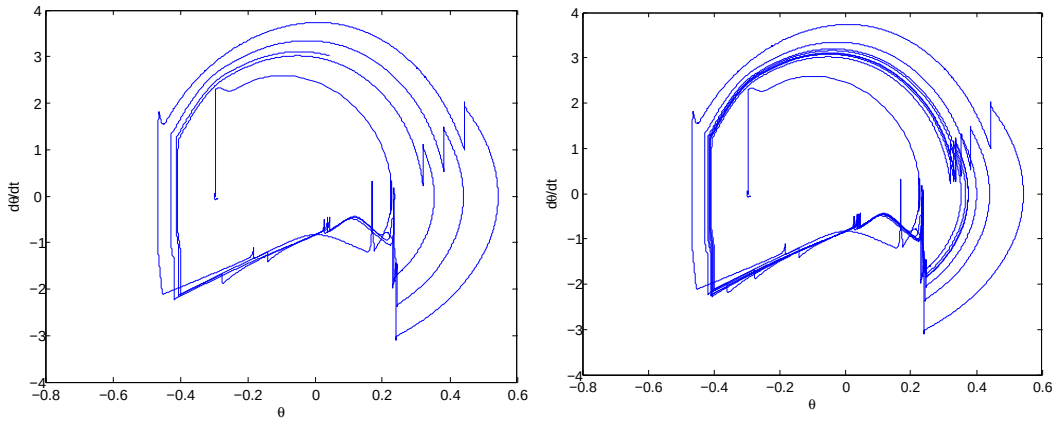


Figure 5.52: Screenshots from dynamic simulation for the fourth experiment



(a) Phase portrait with 5 seconds

(b) Phase portrait with 10 seconds

Figure 5.53: Phase portrait for the upper link of the first leg for the fourth experiment

As you can see, the robot successfully walks for 5 seconds without falling over. However, from the phase portrait shown in figure 5.53a, it's clear to see that the robot simulation doesn't converge to the stable limit cycle in the first 5 seconds. In order to see if the system is able to converge to the stable limit cycle, a simulation with 10 seconds is implemented. As shown in 5.53b, the robot finally converges to the stable limit cycle after 5 seconds. Comparing with the first two simulation experiments, it takes the robot model simulation more time to converge to the stable limit cycle in this fourth experiment.

Dynamic analysis of a passive dynamic robot with dry friction force

In the fifth numerical experiment for the passive dynamic robot with surface contact, Ambrósio friction forces with $v_0 = 1^{-8}$, $v_1 = 1^{-5}$, and $\mu = 0.9$ are employed for all the contact points. The screen shots from walking simulation is shown in figure 5.54 as follows:

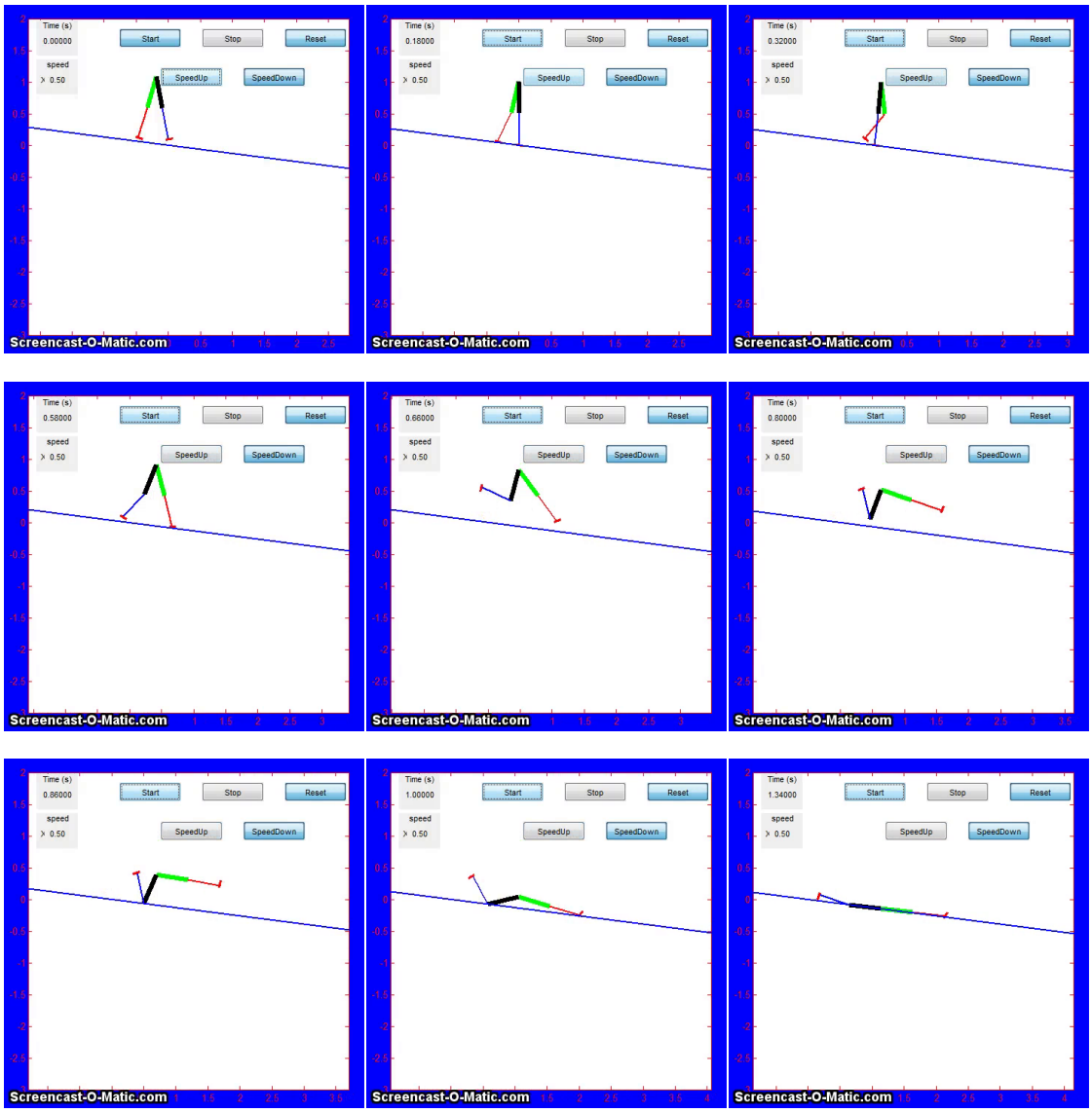


Figure 5.54: Screenshots from dynamic simulation for the fifth experiment

As shown in Figure 5.54, the robot becomes unstable and falls over after one step.

Chapter 6

Multibody dynamics software development

The evolution of multibody dynamics during the last decades makes possible not only to think about the analysis of mechanical systems, but also to develop tools that can help to improve the design of them. There are many commercial packages that can model multibody systems, such as ADAMS, SIMPACK, SimMechanics, LMS VirtualLab Motion, and RecurDyn. One of the drawbacks of these commercial packages is that they don't release the source code to the users. Moreover, commercial multibody packages mainly focus on kinematics and dynamics capabilities, their sensitivity analysis and optimization capabilities are not efficient. There are also some packages developed in academia, but these packages mainly focus on specific application and algorithm.

