

Diss. ETH No. 19715

# Consistent Integrators for Non-Smooth Dynamical Systems

A dissertation submitted to

ETH ZURICH

for the degree of  
Doctor of Sciences

presented by

MICHAEL HEINER MÖLLER

Dipl. Masch.-Ing. ETH  
born June 11, 1980  
citizen of Zurich (Switzerland) and Germany

accepted on the recommendation of

Prof. Dr.-Ing. Dr.-Ing. habil. Ch. Glocker, examiner  
Prof. Dr.-Ing. habil. P. Betsch, co-examiner

2011



# Acknowledgements

---

The work presented in this thesis has been carried out at the Center of Mechanics in the Department of Mechanical and Process Engineering at the ETH Zurich.

First of all, I would like to thank my supervisor Prof. Dr.-Ing. Dr.-Ing. habil. Ch. Glocker for supporting and guiding my research. He introduced me into the field of non-smooth dynamical systems. With his deep knowledge of mechanics and dynamical systems he could always provide very valuable help. His enthusiasm about mechanics and his quest for consistent and elegant formulations have been a great source of motivation for me. I very much appreciated the many interesting discussions.

Furthermore, I would like to thank Prof. Dr.-Ing. habil. P. Betsch for reviewing my thesis and for the valuable remarks. His research on consistent integrators has been an essential basis for my work.

I really enjoyed the good atmosphere and the nice time at the Center of Mechanics wherefore all colleagues are sincerely thanked. Especially I would like to thank PD Dr.ir. habil. Remco I. Leine for his support with many numerical and non-smooth problems, Dr. Ueli Aeberhard for his help with numerous mathematical questions, Dr. Christian Studer for the many discussions about non-smooth problems and numerics. Furthermore, I would like to thank Ondrej Papes for the many interesting discussions about mechanics.

Finally I would like to thank my family and friends for the support and continuous encouragement.



# Abstract

---

Non-smooth mechanical models with set-valued force laws of normal cone type are very well suited for the analysis of the dynamics of mechanical systems with unilateral contacts and Coulomb friction. In a non-smooth formulation, discontinuous velocities and impulsive forces may occur, which allows to describe not only structural changes like stick-slip transitions but impacts as well. Like the force laws, also impact laws can be formulated as normal cone inclusion. For the numerical integration of the dynamics of non-smooth models, event capturing time-stepping schemes based on Moreau's midpoint rule in combination with a reformulation of the normal cone inclusions as proximal point problems have been proven to be very robust. Usually these time-stepping integrators use a fully implicit discretization of all non-smooth forces, while integrating all classical and smooth forces as well as displacements related to unilateral contacts with an explicit scheme. If the time steps cannot be chosen small enough (e.g. for performance reasons), this can lead to an unstable integration, drift in the energy balance or drift in the unilateral constraints.

In this work, different variants of consistent integrators for non-smooth mechanical systems are developed. The integrators discretize all terms with an implicit scheme, which allows to achieve an energetically consistent integration. This means, if the total energy was preserved or strictly decreasing in the model, then the difference scheme used for integration only allows approximations which have this property as well. Similarly drift problems with unilateral contacts and bilateral constraints are addressed. The integrators are based on one- or two-step schemes and allow for inertial and potential forces, bilateral constraints and set-valued force laws of normal cone type. The set-valued force laws can be equipped with impact laws of Newton-type. The consistency properties of the integrators are useful for increasing the overall robustness of the integration.

When formulating consistent integrators based on equations of motion in differential-algebraic form, one has to be careful not to increase the number of equations too drastically, or otherwise performance will suffer. To address this problem, a scalable body model with three translational, three rotation and one uniform scaling degree of freedom is introduced in this work. The rotational and scaling degrees of freedom are parametrized with an unconstrained quaternion. To the scalable body a perfect bilateral constraint is added, restricting the quaternion to unit length and making the body rigid. This way a quaternion based differential algebraic equation formulation for the dynamics of a rigid body is obtained, where the mass matrix is regular and the unit length restriction of the quaternion is enforced by a mechanical constraint.

The completely implicit discretization of the equations of motion yields a set of coupled nonlinear equations and normal cone inclusions that has to be solved for every time step. Normal cone inclusion problems with linear equations, as obtained for Moreau's midpoint rule, can be solved with a projected Jacobi iteration or a projected Gauss-Seidel iteration. For the nonlinear case, a projected Newton iteration is proposed in this work. The iterative methods for the solution of normal cone inclusions rely on an efficient implementation of the proximal point functions for the corresponding convex sets. To help with this, a transformation technique for proximal point iterations is developed. This can be used for multidimensional normal cone inclusion problems using convex sets with a complex shape that can be simplified through a linear transformation. The properties of the consistent integrators and solution methods developed in this work are demonstrated on some examples of non-smooth systems.

# Zusammenfassung

---

Nicht-glatte mechanische Modelle mit mengenwertigen Kraftgesetzen vom Typ Normalkegel sind sehr gut für die Analyse der Dynamik von mechanischen Systemen mit einseitigen Kontakten und Coulomb-Reibung geeignet. In der nicht-glaten Formulierung können diskontinuierliche Geschwindigkeiten und impulsive Kräfte auftreten, was es erlaubt nicht nur strukturelle Veränderungen wie Haft-Gleit-Übergänge sondern auch Stöße zu beschreiben. Die Stossgesetze können genau wie die Kraftgesetze auch als Normalkegelinklusion formuliert werden. Time-Stepping-Verfahren basierend auf Moreau's Mittelpunktsregel in Kombination mit der Umformulierung der Normalkegelinklusionen als Prox-Gleichung haben sich als sehr robuste Methode für die numerische Integration der Dynamik von nicht-glaten Modellen erwiesen. Normalerweise verwenden diese Time-Stepping-Integratoren eine vollständig implizite Diskretisierung aller nicht-glaten Kräfte, während alle klassischen und glatten Kräfte sowie die Auslenkungen der einseitigen Kontakte mit einem expliziten Schema integriert werden. Falls die Zeitschritte nicht klein genug gewählt werden können (z.B. aus Performance-Gründen), kann dies zu einer instabilen Integration, zu Drift in der Energiebilanz oder zu Drift in den einseitigen Bindungen führen.

In dieser Arbeit werden verschiedene Varianten konsistenter Integratoren für nicht-glatte mechanische Systeme entwickelt. Die Integratoren diskretisieren alle Terme mit einem impliziten Schema, wodurch eine energetisch konsistente Integration erreicht werden kann. Das heisst, wenn die Gesamtenergie im Model erhalten bleibt oder monoton fällt, so erlaubt das Differenzenschema des Integrators nur Approximationen welche diese Eigenschaft ebenfalls haben. In ähnlicher Art und Weise können auch Driftprobleme bei einseitigen Kontakten und bilateralen Bindungen behandelt werden. Die Integratoren basieren auf Ein- oder Zweisritt-Schemen und können Trägheitskräfte, Potentialkräfte, bilaterale Bindungen und mengenwertige Kraftgesetze vom Typ Normalkegel berücksichtigen. Die mengenwertige Kraftgesetze können mit Newton-Stossgesetzen ausgerüstet werden. Die Konsistenzeigenschaften der Integratoren sind nützlich für die Erhöhung der Robustheit der Integration.

Bei der Formulierung konsistenter Integratoren basierend auf Bewegungsgleichungen in differential-algebraischer Form muss darauf geachtet werden, dass die Anzahl der Gleichungen nicht zu stark erhöht wird, da sonst die Performance leidet. Um dieses Problem anzugehen wird in dieser Arbeit das Modell eines skalierbaren Körpers eingeführt, welcher drei translatorische, drei rotatorische und einen uniform skalierenden Freiheitsgrad hat. Die rotatorischen und skalierenden Freiheitsgrade werden mit einer uneingeschränkten Quaternion parametrisiert. Dem skalierbaren Körper wird eine perfekte bilaterale Bin-

zung hinzugefügt, welche die Quaternion auf Einheitslänge einschränkt und das Modell auf einen Starrkörper reduziert. Auf diesem Weg erhält man eine auf Quaternionen basierte differential-algebraische Formulierung für die Dynamik eines Starrkörpers für welche die Massenmatrix regulär ist und die Einheitslänge des Quaternions von einer mechanischen Bindung erzwungen wird.

Die komplett implizite Diskretisierung der Bewegungsgleichungen ergibt ein System von nichtlinearen Gleichungen und Normalkegelinklusionen, welches für jeden Zeitschritt numerisch gelöst werden muss. Normalkegelinklusionen mit linearen Gleichungen, wie man sie mit Moreau's Mittelpunktsregel erhält, können mit einer projizierten Jacobi-Iteration oder einer projizierten Gauss-Seidel-Iteration gelöst werden. Für den nichtlinearen Fall wird in dieser Arbeit eine projizierte Newton-Iteration vorgeschlagen. Die iterativen Methoden zur Lösung von Normalkegelinklusionen setzen auf eine effiziente Implementierung der Prox-Funktionen für die entsprechenden konvexen Mengen. Um die Effizienz zu verbessern wurde eine Transformationsmethode für Prox-Iterationen entwickelt. Diese Methode kann für mehrdimensionale Normalkegelinklusionen verwendet werden, sofern die konvexe Menge durch eine lineare Transformation vereinfacht werden kann. Die Eigenschaften der in dieser Arbeit entwickelten konsistenten Integratoren und Lösungsverfahren werden anhand von einigen Beispielen nicht-glatte mechanischer Systeme veranschaulicht.



# Contents

---

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Motivation . . . . .	1
1.2	Literature Survey . . . . .	2
1.2.1	Non-smooth Dynamical Systems . . . . .	2
1.2.2	Consistent Integrators . . . . .	4
1.3	Aim and Scope . . . . .	5
1.4	Outline . . . . .	6
<b>2</b>	<b>Normal Cone Inclusions</b>	<b>9</b>
2.1	Convex Sets and Convex Functions . . . . .	9
2.2	Normal Cones and Proximal Points . . . . .	11
2.3	Basic Proximal Point Functions . . . . .	13
2.4	Transformation of Normal Cone Inclusions . . . . .	14
2.5	Solving Normal Cone Inclusions . . . . .	18
2.5.1	Projected Jacobi Iteration . . . . .	18
2.5.2	Transformation of Proximal Point Iterations . . . . .	20
2.5.3	Recovering the Jacobi Iteration . . . . .	22
2.5.4	Projected Gauss-Seidel Iteration . . . . .	24
2.5.5	Projected Newton Iteration . . . . .	24
<b>3</b>	<b>Equations of Motion</b>	<b>27</b>
3.1	Principles of Virtual Work and Virtual Power . . . . .	28
3.2	Generalized Coordinates . . . . .	29
3.3	Generalized Velocities . . . . .	31
3.4	Coordinate Dependent Mass Matrix . . . . .	32
3.5	Decomposed Mass Matrix . . . . .	34
3.6	Generalized Forces . . . . .	36
3.6.1	Local Force Laws on Displacement Level . . . . .	36
3.6.2	Local Force Laws on Velocity Level . . . . .	38
3.6.3	Potential Force Laws . . . . .	41
3.7	Set-Valued Force Laws of Normal Cone Type . . . . .	42
3.7.1	Bilateral Constraint . . . . .	42
3.7.2	Unilateral Contact . . . . .	43
3.7.3	Coulomb Friction . . . . .	44

3.7.4	Spatial Coulomb Friction . . . . .	46
3.7.5	Coulomb-Contensou Friction . . . . .	48
3.8	Impacts . . . . .	51
3.9	Total Energy . . . . .	52
<b>4</b>	<b>Rigid Body Formulations</b>	<b>59</b>
4.1	Mass Distribution . . . . .	60
4.2	Affine Body . . . . .	61
4.2.1	Kinematics . . . . .	61
4.2.2	Equations of Motion . . . . .	63
4.2.3	Rigidity Constraints . . . . .	64
4.3	Scalable Body . . . . .	65
4.3.1	Quaternions . . . . .	66
4.3.2	Rotation and Scaling . . . . .	68
4.3.3	Kinematics . . . . .	70
4.3.4	Equations of Motion . . . . .	73
4.3.5	Scaling Constraint . . . . .	75
4.3.6	Generalized Velocities . . . . .	77
<b>5</b>	<b>Integrators</b>	<b>81</b>
5.1	Moreau's Midpoint Rule . . . . .	81
5.2	Consistent Integrators . . . . .	83
5.2.1	Discrete Derivative . . . . .	84
5.2.2	Constant Mass Matrix . . . . .	87
5.2.3	Coordinate Dependent Mass Matrix . . . . .	92
5.2.4	Decomposed Mass Matrix . . . . .	95
5.2.5	Quaternion Based Formulation . . . . .	107
<b>6</b>	<b>Examples</b>	<b>115</b>
6.1	Oscillator . . . . .	115
6.2	Unilateral Double Pendulum . . . . .	118
6.3	Overbalanced Wheel . . . . .	122
6.4	Woodpecker . . . . .	129
6.5	Tippe Top . . . . .	137
<b>7</b>	<b>Conclusions</b>	<b>145</b>

# Introduction

---

This first chapter starts with a short introduction and motivation for the thesis. After a literature survey in Section 1.2, the aim and scope of the thesis is presented in Section 1.3. An outline of the chapters of this thesis is given in Section 1.4.

## 1.1 Motivation

Unilateral constraints and friction are present in many mechanical systems. Sometimes unilateral constraints can be replaced with bilateral formulations and the influence of friction can be neglected. This is no surprise, as most technical realizations of what in mechanics is considered a bilateral constraint rely on a special arrangement of unilateral contacts with reduced friction. Perfect bilateral constraints have a long tradition in mechanics, and according methods are very well developed. Still one is occasionally confronted with unilateral contacts or friction that can not be replaced with perfect bilateral constraints. For these cases, one approach is to use smooth force laws for the description of the unilateral contacts with friction in a regularized form. Depending on the situation, this can have disadvantages, like adding unnecessary parameters to the model or numerical difficulties associated with stiff differential equations. An alternative approach is to describe unilateral constraints and friction in an idealized way. In the last decades this lead to a very well developed framework for the formulation of non-smooth mechanical models with unilateral contacts and friction. The word *non-smooth* refers to the fact, that the displacements may not be differentiable everywhere, the velocities can be discontinuous, and the forces can even contain Dirac impulses.

In a non-smooth mechanical model, the unilateral contacts and friction can be described with set-valued force and impact laws of normal cone type. The formulation as normal cone inclusion provides a very general structure for a large class of force laws, even classical bilateral constraints can be formulated as normal cone inclusion. Time-stepping schemes based on Moreau's midpoint rule in combination with a reformulation of the normal cone inclusions as proximal point problems have been proven to be very useful for the numerical integration of the dynamics of non-smooth models. Moreau's midpoint rule uses an implicit discretization of all set-valued force laws, while all classical forces

are discretized in an explicit way. For larger time steps, the explicit discretization can lead to an unstable integration. For systems where high frequency dynamics is created by classical forces this can force Moreau's midpoint rule to very small time steps in order to avoid instabilities of the integrator. This could be avoided with a completely implicit scheme. The progress on conserving integrators in recent years suggests not to use just a random kind of implicit discretization, but one that yields an energetically consistent integration. This means, if the total energy was preserved or strictly decreasing in the model, then the difference scheme used for integration only allows approximations which have this property as well. The consistency with respect to the total energy guarantees by itself that the numerical solution does not diverge to infinity. In this context also drift problems of unilateral contacts caused by the usual discretization on velocity level could be addressed. Beside stability and robustness enhancements, a differential-algebraic equations based integrator formulation could be useful for larger mechanical systems with a more complicated structure. Differential-algebraic formulations are good for systems where the number of degrees of freedom or kinematic loops make a minimal coordinates based approach difficult. Beyond their direct use for the analysis of non-smooth mechanical applications, consistent integration schemes can be interesting for the interpretation and understanding of the properties of time-stepping schemes closer to Moreau's original midpoint rule.

## 1.2 Literature Survey

In this section, a short literature survey on non-smooth dynamical systems and consistent integrators is given. The intent of this section is not to give a complete overview on the relevant literature, but to give some references which might be helpful in getting more detailed information and background knowledge on the topic.

### 1.2.1 Non-smooth Dynamical Systems

There exists a large number of publications on the modeling and numerical simulation of non-smooth dynamical systems. A review of the main techniques and algorithms for the numerical simulation of non-smooth multibody systems is given in [16]. In the book [3] by Acary and Brogliato a detailed collection of numerical methods for non-smooth dynamical systems in the field of mechanics and electronics is given. The methods discussed in the book include the different non-smooth formulations, numerical time-integration schemes and non-smooth problem solvers.

The non-smooth formulation of the dynamics of mechanical systems and their numerical solution has been pioneered by the work of Moreau [60, 61] and Jean [38, 39]. The framework developed by Moreau and Jean is based on convex analysis [72]. It formulates the dynamics as a measure differential inclusion on velocity level. The midpoint rule published in [60] is a first event capturing (or time-stepping) integrator, that allows to handle all non-smooth events occurring in a time interval with one difference scheme. Non-smooth events are for example slip-stick transitions or impacts. In contrast to event capturing

integrators, an event driven integrator splits the dynamics into non-smooth events and smooth parts. Both parts are then handled separately. An event capturing integrator avoids the difficulties that can occur in an event driven integrator at accumulation points of non-smooth events.

In the work of Glocker [23] and the book by Pfeiffer and Glocker [71], the formulation of dynamical problems with friction and impacts in a linear complementarity framework has been developed. The equations of motion are formulated on acceleration level and integrated with an event driven integrator. A very detailed presentation of the linear complementarity problem (LCP) and numerical solution methods for it can be found in the book of Cottle et al. [19] and in the book of Murty [63]. In [9] Baraff presents a nonlinear extension of the LCP for frictional contact problems and a correspondingly modified version of Dantzig's algorithm [18] for the solution of LCPs. With the linear complementarity problem based formulations one is limited to one-dimensional friction or one has to use pyramidal approximation of the friction cone [24, 41] for the multidimensional case. The approximation of the friction cone can be avoided by using a normal cone based formulation of the friction law as described in the book by Glocker [25]. In the book [25] a detailed framework for set-valued force laws on different kinematic levels and their representation by non-smooth potentials is given. Beside this, also the non-smooth extension of classical principles in dynamics is described. An introduction to impact constitutive laws formulated in terms of set-valued maps can be found in [23, 27].

Based on Moreau's midpoint rule, many variants of event capturing integrators have been developed. These include the  $\theta$ -method by Jean [39], a formulation on displacement level by Paoli and Schatzman [67, 68] or the methods by Anitescu, Stewart and Trinkle [8, 76]. In [81, 78] Studer developed a time-stepping method with step size adjustment and extrapolation. Beside this, in [78] also a review of many of the event capturing integrators can be found. Numerical methods for the time integration of higher order non-smooth dynamical systems have been developed by Acary and Brogliato in [2, 4].

The one-step problem arising from the discretization of normal cone inclusion based formulations can be solved by an augmented Lagrangian approach resulting in an iterative projection method. The augmented Lagrangian approach for contact problems has been described by Alart and Curnier [6] and by Simo and Laursen [74]. The iterative projection method has been applied to dynamic mechanical problems by Leine and Glocker [47, 48, 29]. A detailed discussion of the augmented Lagrangian approach and the resulting iterative methods for the solution of normal cone inclusions as well as some analysis on convergence of the iteration has been published by Studer [80, 78].

Using an implicit discretization of the classical forces in a scheme based on Moreau's midpoint rule yields a one-step problem that consists of nonlinear equations and inclusions. In general these can be difficult to solve, especially if the one-step problem approximates impacts as well. With jumps in the velocities, one might not have a starting point for a Newton-type iteration that is close to the solution. One way to circumvent this problem, is to apply a time-splitting method as developed in [64]. The integration is split into a smooth nonlinear part that can be solved with a Newton-type method and a non-smooth inclusion based part that can be solved with a proximal point iteration.

Examples of the application of the non-smooth formulations for mechanical systems are

the simulation of the woodpecker toy by Glocker [23, 30], the analysis of the curve squealing mechanism of railway vehicles by Cataldi-Spinola [17], the design and simulation of a passively actuated robot manipulator by Welge-Lüßen [86] or the simulation of a snake robot by Transeth et al. [82]. The non-smooth formulations have been successfully applied as well to electrical systems by Glocker and Möller [26, 56, 57] as well as Acary et al. [1].

## 1.2.2 Consistent Integrators

The energy and momentum conserving integrators for particle system by LaBudde and Greenspan [43, 44] are early works on this topic. In [75] Simo and Wong describe an integrator for rigid body dynamics based on the Newton-Euler equations that preserves energy and momentum for force free motion. An important building block for the consistent integrators formulated in this work is the discrete derivative introduced by Gonzalez [31, 32]. It allows to approximate a partial derivative with a discrete version that yields exactly the difference in function values when multiplied with the difference of the arguments used to construct it.

In [32] a conserving integrator for Hamiltonian systems is developed that preserves the total energy and certain first integrals arising from symmetry. The integrator is based on the discrete derivative technique. In [33] Gonzalez develops a conserving integrator for mechanical systems subject to holonomic constraints where the equations of motion are formulated in differential-algebraic form. The integrator preserves the total energy, the constraints on displacement level and integrals arising from symmetry like the linear and angular momentum. The induced constraints on velocity are not preserved in the approximation of the integrator, which leads to oscillations in the velocities [33].

Inspired by the method of Gear et al. [22], Betsch and Steinmann [14] developed a conserving integrator based on a Hamiltonian differential-algebraic equation formulation, which not only preserves the total energy and momentum but also enforces consistent constraints on displacement and velocity level. The director based rigid body formulation described in [14] leads to a system of nonlinear equations with a rather large number of unknowns. To reduce the number of unknowns, Betsch developed the discrete null space method for consistent integrators in [10, 11]. The discrete null space method allows to eliminate the discrete Lagrange multipliers present after the discretization of the equations of motion in differential-algebraic form. The discrete null space method has been extended for flexible multibody systems by Leyendecker et al. [52, 53]. Unfortunately, the integrator developed in [10] enforces the bilateral constraints on displacement level only, and assumes a constant mass matrix.

A consistent integrator using a decomposed mass matrix has been developed by Lens et al. [51]. For a decomposed mass matrix, the kinetic energy can be expressed in terms of a constant matrix and an appropriate set of generalized velocities. In [51] also an example with two pendulums interacting with one contact is considered. Some kind of contact law is formulated by introducing a slack variable, but without formulating a complementarity condition which would guarantee compressive-only forces at the unilateral contact. Beside this, no impact law is formulated. In [51] the bilateral constraints are enforced on displacement level only, which leads as well to oscillations of the velocities. An attempt to

reduce the oscillations with additional numerical dissipation has been developed by Lens and Cardona in [50].

In the work of Hesch [36] consistent integrators for the dynamics of deformable bodies with frictionless unilateral contacts are developed. An impact constitutive law is not stated explicitly. Many modeling techniques for consistent integrators based on the discrete null space method are developed in the work of Uhlar [83]. This includes formulations for rigid and flexible bodies as well as null space matrices for closed loop systems, the formulation of control constraints and the modeling of dissipation. In [83] also a nice collection of example systems and applications is presented.

### 1.3 Aim and Scope

The aim of this thesis is the development of consistent integrators for the dynamics of non-smooth mechanical systems. The focus lies on the consistency of the integrators with respect to the total energy and to the kinematics of unilateral and bilateral constraints. Other consistency properties like the preservation of generalized momenta associated with symmetries of the system are not considered. The motivation for the energy consistency are the inherent stability properties. A consistent discretization of bilateral constraints on displacement level is required as resolving drift problems with projections after each time step is not feasible for energy consistent integrators. For the combination with impacts also a consistent discretization of bilateral constraints on velocity level is required. For unilateral constraints only a reduced form of kinematic consistency is considered. The consistent integrators considered in this work are using the discrete derivative approximation of partial derivatives in order to preserve properties of the equations of motion in the discretized formulation.