This study developed a modular multibody package MBSVT (Multibody Systems at Virginia Tech) as a software library with forward kinematics and dynamics, sensitivity analysis, and optimization capabilities. MBSVT is a package for education and research, which allows access to the source code for customization. This software was developed in Fortran 2003 as a collection of Fortran modules and it was tested on several different platforms using multiple compilers. The kinematic library includes dot-1 constraint, revolute, spherical, Euler, and translational joints, as well as distance and coordinates driving constraints. The library implements external forces, such as translational spring-damper-actuator, bump stop, a static friction model, Ambrósio dry friction model, Kelvin-Voigt viscous-elastic model, and a simplified tire model. In MBSVT, Newton's method, explicit and implicit Runge-Kutta method are used to perform kinematic analysis and dynamic analysis respectively. Moreover, the adjoint variable method based on the penalty formulation is employed to calculate the sensitivities. Furthermore, with the outputs of sensitivity analysis, MBSVT uses L-BFGS-B [101], a very popular optimization package, to perform gradient-based optimization. MBSVT also provides a connection with Matlab by means of the Matlab

engine.

To show the functionality of the library, the application of MBSVT to a full vehicle and a passive dynamic robot are discussed in section 5.2 and section 5.3 respectively.

6.1 The MBSVT features

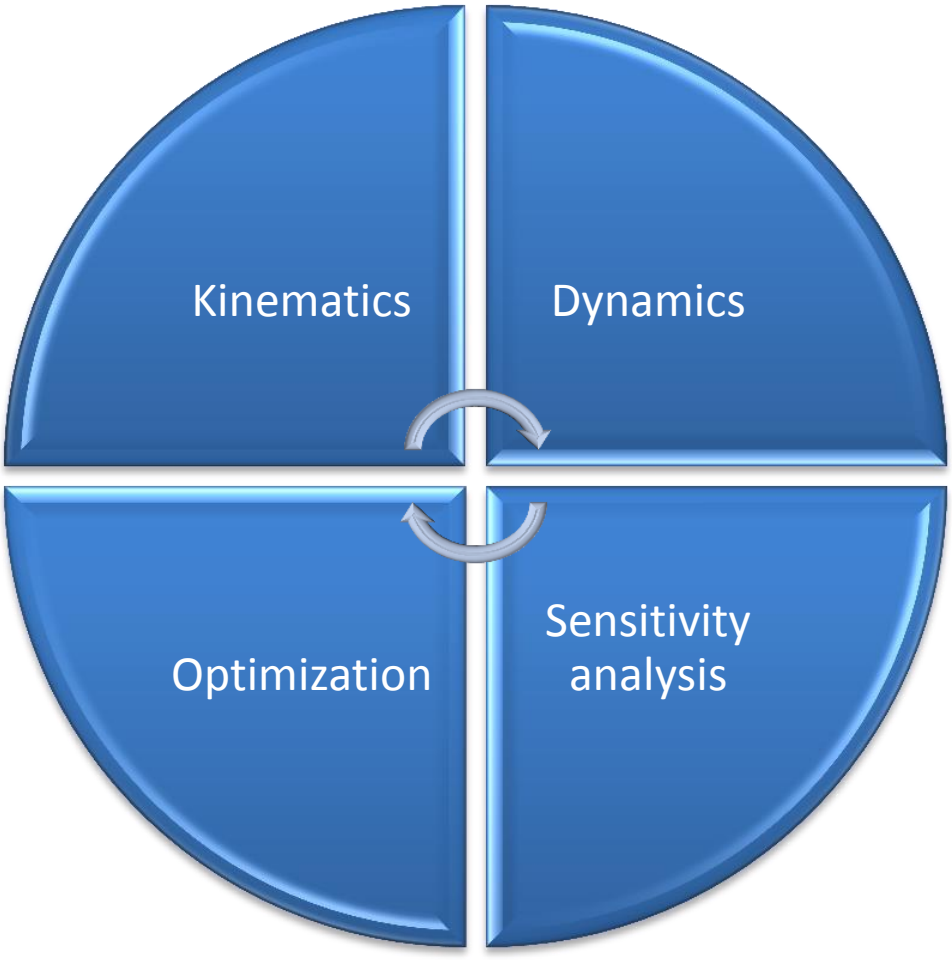


Figure 6.1: MBSVT features

As shown in figure 6.1, the main features of MBSVT include kinematics, dynamics, sensitivity analysis, and optimization. Newton's method and Runge-Kutta method are used for the kinematic analysis and dynamic analysis respectively. The adjoint variable method based on the penalty formulation and L-BFGS-B are used to perform sensitivity analysis and optimization respectively.

Kinematics

- Position equation

$$\mathbf{\Phi}(\mathbf{q}, t) = \begin{bmatrix} \mathbf{\Phi}^K(\mathbf{q}) \\ \mathbf{\Phi}^D(\mathbf{q}, t) \end{bmatrix} = \mathbf{0}. \tag{6.1}$$

- Velocity equation

$$\mathbf{\Phi}_\mathbf{q}(\mathbf{q}, t)\dot{\mathbf{q}} = -\mathbf{\Phi}_t. \tag{6.2}$$

- Acceleration equation

$$\mathbf{\Phi}_\mathbf{q}(\mathbf{q}, t)\ddot{\mathbf{q}} = -\dot{\mathbf{\Phi}}_t - \dot{\mathbf{\Phi}}_\mathbf{q}\dot{\mathbf{q}}. \tag{6.3}$$

The main task of kinematic analysis is to solve position, velocity, and acceleration equations.

As shown in equation 6.1, the position equation is composed of kinematic constraints $\mathbf{\Phi}^K(\mathbf{q})$ and driving constraints $\mathbf{\Phi}^D(\mathbf{q}, t)$. The velocity equation 6.2 is obtained by differentiating the position equation while the acceleration equation 6.3 is obtained by differentiating the velocity equation. To solve these equations, one only needs to simply employs newton's method or other similar methods for solving nonlinear system.

Dynamics

MBSVT employs the penalty formulation as the EOM for dynamic analysis. For more details about the penalty formulation, the reader is referred to section 2.4. Comparing with other formulations, there are several advantages to use the penalty formulation. First, the penalty formulation is more stable, and it doesn't fail around kinematic singularity. Second, the penalty formulation allows redundant constraints. Third, the penalty formulation is an ODE-like formulation, it is more

computationally efficient to be solved than DAE formulations. Fourth, unlike Maggi's formulation, the penalty formulation doesn't need to restart the numerical integrator for each time step. Based on these advantages, the penalty formulation is chosen to be the dynamic equation for MBSVT. 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. To solve the EOM, explicit Runge-Kutta method and implicit Runge-Kutta method are employed in MBSVT.

Sensitivity analysis

In order to perform sensitivity analysis for large and complex multibody systems with a large number of parameters, the adjoint variable method based on the penalty formulation is used in MBSVT. The advantages of the penalty formulation has already been described in the previous section. The reason to choose the adjoint variable method instead of the direct differentiation method is the adjoint variable method works well when the number of design parameters is large while the direct differentiation method doesn't work well when the number of design parameters is large. For more details about the adjoint variable method based on the penalty formulation, the reader is referred to section 3.2.3.