Basic support for rigid multibody formulations is an essential requirement, in order to be able to formulate any model that goes beyond simple point masses. The rigid body formulation requirement together with the consistency of bilateral constraints on displacement and velocity level can lead to formulations that use a rather large number of unknowns. This problem is addressed with the development of a quaternion based rigid body formulation in differential-algebraic form suitable for the construction of consistent integrators.

In this work, the non-smooth forces are described with set-valued force laws of normal cone type. This allows for models with unilateral contacts and one- or multidimensional Coulomb friction. The impact-free dynamics is described with equations of motion and set-valued force laws in terms of the time derivatives of the generalized coordinates and velocities. The impacts are described with separate impact equations and impact laws relating the pre- and post-impact magnitudes. In order to keep the argumentation and notation simple, no differential measures based formulation is used. The combination of the continuous differential inclusions and the impact mappings into one differential measure inclusion formulation creates additional problems which are not further investigated here.

The discretization of the integrator yields a one-step inclusion problem that has to be

solved for every time step. In this work a few solution methods for normal inclusion problems with linear equations are discussed. The one-step inclusion problems resulting from the completely implicit discretization of a consistent integrator consist of coupled normal cone inclusions and nonlinear equations in general. For this problem a projected Newton iteration is proposed in this work. In the context of proximal point iteration methods also a few improvements for multidimensional force laws are developed.

The properties of the consistent integrators are demonstrated for a few examples of non-smooth systems. The detailed integration order and convergence properties of the consistent integrators are not analyzed in this work. This is mainly, because the discussion of the uniqueness and existence of solutions of non-smooth problems is difficult already. Correspondingly, a sound analysis of the convergence of an approximation towards such a solution can be even more complicated. The integration order of non-smooth event capturing integrators is just equal to one in general, if no additional measures are taken to locate the impact events more precisely.

## 1.4 Outline

In Chapter 2, an introduction to normal cone inclusions on convex sets and the relation to proximal point problems is given. The reformulation of a normal cone inclusion problem as a proximal point problem can be used for the iterative solution of the inclusion problem. Two proximal point based iteration methods for the combined solution of linear equations and normal cone inclusions are discussed. The linear transformation of normal cone inclusions and the recovering strategy for multidimensional convex sets presented in the following can be used for the efficient numerical solution of ellipsoid based normal cone inclusions. Finally the Newton based iterative method for the solution of coupled nonlinear equations and normal cone inclusions is proposed.

In Chapter 3, the formulation of the equations of motion of non-smooth mechanical systems is described. This chapter starts with the classical principles of virtual work and virtual power, the reduction to generalized coordinates and the reformulation using generalized velocities. As preparation for the consistent integrators, two special representations of the inertial forces in the equations of motion are discussed. In the following, different classes of force laws and the most important set-valued force laws of normal cone type are presented. The chapter concludes with a short introduction to impacts and a discussion of the different contributions to a total energy balance.

In Chapter 4, two rigid body formulations in differential-algebraic form are described. The first part of the chapter is a short summary of the affine body based formulation using twelve degrees of freedom and 6 bilateral constraints. In the second part a detailed discussion of the scalable body based rigid body formulation is presented. The scalable body uses an unconstrained quaternion and a translation vector as coordinates. Adding one perfect bilateral constraint reduces the scalable body to a rigid body.

In Chapter 5, the event capturing integrator based on Moreau's midpoint rule is described as a reference for the consistent integrators. After a short summary on the discrete derivative, the different consistent integrators based on differential-algebraic equations



---

are developed. This includes a discussion of the consistency of bilateral constraints on displacement and velocity level. Also for each integrator the consistency with respect to the total energy is shown. The chapter concludes with a quaternion based consistent integrator that uses the Newton-Euler formulation of rigid body dynamics.

The numerical solution for a few examples of non-smooth mechanical systems are presented in Chapter 6. The examples demonstrate the properties of the consistent integrator formulations developed in Chapter 5. Also the efficient evaluation of a Coulomb-Contensou friction approximation is shown, using the normal cone transformation technique from chapter Chapter 2.

In Chapter 7, a few conclusions are presented and the main contributions of this thesis are summarized. Finally, some of the remaining open questions are discussed in an outlook.



## Normal Cone Inclusions

---

In this chapter an introduction to normal cones, proximal points and the solution of normal cone inclusion problems is given. Normal cone inclusions can be used to formulate non-smooth force laws like the unilateral contact law or Coulomb's friction law. In the numerical integration of the resulting equations of motion, these force laws of normal cone type give rise to normal cone inclusion problems, which have to be solved numerically. In Section 2.1 a short introduction to the terminology and notation used for convex sets, indicator functions and the subdifferential is given. For more details on the subject the reader is referred to [72, 23, 46]. In Section 2.2 the normal cone and its connection to the proximal point function is described. Some basic proximal point functions are given in Section 2.3. The linear transformation of normal cone inclusions is described in Section 2.4. In Section 2.5 some iterative methods for the solution of normal cone inclusions are described. Parts of this chapter related to the normal cone and its transformation are based on the conference proceeding [59] of the author.

### 2.1 Convex Sets and Convex Functions

A set  $\mathcal{C} \subseteq \mathbb{R}^n$  is convex if for all  $\mathbf{x} \in \mathcal{C}$  and  $\mathbf{y} \in \mathcal{C}$  also  $\lambda \mathbf{x} + (1 - \lambda)\mathbf{y} \in \mathcal{C}$  holds for all  $\lambda \in [0, 1]$ . An application of this definition is shown in Figure 2.1.

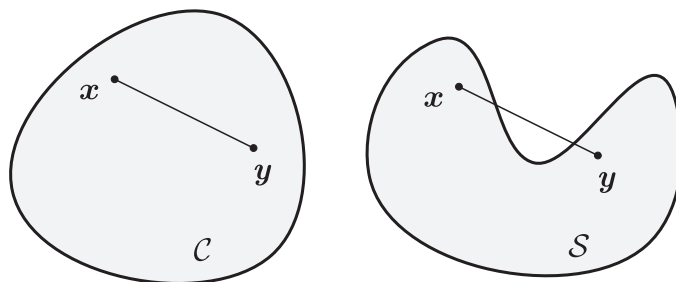


Figure 2.1: Convex set  $\mathcal{C}$  and non-convex set  $\mathcal{S}$ .

The indicator function  $\Psi_{\mathcal{C}} : \mathbb{R}^n \rightarrow \mathbb{R}$  of a set  $\mathcal{C}$  is defined as

$$\Psi_{\mathcal{C}}(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \in \mathcal{C} \\ +\infty, & \mathbf{x} \notin \mathcal{C} \end{cases} \quad (2.1)$$

An illustration of the indicator function of a one-dimensional convex set is shown in Figure 2.2.

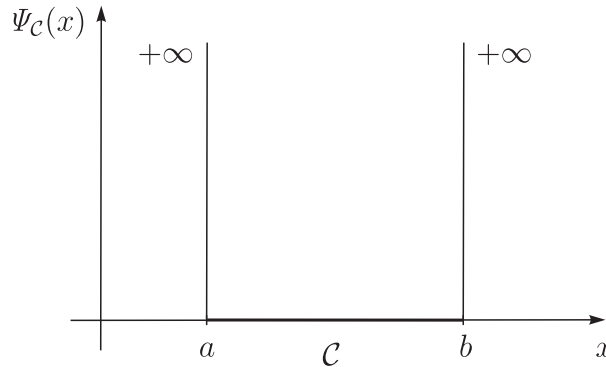


Figure 2.2: Indicator function  $\Psi_{\mathcal{C}}$  of a convex set  $\mathcal{C}$ .

The epigraph of a function  $f : \mathcal{D} \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  is the set of points lying on or above the graph of  $f$ ,

$$\text{epi}(f) = \{(\mathbf{x}, y) \mid \mathbf{x} \in \mathcal{D}, y \in \mathbb{R}, y \geq f(\mathbf{x})\}. \quad (2.2)$$

If the epigraph of a function  $f$  is a convex subset of  $\mathbb{R}^{n+1}$ , then the function  $f$  is convex. Note that the indicator function of a convex set is always a convex function. For a convex functions  $f$  the subdifferential is defined as

$$\partial f(\mathbf{x}) = \{\mathbf{y} \mid f(\mathbf{x}^*) \geq f(\mathbf{x}) + \mathbf{y}^T(\mathbf{x}^* - \mathbf{x}), \forall \mathbf{x}^*\}. \quad (2.3)$$

The epigraph and the subdifferential of a convex function is illustrated in Figure 2.3.

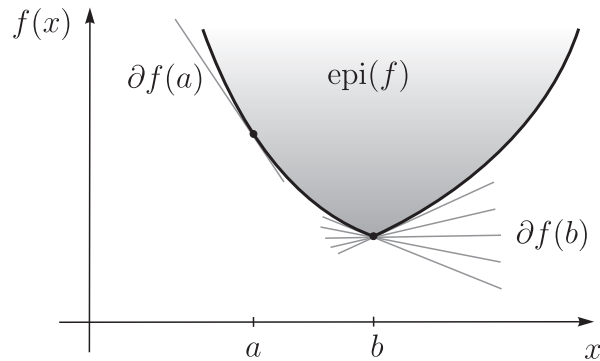


Figure 2.3: Epigraph and supporting hyperplanes of a convex function  $f(x)$ .

## 2.2 Normal Cones and Proximal Points

The normal cone  $\mathcal{N}_{\mathcal{C}}(\mathbf{x})$  to a convex set  $\mathcal{C} \subseteq \mathbb{R}^n$  at the point  $\mathbf{x} \in \mathcal{C}$  is the set of all vectors  $\mathbf{y} \in \mathbb{R}^n$  which do not form an acute angle to any vector  $\mathbf{x}^* - \mathbf{x}$  for all points  $\mathbf{x}^* \in \mathcal{C}$ , i.e.

$$\mathcal{N}_{\mathcal{C}}(\mathbf{x}) := \{\mathbf{y} \mid \mathbf{y}^{\top}(\mathbf{x}^* - \mathbf{x}) \leq 0, \forall \mathbf{x}^* \in \mathcal{C}\}. \quad (2.4)$$

If  $\mathbf{x}$  is in the interior of  $\mathcal{C}$ , then the normal cone contains only the 0-element  $\mathcal{N}_{\mathcal{C}}(\mathbf{x}) = \{0\}$ . An illustration of the normal cone for a two-dimensional convex set is shown in Figure 2.4. The normal cone  $\mathcal{N}_{\mathcal{C}}(\mathbf{x})$  to a convex set  $\mathcal{C}$  is identical to the subdifferential  $\partial \Psi_{\mathcal{C}}(\mathbf{x})$  of the

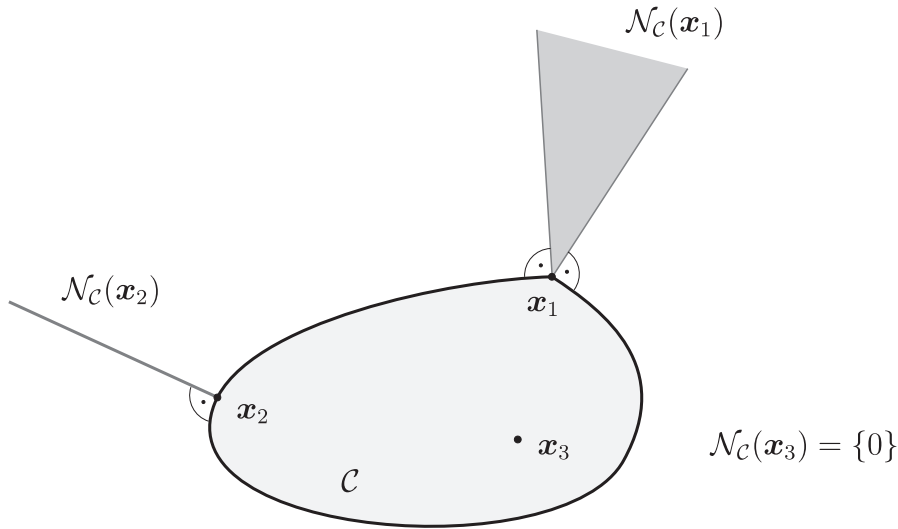


Figure 2.4: Normal cone.