Optimization

With the outputs of the sensitivity analysis, many gradient-based optimization packages can be used to perform the dynamic response optimization. L-BFGS-B, a third party gradient-based optimization package, is used with MBSVT to perform optimization.

$$\begin{aligned} &\text{Minimize } f(\mathbf{P}) \\ &\text{Subject to } \mathrm{l_i} \leq \mathrm{p_i} \leq \mathrm{u_i} \ \forall \mathrm{i} \end{aligned} \tag{6.4}$$

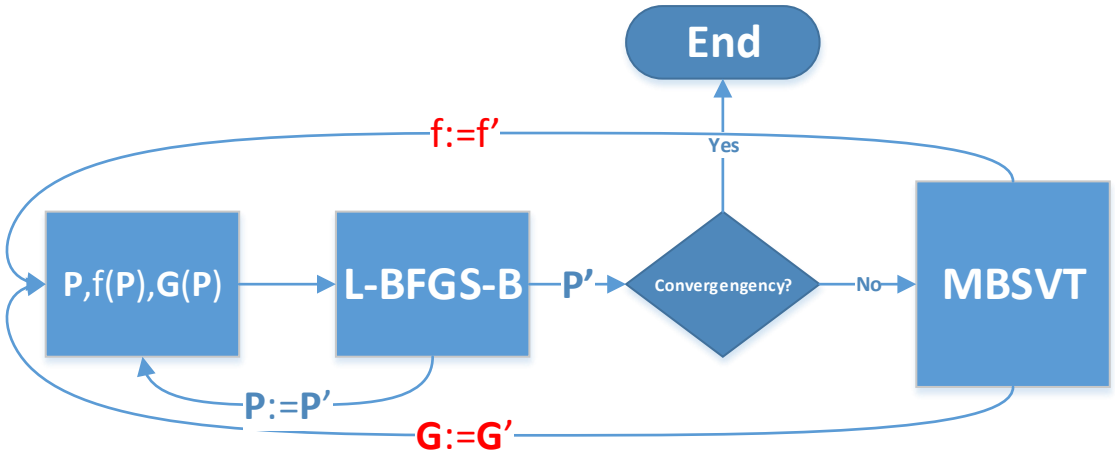


Figure 6.2: L-BFGS-B flowchart

As shown in equation 6.4, L-BFGS-B is able to perform optimization with bounded constraints, when $f(\mathbf{q})$ is the cost function, $\mathbf{q}$ is the set of design parameters.

Figure 6.2 shows the functionality of L-BFGS-B, where $\mathbf{P}$ is the set of parameters, f is the value of cost function, $\mathbf{G}$ is the set of sensitivities. The same initial condition of state vector $\mathbf{Y}$ is used for each iteration. At the beginning, with the initial guess of the set of parameters $\mathbf{P}$, MBSVT computes the value of cost function f and sensitivities $\mathbf{G}$. $\mathbf{P}$, f, and $\mathbf{G}$ are then sent to L-BFGS-B to calculate a new set of parameters $\mathbf{P}'$. If the set of parameters converges at the current iteration, then the process is ended, otherwise $\mathbf{P}'$ is sent to MBSVT to obtain the new value of cost function f' and sensitivities $\mathbf{G}'$ that are then sent to L-BFGS-B along with $\mathbf{P}'$ to start a new iteration.

6.2 The MBSVT algorithm

The MBSVT is a modular library that can be linked to user's projects. The algorithm , as shown in Fig.6.3, follows through the primary steps of model set-up, integrator set-up, and results output .

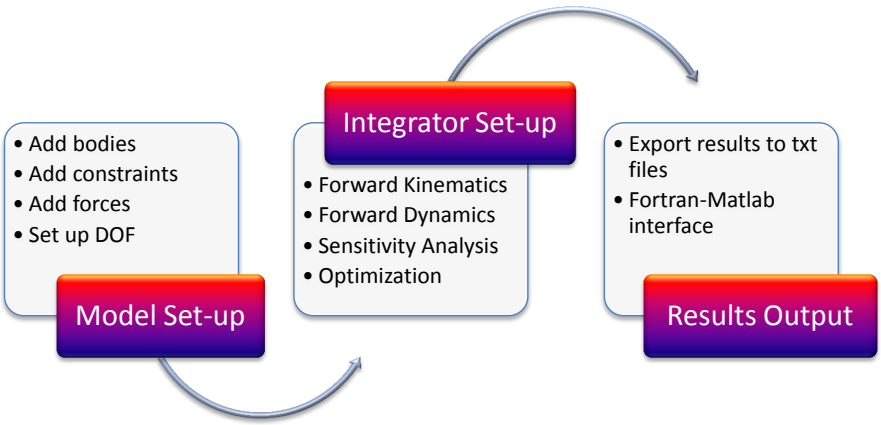


Figure 6.3: MBSVT Algorithm flowchart

First, by simply adding bodies from the user's interface, all the inertia properties are automatically saved. Second, constraints and forces are added to the system. After adding constraints and forces, If the number of independent constraints, m, is smaller than the number of coordinates, n_c, then DOF, $d = n_c - m$, must be chosen and initialized in order to solve the initial position and velocity of the mechanism in function of the DOF. Once the mechanism has been described, MBSVT has different functions and integrators to perform the kinematics, forward dynamics, sensitivity analysis, or optimization. Finally, MBSVT provides a connection with Matlab by means of the Matlab engine in order to visually output the results.

6.3 The MBSVT structure

As a modular package, MBSVT is made up of several modules, as shown in table 6.1 with the structure shown in figure 6.4.

As shown in figure 6.4, the MBSVT is composed of five main parts. The first part contains all of the inertia terms that include mass, moment of inertia, and all the inertia related computational terms.

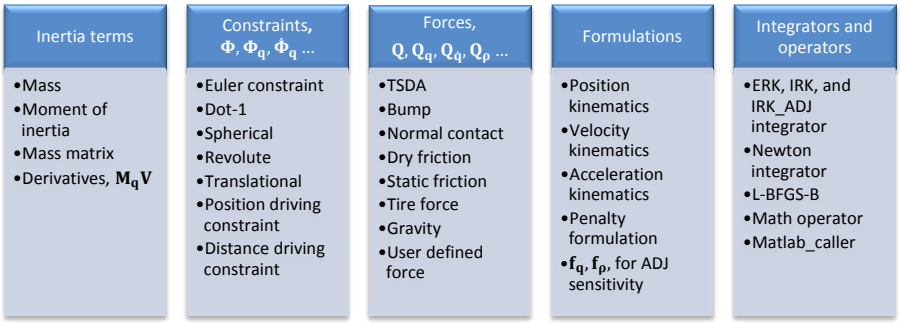


Figure 6.4: MBSVT Modular Structure

Unlike kinematic analysis and dynamic analysis, the computation of sensitivities requires the differentiation of cost function and the EOM, which generates many derivatives. Thus, the second part includes all the expressions of these derivatives, which include: $\mathbf{\Phi}$, $\mathbf{\Phi_q}$, $\dot{\mathbf{\Phi}}_\mathbf{q}$, $\ddot{\mathbf{\Phi}}_\mathbf{q}$, $\dot{\mathbf{\Phi}}_\mathbf{q}\dot{\mathbf{q}}$, $\dot{\mathbf{\Phi}}_\mathbf{q}\dot{\mathbf{q}}_\mathbf{q}$, $\mathbf{\Phi_{qq}V}$, and $\mathbf{\Phi}^{\mathrm{T}}_{\mathbf{qq}}\mathbf{V}$. Manual differentiation is used to compute the expressions of these derivatives, for more details about manual differentiation, reader is referred to section 1.4

The third part contains several kinds of external forces and tire models, such as translational-spring-damper-actuator force, bump stop force, static friction force, Ambrósio dry friction force, Kelvin-Voigt viscous-elastic force, and a simplified tire model. For those forces not defined in MBSVT, callbacks are used for users to provide user defined forces. Gravity is automatically considered by the library.