indicator function of the convex set,

$$\partial \Psi_{\mathcal{C}}(\mathbf{x}) \equiv \mathcal{N}_{\mathcal{C}}(\mathbf{x}). \quad (2.5)$$

The normal cone is related to the problem of finding the proximal point to a convex set. For the definition of a proximal point a norm is required. With a symmetric and positive definite (PD) matrix  $\mathbf{R} \in \mathbb{R}^{n \times n}$ , the inner product of two vectors can be defined as

$$\langle \mathbf{x}, \mathbf{y} \rangle := \mathbf{x}^{\top} \mathbf{R} \mathbf{y}, \quad \mathbf{R} = \mathbf{R}^{\top}, \mathbf{R} \text{ PD}. \quad (2.6)$$

This inner product naturally defines the norm  $\|\mathbf{x}\|_R$  of a vector  $\mathbf{x} \in \mathbb{R}^n$  as

$$\|\mathbf{x}\|_R = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{\mathbf{x}^{\top} \mathbf{R} \mathbf{x}}. \quad (2.7)$$

The proximal point  $\text{prox}_{\mathcal{C}}^R(\mathbf{x})$  of a point  $\mathbf{x} \in \mathbb{R}^n$  to the convex set  $\mathcal{C} \subseteq \mathbb{R}^n$  is the closest point in  $\mathcal{C}$  to  $\mathbf{x}$  with respect to the norm  $\|\cdot\|_R$ ,

$$\text{prox}_{\mathcal{C}}^R(\mathbf{x}) := \underset{\mathbf{x}^* \in \mathcal{C}}{\text{argmin}} \|\mathbf{x} - \mathbf{x}^*\|_R. \quad (2.8)$$

If  $\mathbf{x}$  is in the set  $\mathcal{C}$ , then the proximal point  $\text{prox}_{\mathcal{C}}^R(\mathbf{x})$  is  $\mathbf{x}$  itself. The proximal point function  $\text{prox}_{\mathcal{C}}^R$  is invariant to scaling of the matrix  $\mathbf{R}$  of the associated norm  $\|\cdot\|_R$  by a positive scalar  $\alpha$ ,

$$\text{prox}_{\mathcal{C}}^R(\mathbf{x}) = \text{prox}_{\mathcal{C}}^{\alpha R}(\mathbf{x}), \quad \alpha \in \mathbb{R}^+. \quad (2.9)$$

The function  $\text{prox}_{\mathcal{C}}(\mathbf{x})$  denotes the proximal point function for the special case of the Euclidean norm which is obtained for  $\mathbf{R} = \mathbf{I}$ ,

$$\text{prox}_{\mathcal{C}}(\mathbf{x}) := \text{prox}_{\mathcal{C}}^I(\mathbf{x}). \quad (2.10)$$

If a point  $\mathbf{x}$  is the proximal point to the convex set  $\mathcal{C}$  of the point  $\mathbf{z}$ , then the vector  $\mathbf{z} - \mathbf{x}$  is an element of the normal cone  $\mathcal{N}_{\mathcal{C}}(\mathbf{x})$  as shown in Figure 2.5. This means a normal

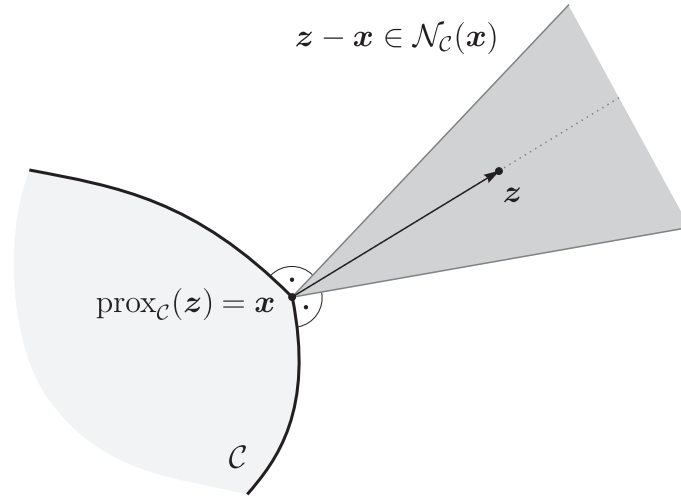


Figure 2.5: Proximal point and normal cone.

cone inclusion can be reformulated equivalently as a proximal point problem,

$$\mathbf{z} - \mathbf{x} \in \mathcal{N}_{\mathcal{C}}(\mathbf{x}) \quad \Leftrightarrow \quad \mathbf{x} = \text{prox}_{\mathcal{C}}(\mathbf{z}). \quad (2.11)$$

In the more general case where the proximal point function  $\text{prox}_{\mathcal{C}}^R(\mathbf{x})$  is used, this relationship becomes

$$\mathbf{y} \in \mathcal{N}_{\mathcal{C}}(\mathbf{x}) \quad \Leftrightarrow \quad \mathbf{x} = \text{prox}_{\mathcal{C}}^R(\mathbf{x} + \mathbf{R}^{-1}\mathbf{y}). \quad (2.12)$$

To obtain this relationship, the proximal point function in the equation  $\mathbf{x} = \text{prox}_{\mathcal{C}}^R(\mathbf{x} + \mathbf{R}^{-1}\mathbf{y})$  is written as a minimization problem restricted to the convex set  $\mathcal{C}$  by using the definition (2.8). Then, the restriction is replaced by adding the indicator function  $\Psi_{\mathcal{C}}(\mathbf{x}^*)$  of the convex set  $\mathcal{C}$  to the objective function of the minimization problem. If  $\mathbf{x}$  is a point at which the minimum of  $f(\mathbf{x}^*)$  is attained, then 0 has to be an element of the subdifferential  $\partial f$ . The subdifferential  $\partial \Psi_{\mathcal{C}}(\mathbf{x})$  of the indicator function is just the normal

cone  $\mathcal{N}_C(\mathbf{x})$  and one obtains the equivalent inclusion  $\mathbf{y} \in \mathcal{N}_C(\mathbf{x})$ :

$$\begin{aligned}
\mathbf{x} &= \text{prox}_C^R(\mathbf{x} + \mathbf{R}^{-1}\mathbf{y}) \\
&\Leftrightarrow \mathbf{x} = \underset{\mathbf{x}^* \in C}{\text{argmin}} \|\mathbf{x} + \mathbf{R}^{-1}\mathbf{y} - \mathbf{x}^*\|_R \\
&\Leftrightarrow \mathbf{x} = \underset{\mathbf{x}^* \in C}{\text{argmin}} \frac{1}{2} \|\mathbf{x} + \mathbf{R}^{-1}\mathbf{y} - \mathbf{x}^*\|_R^2 \\
&\Leftrightarrow \mathbf{x} = \underset{\mathbf{x}^*}{\text{argmin}} \left( \frac{1}{2} \|\mathbf{x} + \mathbf{R}^{-1}\mathbf{y} - \mathbf{x}^*\|_R^2 + \Psi_C(\mathbf{x}^*) \right) \\
&\Leftrightarrow \mathbf{x} = \underset{\mathbf{x}^*}{\text{argmin}} \left( \frac{1}{2} (\mathbf{x} - \mathbf{x}^*)^\top \mathbf{R} (\mathbf{x} - \mathbf{x}^*) + \mathbf{y}^\top (\mathbf{x} - \mathbf{x}^*) + \Psi_C(\mathbf{x}^*) \right) \\
&\quad f(\mathbf{x}^*) := \frac{1}{2} (\mathbf{x} - \mathbf{x}^*)^\top \mathbf{R} (\mathbf{x} - \mathbf{x}^*) + \mathbf{y}^\top (\mathbf{x} - \mathbf{x}^*) + \Psi_C(\mathbf{x}^*) \\
&\Leftrightarrow \mathbf{x} = \underset{\mathbf{x}^*}{\text{argmin}} f(\mathbf{x}^*) \\
&\Leftrightarrow 0 \in \partial f(\mathbf{x}^*) \Big|_{\mathbf{x}^*=\mathbf{x}} \\
&\Leftrightarrow 0 \in (-\mathbf{R}(\mathbf{x} - \mathbf{x}^*) - \mathbf{y} + \partial \Psi_C(\mathbf{x}^*)) \Big|_{\mathbf{x}^*=\mathbf{x}} \\
&\Leftrightarrow 0 \in -\mathbf{y} + \partial \Psi_C(\mathbf{x}) \\
&\Leftrightarrow \mathbf{y} \in \mathcal{N}_C(\mathbf{x}).
\end{aligned} \tag{2.13}$$

Using this relationship, a normal cone inclusion can be formulated as a nonlinear equation, where the matrix  $\mathbf{R}$  can still be chosen freely as long as it is symmetric and positive definite.

## 2.3 Basic Proximal Point Functions

The proximal point functions for one-dimensional convex sets are very simple to evaluate. If the point is in the set, then the proximal point is the point itself, otherwise it is the point on the corresponding boundary. For the convex set  $\mathbb{R}_0^-$  the proximal point function can be written as

$$\text{prox}_{\mathbb{R}_0^-}(x) = \min(x, 0) = \begin{cases} x, & x \leq 0 \\ 0, & x > 0 \end{cases}. \tag{2.14}$$

For the one-dimensional interval  $[a, b]$  one gets the following proximal point function:

$$\text{prox}_{[a,b]}(x) = \max(\min(x, b), a) = \begin{cases} b, & x > b \\ x, & a \leq x \leq b \\ a, & x < a \end{cases}. \tag{2.15}$$

Note that the function  $\text{prox}_{[a,b]}$  also includes the function  $\text{prox}_{\mathbb{R}_0^-}$  for  $a = -\infty$  and  $b = 0$ . An illustration of these functions is shown in Figure 2.6. The Euclidean proximal point function for multidimensional convex sets obtained as Cartesian product of (left- or

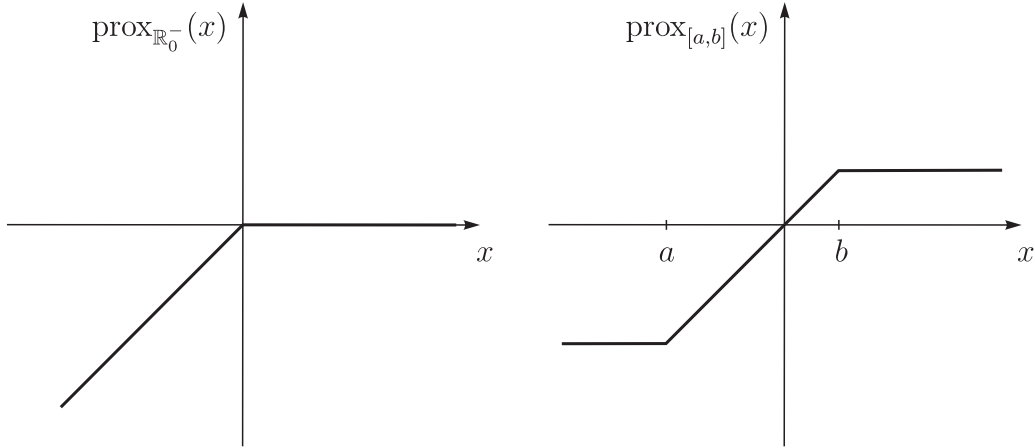


Figure 2.6: The proximal point functions  $\text{prox}_{\mathbb{R}_0^-}$  and  $\text{prox}_{[a,b]}$ .

right-unbounded) one-dimensional intervals can be calculated component-by-component using (2.15). This means, for a multidimensional interval given by

$$[\mathbf{a}, \mathbf{b}] = \{\mathbf{x} \in \mathbb{R}^n \mid a_i \leq x_i \leq b_i\} \quad (2.16)$$

the corresponding proximal point function can be evaluated as

$$\text{prox}_{[\mathbf{a}, \mathbf{b}]}(\mathbf{x}) = \begin{pmatrix} \vdots \\ \text{prox}_{[a_i, b_i]}(x_i) \\ \vdots \end{pmatrix}. \quad (2.17)$$

The  $n$ -dimensional closed Euclidean ball with radius  $r$  can be described as

$$\mathcal{S}_r = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| \leq r\} \quad (2.18)$$

where  $\|\cdot\|$  denotes the Euclidean norm. In the two-dimensional case the set  $\mathcal{S}_r$  is a circular disc. The proximal point function for the set  $\mathcal{S}_r$  is given by

$$\text{prox}_{\mathcal{S}_r}(\mathbf{x}) = \begin{cases} \mathbf{x}, & \|\mathbf{x}\| \leq r \\ \frac{r}{\|\mathbf{x}\|} \mathbf{x}, & \|\mathbf{x}\| > r \end{cases}. \quad (2.19)$$

All the proximal point functions listed in this section are formulated in explicit form yielding an efficient numerical implementation. This is important for the iterative methods presented in Section 2.5.

## 2.4 Transformation of Normal Cone Inclusions

In this section, the transformation of normal cone inclusions and proximal point problems under linear mappings of the underlying convex sets are discussed. For a convex



set  $\mathcal{C} \subseteq \mathbb{R}^n$  and a nonsingular linear mapping  $\mathbf{A} \in \mathbb{R}^{n \times n}$ ,  $\det(\mathbf{A}) \neq 0$  the transformed set  $\mathbf{A}\mathcal{C}$  consists of all the vectors  $\mathbf{A}\mathbf{x}$  with  $\mathbf{x} \in \mathcal{C}$ . The inverse mapped set

$$\mathbf{A}^{-1}\mathcal{C} = \{\mathbf{y} \mid \mathbf{A}\mathbf{y} \in \mathcal{C}\} \quad (2.20)$$

contains all the vectors which are mapped by  $\mathbf{A}$  to the set  $\mathcal{C}$ . The relationship between the normal cone to the set  $\mathcal{C}$  and the normal cone to the set  $\mathbf{A}^{-1}\mathcal{C}$  is given by

$$\mathcal{N}_{\mathcal{C}}(\mathbf{x}) = \mathbf{A}^{-\top} \mathcal{N}_{\mathbf{A}^{-1}\mathcal{C}}(\mathbf{A}^{-1}\mathbf{x}), \quad (2.21)$$

for all  $\mathbf{x} \in \mathcal{C}$ . Of course, (2.21) can be written in the simpler form

$$\mathcal{N}_{\mathcal{C}}(\mathbf{x}) = \mathbf{A}^{\top} \mathcal{N}_{\mathbf{A}\mathcal{C}}(\mathbf{A}\mathbf{x}), \quad (2.22)$$

but (2.21) is used in the following to keep all relations in this section between the set  $\mathcal{C}$  and the inverse mapped set  $\mathbf{A}^{-1}\mathcal{C}$ . The relationship (2.21) can be obtained by injecting the identity matrix  $\mathbf{I} = \mathbf{A}\mathbf{A}^{-1}$  into the definition of the normal cone (2.4) and manipulating the expression such that one part transforms the argument and the set of the normal cone, and the other transforms the normal cone

$$\begin{aligned} \mathcal{N}_{\mathcal{C}}(\mathbf{x}) &= \{\mathbf{y} \mid \mathbf{y}^{\top}(\mathbf{x}^* - \mathbf{x}) \leq 0, \forall \mathbf{x}^* \in \mathcal{C}\} \\ &= \{\mathbf{y} \mid \underbrace{\mathbf{y}^{\top} \mathbf{A} \mathbf{A}^{-1}}_{=: \mathbf{z}^{\top}}(\mathbf{x}^* - \mathbf{x}) \leq 0, \forall \mathbf{x}^* \in \mathcal{C}\} \\ &= \{\mathbf{A}^{-\top} \mathbf{z} \mid \mathbf{z}^{\top}(\mathbf{A}^{-1}\mathbf{x}^* - \mathbf{A}^{-1}\mathbf{x}) \leq 0, \forall \mathbf{A}^{-1}\mathbf{x}^* \in \mathbf{A}^{-1}\mathcal{C}\} \\ &= \mathbf{A}^{-\top} \{\mathbf{z} \mid \mathbf{z}^{\top}(\mathbf{x}^{\circ} - \mathbf{A}^{-1}\mathbf{x}) \leq 0, \forall \mathbf{x}^{\circ} \in \mathbf{A}^{-1}\mathcal{C}\} \\ &= \mathbf{A}^{-\top} \mathcal{N}_{\mathbf{A}^{-1}\mathcal{C}}(\mathbf{A}^{-1}\mathbf{x}). \end{aligned} \quad (2.23)$$

The proximal point function to the convex set  $\mathcal{C}$  in a norm  $\|\cdot\|_R$  can be expressed with the proximal point function to the set  $\mathbf{A}^{-1}\mathcal{C}$  in the norm  $\|\cdot\|_{\mathbf{A}^{\top}R\mathbf{A}}$

$$\text{prox}_{\mathcal{C}}^R(\mathbf{x}) = \mathbf{A} \text{prox}_{\mathbf{A}^{-1}\mathcal{C}}^{\mathbf{A}^{\top}R\mathbf{A}}(\mathbf{A}^{-1}\mathbf{x}). \quad (2.24)$$

This equation can be verified using the definition of the proximal point function (2.8)

$$\begin{aligned} \text{prox}_{\mathcal{C}}^R(\mathbf{x}) &= \underset{\mathbf{x}^* \in \mathcal{C}}{\text{argmin}} \|\mathbf{x} - \mathbf{x}^*\|_R \\ &= \{\mathbf{x}^* \mid \|\mathbf{x} - \mathbf{x}^*\|_R \leq \|\mathbf{x} - \mathbf{y}\|_R, \forall \mathbf{y} \in \mathcal{C}, \mathbf{x}^* \in \mathcal{C}\} \\ &= \{\mathbf{A}\mathbf{x}^{\circ} \mid \|\mathbf{x} - \mathbf{A}\mathbf{x}^{\circ}\|_R \leq \|\mathbf{x} - \mathbf{A}\mathbf{z}\|_R, \forall \mathbf{A}\mathbf{z} \in \mathcal{C}, \mathbf{A}\mathbf{x}^{\circ} \in \mathcal{C}\} \\ &= \mathbf{A} \{\mathbf{x}^{\circ} \mid \|\mathbf{x} - \mathbf{A}\mathbf{x}^{\circ}\|_R \leq \|\mathbf{x} - \mathbf{A}\mathbf{z}\|_R, \forall \mathbf{z} \in \mathbf{A}^{-1}\mathcal{C}, \mathbf{x}^{\circ} \in \mathbf{A}^{-1}\mathcal{C}\} \\ &= \mathbf{A} \underset{\mathbf{x}^{\circ} \in \mathbf{A}^{-1}\mathcal{C}}{\text{argmin}} \|\mathbf{x} - \mathbf{A}\mathbf{x}^{\circ}\|_R \\ &= \mathbf{A} \underset{\mathbf{x}^{\circ} \in \mathbf{A}^{-1}\mathcal{C}}{\text{argmin}} \|\mathbf{A}(\mathbf{A}^{-1}\mathbf{x} - \mathbf{x}^{\circ})\|_R \\ &= \mathbf{A} \underset{\mathbf{x}^{\circ} \in \mathbf{A}^{-1}\mathcal{C}}{\text{argmin}} \|\mathbf{A}^{-1}\mathbf{x} - \mathbf{x}^{\circ}\|_{\mathbf{A}^{\top}R\mathbf{A}} \\ &= \mathbf{A} \text{prox}_{\mathbf{A}^{-1}\mathcal{C}}^{\mathbf{A}^{\top}R\mathbf{A}}(\mathbf{A}^{-1}\mathbf{x}). \end{aligned} \quad (2.25)$$

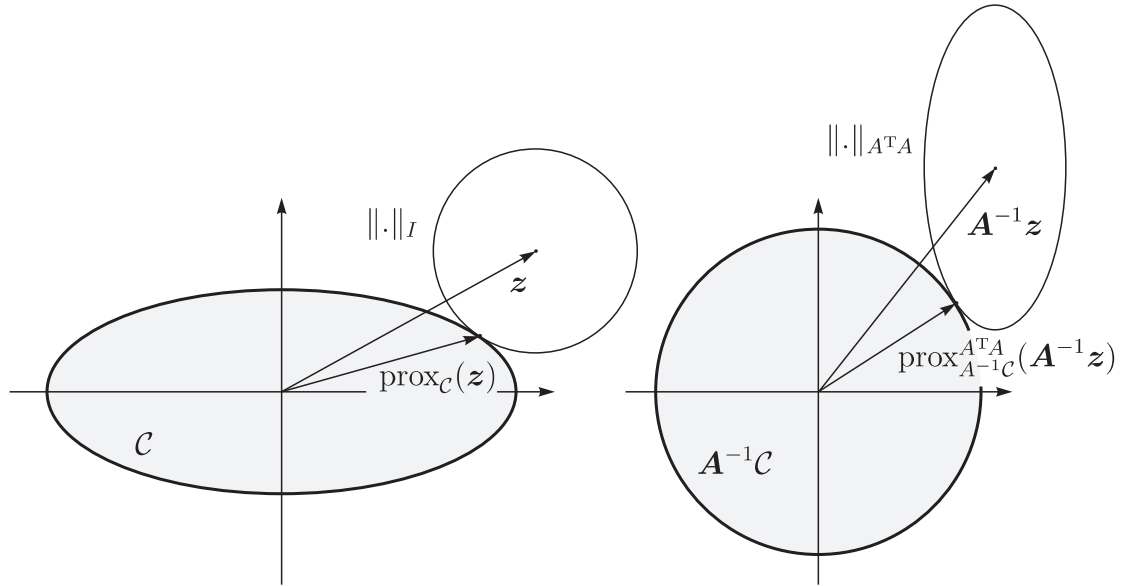


Figure 2.7: Transformation of the proximal point function.

Note that starting with a proximal point function using the Euclidean norm and applying the transformation (2.24) yields a proximal point function which uses no longer the Euclidean norm. This is illustrated for the example of an elliptical set in Figure 2.7. A normal cone inclusion  $\mathbf{y} \in \mathcal{N}_C(\mathbf{x})$  can be transformed into a normal cone inclusion on the inverse mapped set  $\mathbf{A}^{-1}\mathcal{C}$  using (2.21),

$$\mathbf{y} \in \mathcal{N}_C(\mathbf{x}) \quad \Leftrightarrow \quad \mathbf{A}^T \mathbf{y} \in \mathcal{N}_{\mathbf{A}^{-1}\mathcal{C}}(\mathbf{A}^{-1}\mathbf{x}). \quad (2.26)$$

This can be used for example to transform a normal cone inclusion on an elliptical set to a normal cone inclusion on a circular set as illustrated in Figure 2.8.

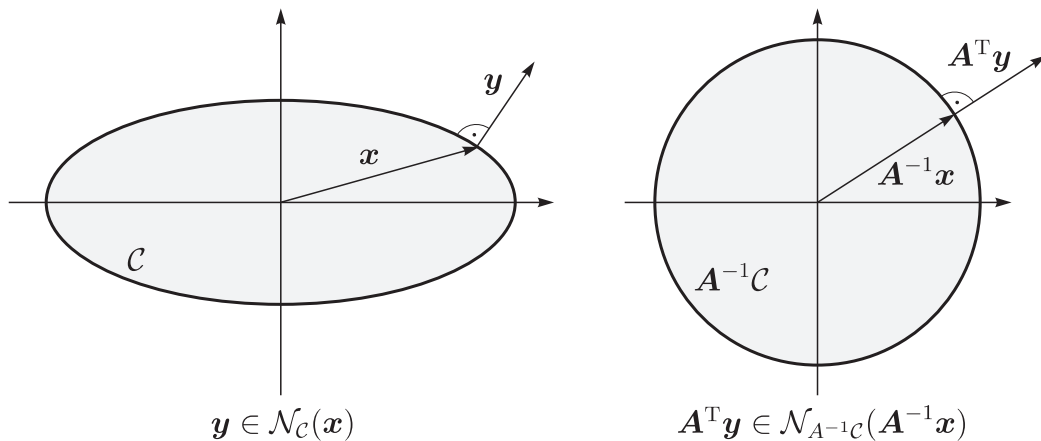


Figure 2.8: Transformation of the normal cone inclusion.