The forth part includes all the generic functions to construct different formulations corresponding to different modes such as kinematics, dynamics, sensitivity analysis, and optimization.

The last part includes all the integrators, operators, and interfaces to perform kinematic analysis, dynamic analysis, sensitivity analysis, and optimization and to visually plot the results. These integrators and operators includes newton's integrator for kinematics, explicit and implicit Runge-Kutta integrator for forward dynamics, implicit Runge-Kutta adjoint integrator for adjoint sensitivity analysis, L-BFGS-B for gradient optimization, and a math_operator module that includes some important user-defined math operators. After obtaining the results, MATLAB_CALLER provides a connection with Matlab by means of the Matlab engine in order to visually output the results.

Table 6.1: Module list

CONSTANTS	Module of solver parameters
CONSTRAINTS	Module that manages the constraints
d2Jacobdt2	Module of $\ddot{\mathbf{\Phi}}_{\mathbf{q}}$
DERIVED_TYPES	Module of solver derived types
dJacobdt	Module of $\dot{\mathbf{\Phi}}_{\mathbf{q}}$
djacobdt_qp	Module of $\dot{\mathbf{\Phi}}_{\mathbf{q}}\dot{\mathbf{q}}$
formulation_Dynamics	Dynamic simulation module
formulation_Kinematics	Kinematic simulation module
formulation_Sensitivity	Sensitivity analyis module
formulations	Module contain the generic functions that manage the use of different formulations
generalized_forces	Generalized forces module
Jacob	Module of $\mathbf{\Phi}_{\mathbf{q}}$
jacob_djacobdt_qp	Module of $(\dot{\mathbf{\Phi}}_{\mathbf{q}}\dot{\mathbf{q}})_{\mathbf{q}}$
jacob_jacob	Module of $\mathbf{\Phi}_{\mathbf{qq}}\mathbf{V}$, which is the jacobian of the primitive jacobian multiplied by a vector
jacobT_jacob	Module of $\mathbf{\Phi}^{\mathrm{T}}_{\mathbf{qq}}\mathbf{V}$, which is the transpose of the jacobian of the primitive jacobian multiplied by a vector
Mass_Massq	Module of $\mathbf{M}_{\mathbf{q}}\mathbf{V}$, which is the jacobian of the mass matrix multiplied by a vector
math_oper	Module of mathematical operations for multibody dynamics computations not supported by the Fortran 2003 standard
matlab_caller	Managment of sessions of MATLAB engine
primitive_forces	Primitive forces module
restric	Module of primitive constraints
SOLIDS	Solids module that adds and manages the bodies of the system
STATE	Module of solver state variables, subroutines and functions. It creates, manages and updates the state variables of the model

Chapter 7

Conclusion

In this study, we bring new contributions to the state-of-the-art in analytical approaches to perform sensitivity analysis of multibody systems. The direct differentiation method and the adjoint variable method are developed in the context of the penalty formulation and Maggi's formulation. The resulting sensitivities are applied to perform dynamical optimization of different multibody systems. The collection of bench-mark problems includes a five-bar mechanism and a full vehicle model. In addition, the dynamic simulation of a passive dynamic walker is implemented by using reference point coordinates and the penalty formulation. Finally, a new multibody dynamics software library MBSVT (Multibody Systems at Virginia Tech) is developed in Fortran 2003, with forward kinematics and dynamics, sensitivity analysis, and optimization capabilities.

7.1 Contributions

The main contributions of this study can be summarized as follows:

- Reviewed and summarized most of the current analytical sensitivity approaches based on the direct differentiation method and the adjoint variable method.
- Brought new contributions to the state-of-the-art in analytical approaches to perform sensitivity analysis of multibody systems. The new analytical sensitivity approaches developed in this study include:
  - The direct differentiation method based on explicit first order Maggi's formulation
  - The adjoint variable method based on explicit first order Maggi's formulation
  - The adjoint variable method based on implicit first order Maggi's formulation
  - The adjoint variable method based on implicit second order Maggi's formulation

– The adjoint variable method based on explicit first order penalty formulation

– The adjoint variable method based on implicit second order penalty formulation

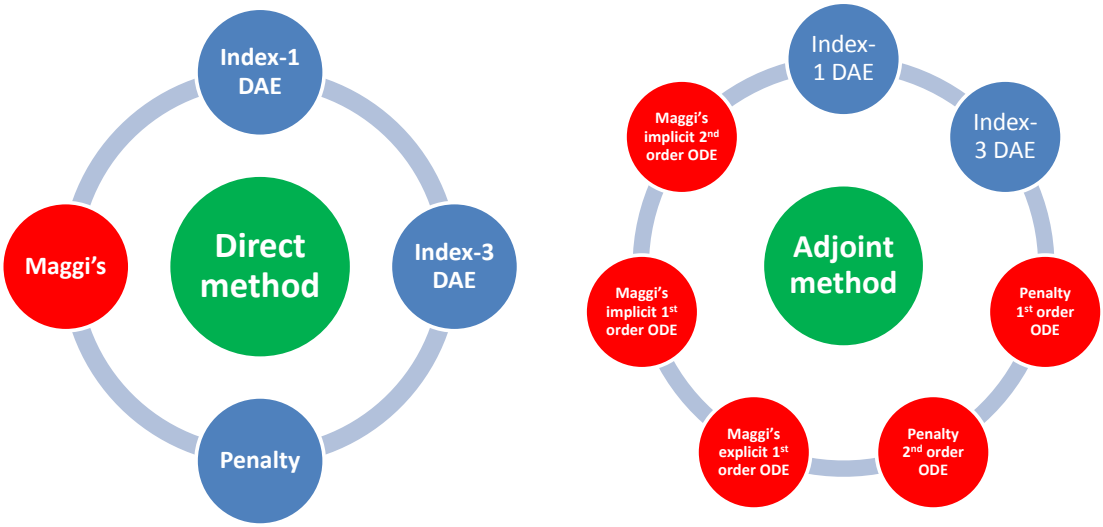


Figure 7.1: Main contributions

The direct differentiation method and the adjoint variable method in the context of several different multibody dynamics formulations are summarized in figure 7.1. The approaches developed in this study are highlighted with red color.

- Validated and tested these new analytical sensitivity approaches by comparing them with other analytical and numerical approaches by using a five-bar mechanism as a case study.

- Demonstrated the capability of the adjoint variable method based on explicit first order penalty formulation to perform sensitivity analysis and gradient-based optimization for large and complex multibody systems with respect to multiple design parameters by perform handling optimization and ride optimization for a complex full vehicle model.

- Developed an open source software package MBSVT, which has forward kinematics and dynamics, sensitivity analysis, and optimization capabilities, for research and education of multibody dynamics and vehicle dynamics, and make it available to the scientific community. Complex finite difference method is used to validate all the derivatives before putting them

in the library, which guarantees the computational accuracy of MBSVT. Checked and tested MBSVT in several different platforms, compilers, and operating systems, such as gfortran, gcc, intel fortran, Windows, and Linux. Furthermore, several different contact and friction models, which can be used to model point contact and surface contact, are developed and included in MBSVT. These models include a static friction model, Ambrósio dry friction model, Kelvin-Voigt viscous-elastic model, and a simplified tire model.