Now, instead of directly rewriting the normal cone inclusion on  $\mathcal{C}$  as a proximal point

equation (cf. Section 2.2),

$$\mathbf{y} \in \mathcal{N}_{\mathcal{C}}(\mathbf{x}) \quad \Leftrightarrow \quad \mathbf{x} = \text{prox}_{\mathcal{C}}^{\mathbf{R}}(\mathbf{x} + \mathbf{R}^{-1}\mathbf{y}), \quad (2.27)$$

the normal cone inclusion using the inverse mapped set  $\mathbf{A}^{-1}\mathcal{C}$  can be rewritten. To achieve this, the right hand side of (2.26) is transformed into a proximal point equation by using the equivalence (2.27) which yields

$$\mathbf{A}^{\top}\mathbf{y} \in \mathcal{N}_{\mathbf{A}^{-1}\mathcal{C}}(\mathbf{A}^{-1}\mathbf{x}) \quad \Leftrightarrow \quad \mathbf{A}^{-1}\mathbf{x} = \text{prox}_{\mathbf{A}^{-1}\mathcal{C}}^{\hat{\mathbf{R}}}(\mathbf{A}^{-1}\mathbf{x} + \hat{\mathbf{R}}^{-1}\mathbf{A}^{\top}\mathbf{y}). \quad (2.28)$$

Here again the matrix  $\hat{\mathbf{R}}$  can be chosen freely as long as it is symmetric and positive definite. If the right hand side of (2.27) is transformed to the inverse mapped set  $\mathbf{A}^{-1}\mathcal{C}$  by using (2.24), one obtains the equation

$$\mathbf{x} = \mathbf{A} \text{prox}_{\mathbf{A}^{-1}\mathcal{C}}^{\mathbf{A}^{\top}\mathbf{R}\mathbf{A}}(\mathbf{A}^{-1}\mathbf{x} + \mathbf{A}^{-1}\mathbf{R}^{-1}\mathbf{y}). \quad (2.29)$$

With the special choice

$$\mathbf{R} = \mathbf{A}^{-\top}\hat{\mathbf{R}}\mathbf{A}^{-1}, \quad (2.30)$$

equation (2.29) yields exactly the right hand side of (2.28). This means that we can transform the original normal cone inclusion  $\mathbf{y} \in \mathcal{N}_{\mathcal{C}}(\mathbf{x})$  either to a normal cone inclusion on the inverse mapped set  $\mathbf{A}^{-1}\mathcal{C}$  and then rewrite it as proximal point equation, or rewrite it directly into an equation and transform then the obtained proximal point problem to the set  $\mathbf{A}^{-1}\mathcal{C}$ . This can be illustrated as:

$$\begin{aligned} \mathbf{y} \in \mathcal{N}_{\mathcal{C}}(\mathbf{x}) & \quad \Leftrightarrow \quad \mathbf{A}^{\top}\mathbf{y} \in \mathcal{N}_{\mathbf{A}^{-1}\mathcal{C}}(\mathbf{A}^{-1}\mathbf{x}) \\ \Downarrow & \quad \quad \quad \Downarrow \\ \mathbf{x} = \text{prox}_{\mathcal{C}}^{\mathbf{R}}(\mathbf{x} + \mathbf{R}^{-1}\mathbf{y}) & \quad \Leftrightarrow \quad \mathbf{A}^{-1}\mathbf{x} = \text{prox}_{\mathbf{A}^{-1}\mathcal{C}}^{\hat{\mathbf{R}}}(\mathbf{A}^{-1}\mathbf{x} + \hat{\mathbf{R}}^{-1}\mathbf{A}^{\top}\mathbf{y}). \end{aligned} \quad (2.31)$$

It has to be noted that the constraint (2.30) on  $\mathbf{R}$  and  $\hat{\mathbf{R}}$  is only required for the identity

$$\text{prox}_{\mathcal{C}}^{\mathbf{R}}(\mathbf{x} + \mathbf{R}^{-1}\mathbf{y}) = \mathbf{A} \text{prox}_{\mathbf{A}^{-1}\mathcal{C}}^{\hat{\mathbf{R}}}(\mathbf{A}^{-1}\mathbf{x} + \hat{\mathbf{R}}^{-1}\mathbf{A}^{\top}\mathbf{y}), \quad (2.32)$$

where just a general proximal point problem is transformed without exploiting the special structure of the normal cone inclusion (2.28) or its equivalent in a proximal point equation. The symmetric and positive definite matrices  $\mathbf{R}$  and  $\hat{\mathbf{R}}$  can be chosen freely in (2.31), because the structure of the normal cone inclusion is available there. By renaming the transformed vectors and sets

$$\mathbf{x} = \mathbf{A}\hat{\mathbf{x}}, \quad \mathcal{C} = \mathbf{A}\hat{\mathcal{C}}, \quad \hat{\mathbf{y}} = \mathbf{A}^{\top}\mathbf{y} \quad (2.33)$$

one can recognize that the structure of the problem remains exactly the same for both the formulations based on  $\mathcal{C}$  and those based on the inverse mapped set  $\mathbf{A}^{-1}\mathcal{C}$ .

$$\begin{aligned} \mathbf{y} \in \mathcal{N}_{\mathcal{C}}(\mathbf{x}) & \quad \Leftrightarrow \quad \hat{\mathbf{y}} \in \mathcal{N}_{\hat{\mathcal{C}}}(\hat{\mathbf{x}}) \\ \Downarrow & \quad \quad \quad \Downarrow \\ \mathbf{x} = \text{prox}_{\mathcal{C}}^{\mathbf{R}}(\mathbf{x} + \mathbf{R}^{-1}\mathbf{y}) & \quad \Leftrightarrow \quad \hat{\mathbf{x}} = \text{prox}_{\hat{\mathcal{C}}}^{\hat{\mathbf{R}}}(\hat{\mathbf{x}} + \hat{\mathbf{R}}^{-1}\hat{\mathbf{y}}) \end{aligned} \quad (2.34)$$

## 2.5 Solving Normal Cone Inclusions

This section describes the solution of normal cone inclusion problems with iterative methods. Beside the normal cone inclusion, there are a few other formulations for non-smooth problems, like the linear complementarity problem for example. And for all the different formulations there is a large variety of solution methods, mostly specialized for certain types of problems. An overview on different formulations and methods is given in [3]. The iterative methods presented in this section are based on the works [6, 47, 48, 55, 78, 80]. In Section 2.5.1 the projected Jacobi iteration for the solution of normal cone inclusion problems with linear equations is described. In Section 2.5.2 the transformation of proximal point iteration problems as presented by the author in [59] is described. This transformation is useful for the solution of normal cone inclusion problems using multidimensional convex sets with a complex shape that can be simplified through a linear transformation. In Section 2.5.3 a modification of the projected Jacobi iteration is presented, which allows to recover the performance of the original Jacobi iteration for problems based on multidimensional convex sets. With a small modification to the projected Jacobi iteration one gets the projected Gauss-Seidel iteration [78, 79], which can be used as well for solving normal cone inclusions with linear equations. The projected Gauss-Seidel iteration is described in Section 2.5.4. For normal cone inclusion problems with nonlinear equations, an iteration scheme based on Newton's method is proposed.

### 2.5.1 Projected Jacobi Iteration

In this section we consider a system of normal cone inclusions and linear equations of the form

$$\boldsymbol{\xi}_i \in \mathcal{N}_{\mathcal{C}_i}(-\boldsymbol{\lambda}_i), \quad \boldsymbol{\xi}_i = \sum_{j \in \mathcal{I}} \mathbf{G}_{ij} \boldsymbol{\lambda}_j + \mathbf{c}_i, \quad i \in \mathcal{I} \quad (2.35)$$

which is to be solved for the unknown vectors  $\boldsymbol{\xi}_i \in \mathbb{R}^{n_i}$  and  $\boldsymbol{\lambda}_i \in \mathbb{R}^{n_i}$  with  $i \in \mathcal{I}$ . For each index in the set  $\mathcal{I}$  there is one normal cone inclusion on the convex set  $\mathcal{C}_i \subseteq \mathbb{R}^{n_i}$  and  $n_i$  linear equations. The matrices  $\mathbf{G}_{ij} \in \mathbb{R}^{n_i \times n_j}$  and the vectors  $\mathbf{c}_i \in \mathbb{R}^{n_i}$  are constant. In this section, the matrix  $\mathbf{G}$  obtained by assembling all  $\mathbf{G}_{ij}$  into one matrix is assumed to be symmetric and positive definite. Note that normal cone inclusions obtained from mechanical applications often have a matrix  $\mathbf{G}$  that is only positive semidefinite, but this is not considered in this section. In order to solve the problem (2.35), the normal cone inclusions are rewritten as proximal point equations by applying (2.12). This yields

$$-\boldsymbol{\lambda}_i = \text{prox}_{\mathcal{C}_i}(-\boldsymbol{\lambda}_i + r_i \boldsymbol{\xi}_i), \quad \boldsymbol{\xi}_i = \sum_{j \in \mathcal{I}} \mathbf{G}_{ij} \boldsymbol{\lambda}_j + \mathbf{c}_i, \quad i \in \mathcal{I} \quad (2.36)$$

where the matrix  $\mathbf{R}_i$  has been chosen as  $\frac{1}{r_i} \mathbf{I}$  to keep the proximal point function  $\text{prox}_{\mathcal{C}_i}$  in the Euclidean norm. All the factors  $r_i$  can be chosen independently in  $\mathbb{R}^+$ . After eliminating  $\boldsymbol{\xi}_i$ , one obtains the system of nonlinear equations

$$-\boldsymbol{\lambda}_i = \text{prox}_{\mathcal{C}_i}(-\boldsymbol{\lambda}_i + r_i \sum_{j \in \mathcal{I}} \mathbf{G}_{ij} \boldsymbol{\lambda}_j + r_i \mathbf{c}_i), \quad i \in \mathcal{I}. \quad (2.37)$$

which has to be solved for the vectors  $\lambda_i$ . Once a solution for  $\lambda_i$  is found, the  $\xi_i$  can be calculated directly from the linear equations in (2.35). The system of nonlinear equations (2.37) can be transformed into the iterative scheme

$$\lambda_i^{\nu+1} = -\text{prox}_{\mathcal{C}_i}(-\lambda_i^\nu + r_i \sum_{j \in \mathcal{I}} G_{ij} \lambda_j^\nu + r_i \mathbf{c}_i), \quad i \in \mathcal{I} \quad (2.38)$$

where the iteration step is denoted by  $\nu$ . Whether this iterative scheme converges to a fixed point depends on  $\mathcal{C}_i$ ,  $G_{ij}$ ,  $\mathbf{c}_i$  and on the factors  $r_i$ . To get convergence for given  $\mathcal{C}_i$ ,  $G_{ij}$  and  $\mathbf{c}_i$ , a good choice of the factors  $r_i$  has to be made. Considering the special case  $\mathcal{C}_i = \mathbb{R}$  the iterative scheme (2.38) reduces to

$$\lambda_i^{\nu+1} = \lambda_i^\nu - r_i \left( \sum_{j \in \mathcal{I}} G_{ij} \lambda_j^\nu + c_i \right), \quad i \in \mathcal{I} \quad (2.39)$$

because  $\text{prox}_{\mathbb{R}}$  is the scalar identity function. The fixed point of the iteration (2.39) is given by the system of linear equations

$$\sum_{j \in \mathcal{I}} G_{ij} \lambda_j + c_i = 0, \quad i \in \mathcal{I}. \quad (2.40)$$

The relaxed Jacobi iteration for solving this system of linear equations is given by

$$\lambda_i^{\nu+1} = \lambda_i^\nu - \frac{\alpha}{G_{ii}} \left( \sum_{j \in \mathcal{I}} G_{ij} \lambda_j^\nu + c_i \right), \quad i \in \mathcal{I} \quad (2.41)$$

where  $\alpha$  is the relaxation parameter. Comparing (2.39) with (2.41) shows that both schemes are identical for  $r_i = \alpha/G_{ii}$ . As shown in [73], the relaxed Jacobi iteration converges if the matrix  $G_{ij}$  is strictly diagonal dominant and  $0 < \alpha \leq 1$ . The idea behind the Jacobi iteration can be seen directly from (2.41): The value

$$e_i^\nu = \sum_{j \in \mathcal{I}} G_{ij} \lambda_j^\nu + c_i \quad (2.42)$$

is the error in the  $i$ -th linear equation in iteration  $\nu$ . If this  $i$ -th equation is dominated by  $G_{ii}$  and  $\lambda_i$ , then the error  $e_i^\nu$  can be reduced by subtracting  $e_i^\nu/G_{ii}$  from  $\lambda_i^\nu$  to get the new approximation  $\lambda_i^{\nu+1}$ . If the matrix is strictly diagonal dominant, then any correction on other lines does not increase the error more than it is lowered by adjusting the diagonal term. Considering the slightly more general case  $\mathcal{C}_i = \mathbb{R}^{n_i}$  of the iterative scheme (2.38) one gets

$$\lambda_i^{\nu+1} = \lambda_i^\nu - r_i \left( \sum_{j \in \mathcal{I}} G_{ij} \lambda_j^\nu + \mathbf{c}_i \right), \quad i \in \mathcal{I} \quad (2.43)$$

Obviously the underlying problem to this iteration scheme is still a system of linear equations, but when comparing with (2.39) we now have only one factor  $r_i$  for the complete  $i$ -th block of equations. If one considers only the special case  $\mathcal{C}_i = \mathbb{R}^{n_i}$  then one could introduce again one  $r_i$  per equation, but this is not possible in the original problem (2.37) while

keeping the Euclidean norm for the  $\text{prox}_{\mathcal{C}_i}$  functions. To make sure that no equation is overcorrected by (2.43), the largest diagonal of the matrix  $\mathbf{G}_{ii}$  has to be used to calculate the factor  $r_i$ , i. e.

$$r_i = \frac{\alpha}{\max_k(\mathbf{G}_{iikk})} \quad (2.44)$$

In this equation  $\mathbf{G}_{iikk}$  denotes the  $k$ -th diagonal entry of the matrix  $\mathbf{G}_{ii}$ . Choosing a correction factor smaller than the one suggested by the non-relaxed Jacobi iteration should not remove convergence, although it is expected to be slower. The key for the convergence of an iterative scheme is the contractivity of the iteration function (cf. Banach's fixed point theorem). When comparing (2.43) with (2.38), the only difference in the iteration function is the subsequent projection in (2.38). Since the proximal point function is non-expansive, the contractivity is not destroyed. This means that the iterative scheme

$$\boldsymbol{\lambda}_i^{\nu+1} = -\text{prox}_{\mathcal{C}_i}(-\boldsymbol{\lambda}_i^\nu + r_i \sum_{j \in \mathcal{I}} \mathbf{G}_{ij} \boldsymbol{\lambda}_j^\nu + r_i \mathbf{c}_i), \quad r_i = \frac{\alpha}{\max_k(\mathbf{G}_{iikk})}, \quad i \in \mathcal{I} \quad (2.45)$$

is expected to converge if there is convergence for the associated Jacobi iteration. For a convergence proof and further background the reader is referred to [78, 55]. In [78, 55] also alternative motivations for the projected Jacobi iteration and the choice of the factors  $r_i$  can be found. Of course, if the iteration (2.45) converges, then one gets a solution to the problem (2.37), which again is equivalent to the original problem (2.35). In practice, relaxation parameters in the range  $0 < \alpha \leq 2$  have shown to be useful. Also the convex sets  $\mathcal{C}_i$  have been assumed constant in this section, weak dependencies of  $\mathcal{C}_i$  on  $\boldsymbol{\lambda}$ , for example caused by a Coulomb friction problem, still yield convergent iterations mostly. To check if convergence has been reached in an implementation of the iteration scheme, one can use for example the stopping criterion

$$\|\boldsymbol{\lambda}_i^{\nu+1} - \boldsymbol{\lambda}_i^\nu\| \leq \varepsilon_{ri} \|\boldsymbol{\lambda}_i^{\nu+1}\| + \varepsilon_{ai}, \quad i \in \mathcal{I} \quad (2.46)$$

where  $\varepsilon_{ri}$  is a relative tolerance and  $\varepsilon_{ai}$  is an absolute tolerance. Note that the relative tolerance  $\varepsilon_{ri}$  is dimensionless, while the absolute tolerance  $\varepsilon_{ai}$  has to be chosen with the same dimensions as the corresponding  $\boldsymbol{\lambda}_i$ .

In this section only proximal point functions based on the Euclidean norm have been considered. While using arbitrary norms for the proximal point functions in (2.37) is possible and can lead to equations which are explicit in  $\boldsymbol{\lambda}_i$ , the evaluation of proximal point functions with respect to arbitrary norms is itself a non-trivial problem [78].

## 2.5.2 Transformation of Proximal Point Iterations

In Section 2.5.1, the projected Jacobi iteration (2.45) for the solution of the normal cone inclusion problem (2.35) has been described. For each step in the projected Jacobi iteration scheme, the proximal point functions  $\text{prox}_{\mathcal{C}_i}$  have to be evaluated. In general, calculating the nearest point in a set involves the solution of a minimization problem or an implicit function. This is not very efficient and it has to be done for every step of the fixed point iteration. The first idea is to transform the proximal point functions

in the equations (2.45) from the sets  $\mathcal{C}_i$  to simpler sets  $\mathbf{A}_i^{-1}\mathcal{C}_i$  with respect to proximal points. For example, if the set  $\mathcal{C}_i$  is an ellipse, then the matrix  $\mathbf{A}_i$  can be chosen such that  $\mathbf{A}_i^{-1}\mathcal{C}_i$  is a circular disc. The Euclidean proximal point function for a circular disc can be implemented very efficiently (see Section 2.3). Unfortunately, the transformation of the proximal point function in 2.5.1 with the help of (2.24) yields the iteration

$$\boldsymbol{\lambda}_i^{\nu+1} = -\mathbf{A}_i \operatorname{prox}_{\mathbf{A}_i^{-1}\mathcal{C}_i}^{\mathbf{A}_i^\top \mathbf{A}_i} (-\mathbf{A}_i^{-1}\boldsymbol{\lambda}_i^\nu + r_i \mathbf{A}_i^{-1} \sum_{j \in \mathcal{I}} \mathbf{G}_{ij} \boldsymbol{\lambda}_j^\nu + r_i \mathbf{A}_i^{-1} \mathbf{c}_i) \quad (2.47)$$

which contains a proximal point function to the simplified set  $\mathbf{A}_i^{-1}\mathcal{C}_i$  but in the  $\|\cdot\|_{\mathbf{A}_i^\top \mathbf{A}_i}$  norm. Evaluating the proximal point function to the set  $\mathbf{A}_i^{-1}\mathcal{C}_i$  in the norm  $\|\cdot\|_{\mathbf{A}_i^\top \mathbf{A}_i}$  is not more efficient than evaluating the original proximal point function to the set  $\mathcal{C}_i$  in the Euclidean norm. The problem here is, that we try to transform a general proximal point problem without using the fact that it originates from a normal cone inclusion. The special structure of the problem can be exploited by applying the transformation directly to the normal cone inclusion and then reformulating the transformed inclusion as a proximal point equation as outlined in Section 2.4. This means that relationship (2.31) is applied to the inclusion (2.35), yielding

$$\boldsymbol{\xi}_i \in \mathcal{N}_{\mathcal{C}_i}(-\boldsymbol{\lambda}_i) \Leftrightarrow -\mathbf{A}_i^{-1}\boldsymbol{\lambda}_i = \operatorname{prox}_{\mathbf{A}_i^{-1}\mathcal{C}_i}^{\hat{\mathbf{R}}_i} (-\mathbf{A}_i^{-1}\boldsymbol{\lambda}_i + \hat{\mathbf{R}}_i^{-1} \mathbf{A}_i^\top \boldsymbol{\xi}_i). \quad (2.48)$$

After choosing  $\hat{\mathbf{R}}_i$  as  $\frac{1}{\hat{r}_i} \mathbf{I}$  one obtains the equation

$$-\boldsymbol{\lambda}_i = \mathbf{A}_i \operatorname{prox}_{\mathbf{A}_i^{-1}\mathcal{C}_i} (-\mathbf{A}_i^{-1}\boldsymbol{\lambda}_i + \hat{r}_i \mathbf{A}_i^\top \sum_j \mathbf{G}_{ij} \boldsymbol{\lambda}_j + \hat{r}_i \mathbf{A}_i^\top \mathbf{c}_i) \quad (2.49)$$

which now uses an Euclidean proximal point function to the inverse mapped set  $\mathbf{A}_i^{-1}\mathcal{C}_i$ . Before iterating this equation for a fixed point it is useful to introduce the definitions

$$\hat{\boldsymbol{\xi}}_i := \mathbf{A}_i^\top \boldsymbol{\xi}_i, \quad \hat{\boldsymbol{\lambda}}_i := \mathbf{A}_i^{-1} \boldsymbol{\lambda}_i, \quad \hat{\mathbf{G}}_{ij} := \mathbf{A}_i^\top \mathbf{G}_{ij} \mathbf{A}_j, \quad \hat{\mathbf{c}}_i := \mathbf{A}_i^\top \mathbf{c}_i, \quad (2.50)$$

simplifying equation (2.49) to

$$-\hat{\boldsymbol{\lambda}}_i = \operatorname{prox}_{\mathbf{A}_i^{-1}\mathcal{C}_i} (-\hat{\boldsymbol{\lambda}}_i + \hat{r}_i \sum_j \hat{\mathbf{G}}_{ij} \hat{\boldsymbol{\lambda}}_j + \hat{r}_i \hat{\mathbf{c}}_i) \quad (2.51)$$

which has again the same structure as equation (2.37). The projected Jacobi iteration scheme then becomes

$$\hat{\boldsymbol{\lambda}}_i^{\nu+1} = -\operatorname{prox}_{\mathbf{A}_i^{-1}\mathcal{C}_i} (-\hat{\boldsymbol{\lambda}}_i^\nu + \hat{r}_i \sum_j \hat{\mathbf{G}}_{ij} \hat{\boldsymbol{\lambda}}_j^\nu + \hat{r}_i \hat{\mathbf{c}}_i), \quad \hat{r}_i = \frac{\alpha}{\max_k(\hat{\mathbf{G}}_{iikk})}, \quad i \in \mathcal{I} \quad (2.52)$$

where  $\hat{\mathbf{G}}_{iikk}$  denotes the  $k$ -th diagonal element of the matrix  $\hat{\mathbf{G}}_{ii}$ . After a solution for  $\hat{\boldsymbol{\lambda}}_i$  has been found, the solution for the original unknown  $\boldsymbol{\lambda}_i$  can be calculated as

$$\boldsymbol{\lambda}_i = \mathbf{A}_i \hat{\boldsymbol{\lambda}}_i. \quad (2.53)$$

Note that the iteration (2.52) together with (2.53) yields the same solution as the non-transformed projected Jacobi iteration (2.45), but only after the iteration has converged. In general, the  $\nu$ -th iteration  $\boldsymbol{\lambda}_i^\nu$  obtained with (2.52) and (2.53) is different from the one obtained with (2.45).