- Finally, unlike those traditional methods to model robotic systems that employ relative coordinates and Lagrange formulation, reference point coordinates and the penalty formulation are employed to perform dynamic analysis for robotic systems in this study. By doing this, the robotic systems become more stable when system goes through a singular or a bifurcation position. In addition, it is easier to model and to write the equations with reference point coordinates than with relative coordinates, simplifying the modeling process. Furthermore, the passive dynamic robot is also used to test and validate all the point contact and surface contact models developed in MBSVT, which is the software to be used to perform the dynamic analysis for the passive dynamic robot.

7.2 Future work

Generally, the future work includes the sensitivity analysis and optimization of multibody systems with contact, friction, and uncertainty. In addition, the comparison of the efficiency and accuracy of between automatic differentiation, manual differentiation and other methods is the next direction of future research. Furthermore, the MBSVT package should continue to be developed with new algorithms and formulations for multibody systems with with contact, friction, and uncertainty. Finally, to develop a visualization package for MBSVT is necessary if one wants to commercialize MBSVT. Thus, another task for future work is to develop a visualization package for MBSVT.

Specifically, the future work includes the following topics:

- In this study, the dynamic analysis for a passive dynamic robot has been done, but the

sensitivity analysis and optimization of the performance of the passive dynamic robot has not been studied in this study. Thus, the first job to do in the future is to perform sensitivity analysis and optimization of the passive dynamic robot. The optimization of the performance of passive dynamic robot has been studies in the last 10 years, such as the optimization of energy consumption, walking speed, and stability. However, these approaches contain specific algorithms, which limit the scope of these approaches to specific applications. Currently there is no general-purpose method to optimize the performance of passive dynamic robot. Thus, it's interesting to see if these new approaches developed in this study are capable to optimize the performance of passive dynamic robot.

- Uncertainty comes from different sources, such as initial conditions, system parameters, and external perturbation and noise. Actually, almost all the real-life multibody systems are affected by uncertainty. Thus, the treatment of uncertainty becomes very important. The kinematic analysis and dynamic analysis for multibody systems with uncertainty has already been studied by many people. Therefore, the second job to do in the future is to perform sensitivity analysis and optimization of multibody systems with uncertainty.

- Since the computation of sensitivities requires the differentiation of cost function and the EOM, which generates many derivatives, a efficient and general-purpose method is required to perform sensitivity analysis. As what's introduced in section 1.4, automatic differentiation and manual differentiation are better than symbolic computational software packages and numerical differentiation. However, a more specific comparison of efficiency, accuracy, generality, and development time between these different methods has not been made yet. Thus, the third job to do in the future is to compare efficiency, accuracy, generality, and development between automatic differentiation method, manual differentiation method and other methods.

- MBSVT is developed by Fortran 2003 in order to increase the simulation speed. By doing this, some characteristics are sacrificed, such as the visualization and animation characteristics. Although MBSVT has powerful functional elements, a good visualization package should be developed and included in MBSVT if one wants to commercialize MBSVT . Currently, MBSVT has a Fortran-Matlab interface that is able to start the matlab engine and plot the

results. However, to visually simulate miscellaneous multibody systems, a better visualization package is required. Thus the fourth job to do in the future is to develop a visualization package for MBSVT.

Bibliography

[1] Frik, S., Leister, G., and Schwartz, W., 1993. "Simulation of the IAVSD road vehicle benchmark bombardier iltis with FASIM, MEDYNA, NEWEUL and SIMPACK". Vehicle System Dynamics, **22**(sup1), pp. 215–253.

[2] Callejo Goena, A., 2013. "Dynamic response optimization of vehicles through efficient multibody formulations and automatic differentiation techniques". PhD thesis, Industriales.

[3] Dopico, D., Luaces, A., Gonzalez, M., and Cuadrado, J., 2011. "Dealing with multiple contacts in a human-in-the-loop application". Multibody System Dynamics, **25**(2), pp. 167–183.

[4] Flores, P., Ambrósio, J., Claro, J. P., and Lankarani, H. M., 2008. "Contact-impact force models for mechanical systems". In Kinematics and Dynamics of Multibody Systems with Imperfect Joints. Springer, pp. 47–66.

[5] Chen, H., et al., 2007. "Passive dynamic walking with knees: A point foot model". PhD thesis, Massachusetts Institute of Technology.

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

[7] Haug, E., and Arora, J., 1978. "Design sensitivity analysis of elastic mechanical systems". Computer Methods in Applied Mechanics and Engineering, **15**, pp. 35–62.

[8] Haug, E.J., W. R., and Mani, N., 1984. "Design sensitivity analysis of largescale constrained dynamic mechanical systems". ASME Journal of Mechanisms, Transmissions, and Automation in Design, **106**, pp. 156–162.

[9] Krishnaswami, P., and Bhatti, M., 1984. "A general approach for design sensitivity analysis of constrained dynamic systems". ASME Journal of Mechanisms, Transmissions, and Automation in Design, pp. 84–DET–132.

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

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

[12] de Jalón, J. G., and Serna, M., 1981. "Computer method for kinematic analysis of lower-pair mechanisms- i. velocities and accelerations.". MECH. & MACH. THEORY., **16**(5), pp. 543–556.

[13] de Jalón, J. G., Serna, M. A., and Avilés, R., 1981. "Computer method for kinematic analysis of lower-pair mechanismsii position problems". Mechanism and Machine Theory, **16**(5), pp. 557–566.

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

[15] Jalón, D., García, J., Unda, J., and Avello, A., 1986. "Natural coordinates for the

computer analysis of multibody systems". Computer Methods in Applied Mechanics and Engineering, **56**(3), pp. 309–327.
[16] de Jalon, J. G., Unda, J., Avello, A., and Jimenez, J., 1987. "Dynamic analysis of three-dimensional mechanisms in natural coordinates". Journal of Mechanical Design, **109**(4), pp. 460–465.
[17] Garcia de Jalon, J., and Bayo, E., 1994. Kinematic and dynamic simulation of multibody systems: The real-time challenge. Springer-Verlag, New York (USA).
[18] Serban, R., Negrut, D., Haug, E. J., and Potra, F. A., 1997. "A topology-based approach for exploiting sparsity in multibody dynamics in cartesian formulation*". Journal of Structural Mechanics, **25**(3), pp. 379–396.
[19] Eich-Soellner, E., and Führer, C., 1998. Numerical Methods in Multibody Dynamics. B.G.Teubner Stuttgart.
[20] Baumgarte, J., 1972. "Stabilization of constraints and integrals of motion in dynamical systems". Computer methods in applied mechanics and engineering, **1**(1), pp. 1–16.
[21] Bayo, E., Garcia de Jalon, Ja de Jalon, J., and Serna, M., 1988. "A modified lagrangian formulation for the dynamic analysis of constrained mechanical systems". Computer Methods in Applied Mechanics and Engineering, **71**(2), 11, pp. 183–195.
[22] Jalon, J. G. d., and Bayo, E., 1994. "Kinematic and dynamic simulation of multibody systems: the real time challenge".
[23] Shabana, A. A., 1997. "Flexible multibody dynamics: review of past and recent developments". Multibody system dynamics, **1**(2), pp. 189–222.
[24] Ambrósio, J. A., and Pereira, M. S., 1994. "Flexibility in multibody dynamics with applications to crashworthiness". In Computer-aided analysis of rigid and flexible mechanical systems. Springer, pp. 199–232.
[25] Bakr, E., and Shabana, A., 1986. "Geometrically nonlinear analysis of multibody systems". Computers & structures, **23**(6), pp. 739–751.
[26] Bakr, E., and Shabana, A., 1987. "Timoshenko beams and flexible multibody system dynamics". Journal of sound and vibration, **116**(1), pp. 89–107.
[27] Book, W. J., 1979. "Analysis of massless elastic chains with servo controlled joints". Journal of Dynamic Systems, Measurement, and Control, **101**(3), pp. 187–192.
[28] Boyer, F., Khalil, W., Benosman, M., and Le Vey, G., 2007. "Modeling and control of flexible robots". Modeling, Performance Analysis and Control of Robot Manipulators, pp. 337–394.
[29] Chu, S.-C., and Pan, K., 1975. "Dynamic response of a high-speed slider-crank mechanism with an elastic connecting rod". Journal of Manufacturing Science and Engineering, **97**(2), pp. 542–550.
[30] Erdman, A., and Sandor, G., 1972. "Kineto-elastodynamicsa review of the state of the art and trends". Mechanism and Machine Theory, **7**(1), pp. 19–33.
[31] Huston, R. L., 1991. "Computer methods in flexible multibody dynamics". International journal for numerical methods in engineering, **32**(8), pp. 1657–1668.
[32] Lowen, G., and Chassapis, C., 1986. "The elastic behavior of linkages: An update". Mechanism and Machine Theory, **21**(1), pp. 33–42.

[33] Lowen, G., and Jandrasits, W., 1972. "Survey of investigations into the dynamic behavior of mechanisms containing links with distributed mass and elasticity". Mechanism and Machine theory, **7**(1), pp. 3–17.

[34] Roberson, R., 1972. "A form of the translational dynamical equations for relative motion in systems of many non-rigid bodies". Acta Mechanica, **14**(4), pp. 297–308.

[35] Schiehlen, W., and Rauh, J., 1986. "Modeling of flexible multibeam systems by rigid-elastic superelements". Revista Brasiliera de Ciencias Mecanicas, **8**(2), pp. 151–163.

[36] Shabana, A. A., and Wehage, R. A., 1983. "A coordinate reduction technique for dynamic analysis of spatial substructures with large angular rotations". Journal of Structural Mechanics, **11**(3), pp. 401–431.

[37] Song, J. O., and Haug, E. J., 1980. "Dynamic analysis of planar flexible mechanisms". Computer Methods in Applied Mechanics and Engineering, **24**(3), pp. 359–381.

[38] Maly, T., and Petzold, L. R., 1996. "Numerical methods and software for sensitivity analysis of differential-algebraic systems". Applied Numerical Mathematics, **20**(1), pp. 57–79.

[39] Sohoni, V., and Haug, E., 1982. "A state space technique for optimal design of mechanisms". Journal of Mechanical Design, **104**(4), pp. 792–798.

[40] Chang, C., and Nikravesh, P., 1985. "Optimal design of mechanical systems with constraint violation stabilization method". Journal of Mechanisms, Transmissions and Automation in Design, **107**(4), pp. 493–498.

[41] Haug, E., 1987. Computer aided optimal design : structural and mechanical systems. No. 27 in NATO ASI series. Series F, Computer and systems sciences. Springer-Verlag, ch. Design sensitivity analysis of dynamic systems.

[42] Haug, E., Wehage, R., and Barman, N., 1981. "Design sensitivity analysis of planar mechanism and machine dynamics". Journal of Mechanical Design, **103**(3), pp. 560–570.

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

[44] Ashrafiuon, H., and Mani, N., 1990. "Analysis and optimal design of spatial mechanical systems". Journal of Mechanical Design, **112**(2), pp. 200–207.

[45] Liu, X., 1996. "Sensitivity analysis of constrained flexible multibody systems with stability considerations". Mechanism and machine theory, **31**(7), pp. 859–863.

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

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

[48] Wang, X., Haug, E., and Pan, W., 2005. "Implicit numerical integration for design sensitivity analysis of rigid multibody systems". Mechanics based design of structures and machines, **33**(1), pp. 1–30.

[49] Schaffer, A. S., 2005. "On the adjoint formulation of design sensitivity analysis of multibody dynamics". PhD thesis, University of Iowa, December.

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

[51] Sonneville, V., and Brüls, O., 2014. "Sensitivity analysis for multibody systems formulated on a lie group". Multibody System Dynamics, **31**(1), pp. 47–67.

[52] Bestle, D., and Eberhard, P., 1992. "Analyzing and optimizing multibody systems". Mechanics of Structures and Machines, **20**(1), pp. 67–92.

[53] Feehery, W. F., Tolsma, J. E., and Barton, P. I., 1997. "Efficient sensitivity analysis of large-scale differential-algebraic systems". Applied Numerical Mathematics, **25**(1), pp. 41 – 54.

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

[55] ANDERSSON, D., and ERIKSSON, P., 2004. "Handling and ride comfort optimisation of an intercity bus". Vehicle system dynamics, **41**, pp. 547–556.

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

[57] Neto, M. A., Ambrosio, J.osio, J. A. C., and Leal, R. P., 2009. "Sensitivity analysis of flexible multibody systems using composite materials components". International Journal for Numerical Methods in Engineering, **77**(3), pp. 386–413.

[58] Bhalerao, K., Poursina, M., and Anderson, K., 2010. "An efficient direct differentiation approach for sensitivity analysis of flexible multibody systems". Multibody System Dynamics, **23**, pp. 121–140. 10.1007/s11044-009-9176-0.

[59] Banerjee, J. M., and McPhee, J., 2013. Multibody Dynamics. Computational methods and applications., Vol. 28 of Computational Methods in Applied Sciences. Springer, ch. Symbolic Sensitivity Analysis of Multibody Systems, pp. 123–146.

[60] Adelman, H. M., and Haftka, R. T., 1986. "Sensitivity analysis of discrete structural systems". AIAA journal, **24**(5), pp. 823–832.

[61] Wasfy, T. M., and Noor, A. K., 1996. "Modeling and sensitivity analysis of multibody systems using new solid, shell and beam elements". Computer Methods in Applied Mechanics and Engineering, **138**(1), pp. 187–211.

[62] Kim, N. H., Park, Y. H., and Choi, K. K., 2001. "Optimization of a hyper-elastic structure with multibody contact using continuum-based shape design sensitivity analysis". Structural and multidisciplinary optimization, **21**(3), pp. 196–208.

[63] Etman, L. F. P., Van Campen, D., and Schoofs, A., 1996. "Optimization of multibody systems using approximation concepts". In IUTAM Symposium on Optimization of Mechanical Systems, Springer, pp. 81–88.

[64] Choi, K. K., and Kim, N.-H., 2006. Structural sensitivity analysis and optimization 1: linear systems, Vol. 1. Springer.

[65] Xiang, Y., Arora, J. S., and Abdel-Malek, K., 2009. "Optimization-based motion prediction of mechanical systems: sensitivity analysis". Structural and Multidisciplinary Optimization, **37**(6), pp. 595–608.

[66] Eberhard, P., Schiehlen, W., and Sierts, J., 2007. "Sensitivity analysis of inertia parameters in multibody dynamics simulations". In Proceedings of the 12th World Congress in Mechanism and Machine Science, Besançon, French IFToMM Committee, Vol. 4, pp. 101–106.

[67] Mukherjee, R. M., Bhalerao, K. D., and Anderson, K. S., 2008. "A divide-and-conquer direct differentiation approach for multibody system sensitivity analysis". Structural and Multidisciplinary Optimization, **35**(5), pp. 413–429.

[68] Eberhard, P., 1996. "Adjoint variable method for sensitivity analysis of multibody systems interpreted as a continuous, hybrid form of automatic differentiation". In Proceedings of the Second International Workshop of Computational Differentiation, Santa Fe, NM, pp. 12–14.

[69] Bhalerao, K. D., Poursina, M., and Anderson, K. S., 2010. "An efficient direct differentiation approach for sensitivity analysis of flexible multibody systems". Multibody System Dynamics, **23**(2), pp. 121–140.

[70] Serban, R., et al., 1998. Dynamic and sensitivity analysis of multibody systems. The University of Iowa.

[71] Fan, K.-C., Wang, H., Zhao, J.-W., and Chang, T.-H., 2003. "Sensitivity analysis of the 3-PRS parallel kinematic spindle platform of a serial-parallel machine tool". International Journal of Machine Tools and Manufacture, **43**(15), pp. 1561–1569.

[72] Ma, Z.-D., and Hagiwara, I., 1991. "Sensitivity analysis methods for coupled acoustic-structural systems part i: modal sensitivities". AIAA journal, **29**(11), pp. 1787–1795.

[73] Pollock, G., and Noor, A., 1996. "Sensitivity analysis of the contact/impact response of composite structures". Computers & structures, **61**(2), pp. 251–269.

[74] Bestle, D. Analyse und optimierung von mehrkörpersystemen, 1994.

[75] Pagalday, J., 1994. "Optimización del comportamiento dinámico de mecanismos". PhD thesis, Tesis Doctoral, Universidad de Navarra.

[76] Serban, R., and Freeman, J., 1996. "Direct differentiation methods for the design sensitivity of multi-body dynamic systems". In Proceedings of the 1996 ASME Design Engineering Technical Conferences and Computers in Engineering Conference, pp. 18–22.

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

[78] Chatillon, M., Jezequel, L., Coutant, P., and Baggio, P., 2006. "Hierarchical optimisation of the design parameters of a vehicle suspension system". Vehicle System Dynamics, **44**(11), pp. 817–839.

[79] "http://www.neos-server.org/".

[80] Besselink, I., and Van Asperen, F., 1994. "Numerical optimization of the linear dynamic behaviour of commercial vehicles". Vehicle System Dynamics, **23**(1), pp. 53–70.

[81] Eberhard, P., Schiehlen, W., and Bestle, D., 1999. "Some advantages of stochastic methods in multicriteria optimization of multibody systems". Archive of Applied Mechanics, **69**(8), pp. 543–554.

[82] Goncalves, J. P., and Ambrosio, J. A., 2005. "Road vehicle modeling requirements for optimization of ride and handling". Multibody System Dynamics, **13**(1), pp. 3–23.
[83] Anitescu, M., 2006. "Optimization-based simulation of nonsmooth rigid multibody dynamics". Mathematical Programming, **105**(1), pp. 113–143.
[84] Gonçalves, J. P., and Ambrósio, J. A., 2003. "Optimization of vehicle suspension systems for improved comfort of road vehicles using flexible multibody dynamics". Nonlinear Dynamics, **34**(1-2), pp. 113–131.
[85] Pagalday, J., Aranburu, I., Avello, A., and de Jalón, J. G., 1996. "Multibody dynamics optimization by direct differentiation methods using object oriented programming". In IUTAM Symposium on Optimization of Mechanical Systems, Springer, pp. 213–220.
[86] Ambrósio, J. A., and Kecskeméthy, A., 2007. "Multibody dynamics of biomechanical models for human motion via optimization". In Multibody Dynamics. Springer, pp. 245–272.
[87] Bestle, D., and Eberhard, P., 1992. "Analyzing and optimizing multibody systems". Journal of Structural Mechanics, **20**(1), pp. 67–92.
[88] He, Y., and McPhee, J., 2005. "A design methodology for mechatronic vehicles: application of multidisciplinary optimization, multibody dynamics and genetic algorithms". Vehicle System Dynamics, **43**(10), pp. 697–733.
[89] Kang, B.-S., Park, G.-J., and Arora, J. S., 2005. "Optimization of flexible multibody dynamic systems using the equivalent static load method". AIAA journal, **43**(4), pp. 846–852.
[90] Reinhart, G., and Weissenberger, M., 1999. "Multibody simulation of machine tools as mechatronic systems for optimization of motion dynamics in the design process". In Advanced Intelligent Mechatronics, 1999. Proceedings. 1999 IEEE/ASME International Conference on, IEEE, pp. 605–610.
[91] He, Y., and McPhee, J., 2003. "Design optimization of rail vehicles with passive and active suspensions: A combined approach using genetic algorithms and multibody dynamics". Vehicle System Dynamics, **37**, pp. 397–408.
[92] Biegler, L. T., 2007. "An overview of simultaneous strategies for dynamic optimization". Chemical Engineering and Processing: Process Intensification, **46**(11), pp. 1043–1053.
[93] Betts, J. T., 1998. "Survey of numerical methods for trajectory optimization". Journal of guidance, control, and dynamics, **21**(2), pp. 193–207.
[94] Eberhard, P., Dignath, F., and Kübler, L., 2003. "Parallel evolutionary optimization of multibody systems with application to railway dynamics". Multibody System Dynamics, **9**(2), pp. 143–164.
[95] Samin, J.-C., Brüls, O., Collard, J.-F., Sass, L., and Fisette, P., 2007. "Multiphysics modeling and optimization of mechatronic multibody systems". Multibody System Dynamics, **18**(3), pp. 345–373.
[96] Eberhard, P., Schiehlen, W., and Bestle, D., 1999. "Some advantages of stochastic methods in multicriteria optimization of multibody systems". Archive of Applied Mechanics, **69**(8), pp. 543–554.

[97] Xiang, Y., Chung, H.-J., Kim, J. H., Bhatt, R., Rahmatalla, S., Yang, J., Marler, T., Arora, J. S., and Abdel-Malek, K., 2010. "Predictive dynamics: an optimization-based novel approach for human motion simulation". Structural and Multidisciplinary Optimization, **41**(3), pp. 465–479.

[98] Hull, D. G., 1997. "Conversion of optimal control problems into parameter optimization problems". Journal of Guidance, Control, and Dynamics, **20**(1), pp. 57–60.

[99] Crolla, D., and Abdel-Hady, M., 1991. "Active suspension control; performance comparisons using control laws applied to a full vehicle model". Vehicle System Dynamics, **20**(2), pp. 107–120.

[100] Sharp, R., and Wilson, D., 1990. "On control laws for vehicle suspensions accounting for input correlations". Vehicle System Dynamics, **19**(6), pp. 353–363.

[101] Zhu, C., Byrd, R. H., Lu, P., and Nocedal, J., 1997. "Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization". ACM Trans. Math. Softw., **23**(4), Dec., pp. 550–560.

[102] Ambrósio, J., 2003. "Impact of rigid and flexible multibody systems: deformation description and contact models". In Virtual nonlinear multibody systems. Springer, pp. 57–81.

[103] Zukas, J., 1982. "Impact dynamics". John Wiley & Sons, Inc, 605 Third Ave, New York, N. Y. 10158, U. S. A, 1982. 452.

[104] Bischof, C., Carle, A., Corliss, G., Griewank, A., and Hovland, P., 1992. "ADIFOR–generating derivative codes from fortran programs". Scientific Programming, **1**(1), pp. 11–29.

[105] Bischof, C. H., 1996. "On the automatic differentiation of computer programs and an application to multibody systems". In IUTAM Symposium on Optimization of Mechanical Systems, Springer, pp. 41–48.

[106] Bischof, C. H., Bouaricha, A., Khademi, P. M., and Moré, J. J., 1997. "Computing gradients in large-scale optimization using automatic differentiation". INFORMS Journal on Computing, **9**(2), pp. 185–194.

[107] Martins, J. R., Sturdza, P., and Alonso, J. J., 2003. "The complex-step derivative approximation". ACM Transactions on Mathematical Software (TOMS), **29**(3), pp. 245–262.

[108] Hairer, E., Lubich, C., and Wanner, G., 2006. Geometric numerical integration: structure-preserving algorithms for ordinary differential equations, Vol. 31. Springer.

[109] Cuadrado, J., Cardenal, J., Morer, P., and Bayo, E., 2000. "Intelligent simulation of multibody dynamics: Space-state and descriptor methods in sequential and parallel computing environments". Multibody System Dynamics, **4**(1), 02, pp. 55–73.

[110] Cuadrado, J., Gutierrez, R., Naya, M., and Morer, P., 2001. "A comparison in terms of accuracy and efficiency between a mbs dynamic formulation with stress analysis and a non-linear fea code". International Journal for Numerical Methods in Engineering, **51**(9), 07, pp. 1033–1052.

[111] Bottasso, C., Dopico, D., and Trainelli, L., 2008. "On the optimal scaling of index three DAEs in multibody dynamics". Multibody System Dynamics, **19**(1-2), pp. 3–20.

cited By (since 1996) 7.

[112] Petzold, L., and Lötstedt, P., 1986. "Numerical solution of nonlinear differential equations with algebraic constraints ii: Practical implications". SIAM Journal on Scientific and Statistical Computing, **7**(3), pp. 720–733.

[113] Bauchau, O. A., Epple, A., and Bottasso, C. L., 2009. "Scaling of Constraints and Augmented Lagrangian Formulations in Multibody Dynamics Simulations". Journal of Computational and Nonlinear Dynamics, **4**(2), APR.

[114] de Jalón, J. G., Callejo, A., and Hidalgo, A. F., 2012. "Efficient solution of maggis equations". Journal of computational and nonlinear dynamics, **7**(2), p. 021003.

[115] Pagalday, J., 1994. "Optimizacion del comportamiento dinamico de mecanismos". PhD thesis, Escuela Superior de Ingenieros Industriales de S. Sebastian.

[116] Sonneville, V., and Bruls, O., 2013. "Sensitivity analysis for multibody systems formulated on a lie group". Multibody System Dynamics, pp. 1–21.

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

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

[119] Venkataraman, P., 2009. Applied optimization with MATLAB programming. John Wiley & Sons.

[120] Bazaraa, M. S., Sherali, H. D., and Shetty, C. M., 2013. Nonlinear programming: theory and algorithms. John Wiley & Sons.

[121] Fletcher, R., 2013. Practical methods of optimization. John Wiley & Sons.

[122] Wright, S., and Nocedal, J., 1999. Numerical optimization, Vol. 2. Springer New York.

[123] Byrd, R. H., Lu, P., Nocedal, J., and Zhu, C., 1995. "A limited memory algorithm for bound constrained optimization". SIAM Journal on Scientific Computing, **16**(5), pp. 1190–1208.

[124] Pablo Luque Rodrguez, Daniel lvarez Mntaras, C. V., 2004. Ingeniera del automvil: sistemas y comportamiento dinmico. THOMSON.

[125] McGeer, T., 1990. "Passive dynamic walking". the international journal of robotics research, **9**(2), pp. 62–82.

[126] Collins, S., Ruina, A., Tedrake, R., and Wisse, M., 2005. "Efficient bipedal robots based on passive-dynamic walkers". Science, **307**(5712), pp. 1082–1085.

[127] Kuo, A. D., 1999. "Stabilization of lateral motion in passive dynamic walking". The International journal of robotics research, **18**(9), pp. 917–930.

[128] Collins, S. H., Wisse, M., and Ruina, A., 2001. "A three-dimensional passive-dynamic walking robot with two legs and knees". The International Journal of Robotics Research, **20**(7), pp. 607–615.

[129] Wisse, M., and Van Frankenhuyzen, J., 2006. "Design and construction of mike; a 2-D autonomous biped based on passive dynamic walking". In Adaptive motion of animals and machines. Springer, pp. 143–154.

[130] Tedrake, R., Zhang, T. W., Fong, M.-f., and Seung, H. S., 2004. "Actuating a simple 3D

passive dynamic walker". In Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on, Vol. 5, IEEE, pp. 4656–4661.

[131] McGeer, T., 1990. "Passive walking with knees". In Robotics and Automation, 1990. Proceedings., 1990 IEEE International Conference on, IEEE, pp. 1640–1645.

[132] Asano, F., Yamakita, M., and Furuta, K., 2000. "Virtual passive dynamic walking and energy-based control laws". In Intelligent Robots and Systems, 2000.(IROS 2000). Proceedings. 2000 IEEE/RSJ International Conference on, Vol. 2, IEEE, pp. 1149–1154.

[133] Garcia, M., Chatterjee, A., and Ruina, A., 1998. "Speed, efficiency, and stability of small-slope 2d passive dynamic bipedal walking". In Robotics and Automation, 1998. Proceedings. 1998 IEEE International Conference on, Vol. 3, IEEE, pp. 2351–2356.

[134] Wisse, M., Schwab, A. L., and van der Helm, F. C., 2004. "Passive dynamic walking model with upper body". Robotica, **22**(06), pp. 681–688.

[135] Wisse, M., Schwab, A. L., and Linde, R. v., 2001. "A 3D passive dynamic biped with yaw and roll compensation". Robotica, **19**(03), pp. 275–284.

[136] McGeer, T., 1993. "Passive dynamic biped catalogue, 1991". In Experimental Robotics II. Springer, pp. 463–490.

[137] Wisse, M., Schwab, A. L., van der Linde, R. Q., and van der Helm, F. C., 2005. "How to keep from falling forward: elementary swing leg action for passive dynamic walkers". Robotics, IEEE Transactions on, **21**(3), pp. 393–401.

[138] Wisse, M., 2005. "Three additions to passive dynamic walking: actuation, an upper body, and 3D stability". International Journal of Humanoid Robotics, **2**(04), pp. 459–478.

[139] Hitomi, K., Shibata, T., Nakamura, Y., and Ishii, S., 2006. "Reinforcement learning for quasi-passive dynamic walking of an unstable biped robot". Robotics and Autonomous Systems, **54**(12), pp. 982–988.

[140] Lankarani, H., and Nikravesh, P., 1990. "A contact force model with hysteresis damping for impact analysis of multibody systems". Journal of Mechanical Design, **112**(3), pp. 369–376.

[141] Djerassi, S., 2009. "Collision with friction; part a: Newtons hypothesis". Multibody System Dynamics, **21**(1), pp. 37–54.

[142] Djerassi, S., 2009. "Collision with friction; part b: Poissons and stornges hypotheses". Multibody System Dynamics, **21**(1), pp. 55–70.

[143] Hunt, K., and Crossley, F., 1975. "Coefficient of restitution interpreted as damping in vibroimpact". Journal of applied mechanics, **42**(2), pp. 440–445.