### 2.5.3 Recovering the Jacobi Iteration

The projected Jacobi iteration for the solution of a normal cone inclusion has the form

$$\boldsymbol{\lambda}_i^{\nu+1} = -\text{prox}_{\mathcal{C}_i}(-\boldsymbol{\lambda}_i^\nu + r_i \sum_{j \in \mathcal{I}} \mathbf{G}_{ij} \boldsymbol{\lambda}_j^\nu + r_i \mathbf{c}_i), \quad r_i = \frac{\alpha}{\max_k(\mathbf{G}_{iikk})}, \quad i \in \mathcal{I} \quad (2.54)$$

as described in Section 2.5.1. For every multidimensional convex set  $\mathcal{C}_i$  there is only one factor  $r_i$ , while in the corresponding Jacobi iteration for systems of linear equations one has an individual factor  $r_i$  for each equation. This reduced number of correction scaling factors  $r_i$  can lead to very slow convergence of the iteration in certain situations, when compared with the Jacobi iteration. To show this we assume that the argument of the proximal point function in (2.54) is in the set  $\mathcal{C}_i$ , i. e.

$$-\boldsymbol{\lambda}_i^\nu + r_i \sum_{j \in \mathcal{I}} \mathbf{G}_{ij} \boldsymbol{\lambda}_j^\nu + r_i \mathbf{c}_i \in \mathcal{C}_i. \quad (2.55)$$

This will happen always after some point in the iteration if the actual solution is in the inner of  $\mathcal{C}_i$ . In this case the iteration (2.54) becomes

$$\boldsymbol{\lambda}_i^{\nu+1} = \boldsymbol{\lambda}_i^\nu - \frac{\alpha}{\max_k(\mathbf{G}_{iikk})} \left( \sum_{j \in \mathcal{I}} \mathbf{G}_{ij} \boldsymbol{\lambda}_j^\nu + \mathbf{c}_i \right), \quad i \in \mathcal{I}. \quad (2.56)$$

In cases where the diagonal elements of the matrix  $\mathbf{G}_{ii}$  are of different order of magnitude the convergence of the iteration can be slow compared to the original Jacobi iteration. Lets assume the system consists only of one two-dimensional inclusion with the following values

$$\mathbf{G}_{ii} = \begin{pmatrix} 10^4 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{c}_i = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \alpha = 1. \quad (2.57)$$

The iteration (2.56) requires more than  $10^5$  steps for a solution with a relative accuracy of  $10^{-9}$  starting from zero. The corresponding Jacobi iteration is given by

$$\boldsymbol{\lambda}_i^{\nu+1} = \boldsymbol{\lambda}_i^\nu - \alpha \mathbf{D}_{ii}^{-1} \left( \sum_{j \in \mathcal{I}} \mathbf{G}_{ij} \boldsymbol{\lambda}_j^\nu + \mathbf{c}_i \right), \quad \mathbf{D}_{ii} = \text{diag}(\mathbf{G}_{ii}) \quad i \in \mathcal{I}, \quad (2.58)$$

where  $\mathbf{D}_{ii}$  is the diagonal part of the matrix  $\mathbf{G}_{ii}$ . For the example (2.57), the Jacobi iteration (2.58) needs just one step for the same accuracy starting from zero. To address this slow convergence problem of the projected Jacobi iteration we go back to the proximal point equations

$$-\boldsymbol{\lambda}_i = \text{prox}_{\mathcal{C}_i}(-\boldsymbol{\lambda}_i + r_i \sum_{j \in \mathcal{I}} \mathbf{G}_{ij} \boldsymbol{\lambda}_j + r_i \mathbf{c}_i), \quad i \in \mathcal{I} \quad (2.59)$$

as given in (2.37). Bringing  $-\boldsymbol{\lambda}_i$  in (2.59) from the left to the right, premultiplying everything from the left with the negative of a regular matrix  $\mathbf{Q}_i \in \mathbb{R}^{n_i}$  and adding  $\boldsymbol{\lambda}_i$  to both sides yields the equations

$$\boldsymbol{\lambda}_i = \boldsymbol{\lambda}_i - \mathbf{Q}_i \left( \text{prox}_{\mathcal{C}_i}(-\boldsymbol{\lambda}_i + r_i \sum_{j \in \mathcal{I}} \mathbf{G}_{ij} \boldsymbol{\lambda}_j + r_i \mathbf{c}_i) + \boldsymbol{\lambda}_i \right), \quad i \in \mathcal{I} \quad (2.60)$$



which have exactly the same solution as (2.59). From (2.60) one can construct the iterative scheme

$$\boldsymbol{\lambda}_i^{\nu+1} = \boldsymbol{\lambda}_i^\nu - \mathbf{Q}_i^\nu \left( \text{prox}_{\mathcal{C}_i} \left( -\boldsymbol{\lambda}_i^\nu + r_i \sum_{j \in \mathcal{I}} \mathbf{G}_{ij} \boldsymbol{\lambda}_j^\nu + r_i \mathbf{c}_i \right) + \boldsymbol{\lambda}_i^\nu \right), \quad i \in \mathcal{I} \quad (2.61)$$

which is identical with the projected Jacobi iteration (2.54) for  $\mathbf{Q}_i^\nu = \mathbf{I}$ . If we choose

$$\mathbf{Q}_i^\nu = \frac{\alpha}{r_i} \text{diag}^{-1}(\mathbf{G}_{ii}) \quad (2.62)$$

then (2.61) reduces to the original Jacobi iteration (2.58) for cases where the proximal point function is non-projecting. This idea can be combined into the complete iteration scheme

$$\begin{aligned} \boldsymbol{\lambda}_i^{\nu+1} &= \boldsymbol{\lambda}_i^\nu - \mathbf{Q}_i^\nu \left( \text{prox}_{\mathcal{C}_i} \left( -\boldsymbol{\lambda}_i^\nu + r_i \sum_{j \in \mathcal{I}} \mathbf{G}_{ij} \boldsymbol{\lambda}_j^\nu + r_i \mathbf{c}_i \right) + \boldsymbol{\lambda}_i^\nu \right) \\ \mathbf{Q}_i^\nu &= \begin{cases} \frac{\alpha}{r_i} \text{diag}^{-1}(\mathbf{G}_{ii}), & -\boldsymbol{\lambda}_i^\nu + r_i \sum_{j \in \mathcal{I}} \mathbf{G}_{ij} \boldsymbol{\lambda}_j^\nu + r_i \mathbf{c}_i \in \mathcal{C}_i \\ \mathbf{I}, & \text{else} \end{cases} \\ r_i &= \frac{\alpha}{\max_k(\mathbf{G}_{iikk})}, \end{aligned} \quad (2.63)$$

which recovers the performance of the original Jacobi iteration in the non-projecting case while keeping the properties of the projected Jacobi iteration in projecting situations. And of course, this iteration scheme has the same fixed point as (2.54) since the matrix  $\mathbf{Q}_i$  is always regular. The implementation of this scheme is almost as simple as the original projected Jacobi iteration. From evaluating the proximal point function one knows already, whether the argument is in the convex set or not. If it is outside the convex set, then the iteration step can be finished like the Jacobi iteration. Otherwise the matrix  $\mathbf{Q}_i$  as given by (2.62) has to be applied before finishing the iteration step.

An alternative to the method presented in this section is applying a linear transformation to the proximal point iteration as discussed in Section 2.5.2. With the right transformation one can get a resulting matrix  $\mathbf{G}_{ii}$  where all the diagonal elements are equal and one single factor  $r_i$  is optimal for all components. This method has been suggested in [78]. A drawback of the transformation approach is, that also the convex sets  $\mathcal{C}_i$  get transformed, which can result in more complex proximal point functions. Obviously, transforming the proximal point iteration in order to get equal diagonal entries in  $\mathbf{G}_{ii}$  is not an orthogonal method to transforming the iteration with the goal of simplifying the proximal point functions.

A second alternative to the recovering method (2.63) is to use dynamic factors  $r_i$ . The iteration is started with a value of  $r_i$  based on the largest  $\mathbf{G}_{iikk}$  in a multidimensional proximal point equation. Once the error in the equation associated with the largest  $\mathbf{G}_{iikk}$  is small enough, the factor  $r_i$  could be increased to accelerate the reduction of the error in the next equation. Unfortunately this approach would need a very good strategy for the dynamic adaption of the factors  $r_i$ , in order to avoid running into an overall unstable iteration.

### 2.5.4 Projected Gauss-Seidel Iteration

The Gauss-Seidel method for solving systems of linear equations consists of modified Jacobi iterations, where the updates for each equation are done sequentially and every update uses the most up-to-date values available. This means, if the Jacobi iteration is given by

$$\lambda_i^{\nu+1} = \lambda_i^\nu - \frac{\alpha}{G_{ii}} \left( \sum_{j=1}^n G_{ij} \lambda_j^\nu + c_i \right) \quad (2.64)$$

then the corresponding Gauss-Seidel iteration has the form

$$\lambda_i^{\nu+1} = \lambda_i^\nu - \frac{\alpha}{G_{ii}} \left( \sum_{j=1}^{i-1} G_{ij} \lambda_j^{\nu+1} + \sum_{j=i}^n G_{ij} \lambda_j^\nu + c_i \right). \quad (2.65)$$

As can be seen from (2.65), the calculation of the component  $\lambda_i^{\nu+1}$  uses all the values  $\lambda_j^{\nu+1}$  with  $j < i$ , while the Jacobi iteration relies only on the values from the last iteration. Applying this concept to the proximal point equations

$$-\lambda_i = \text{prox}_{\mathcal{C}_i} \left( -\lambda_i + r_i \sum_{j \in \mathcal{I}} \mathbf{G}_{ij} \lambda_j + r_i \mathbf{c}_i \right), \quad i \in \mathcal{I} \quad (2.66)$$

from Section 2.5.1 yields the projected Gauss-Seidel iteration

$$\lambda_i^{\nu+1} = -\text{prox}_{\mathcal{C}_i} \left( -\lambda_i^\nu + r_i \sum_{j < i} \mathbf{G}_{ij} \lambda_j^{\nu+1} + r_i \sum_{j \geq i} \mathbf{G}_{ij} \lambda_j^\nu + r_i \mathbf{c}_i \right), \quad i \in \mathcal{I}. \quad (2.67)$$

For the factors  $r_i$  the same values

$$r_i = \frac{\alpha}{\max_k (\mathbf{G}_{iikk})} \quad (2.68)$$

as for the projected Jacobi iteration can be used. Of course a converged solution of the projected Gauss-Seidel iteration still solves the proximal point equations (2.66), which themselves are equivalent to the original normal cone inclusion problem (2.35). The convex set transformation described in Section 2.5.2 and the performance recovering strategy described in Section 2.5.3 can be applied as well to the projected Gauss-Seidel iteration. The Gauss-Seidel iteration for linear equations converges on a wider class of matrices. The same is to be expected for the projected Gauss-Seidel iteration. Also experience shows, that the projected Gauss-Seidel iteration usually requires less iteration steps than the projected Jacobi iteration.

### 2.5.5 Projected Newton Iteration

In Section 2.5.1, the solution of normal cone inclusions coupled with linear equations have been discussed. While solving inclusion problems with linear equations can be difficult already, sometimes one is required to solve inclusions coupled with nonlinear equations. In

this section one iterative scheme which can be useful for the solution of systems of normal cone inclusions and nonlinear equations is described. The iterative scheme is based on proximal point iteration and Newton's method. We consider a problem of the form

$$\mathbf{g}_i(\mathbf{x}) \in \mathcal{N}_{\mathcal{C}_i}(-\boldsymbol{\lambda}_i), \quad \mathbf{f}(\mathbf{x}, \boldsymbol{\lambda}) = 0, \quad i \in \mathcal{I} \quad (2.69)$$

which is to be solved for the unknown vectors  $\mathbf{x} \in \mathbb{R}^m$  and  $\boldsymbol{\lambda}_i \in \mathbb{R}^{n_i}$  with  $i \in \mathcal{I}$ . The vector  $\boldsymbol{\lambda}$  is the assembled version of all vectors  $\boldsymbol{\lambda}_i$ . For each index in the set  $\mathcal{I}$  there is one normal cone inclusion on the convex set  $\mathcal{C}_i \subseteq \mathbb{R}^{n_i}$ . The nonlinear equations are given by the function  $\mathbf{f} : \mathbb{R}^{m+\sum n_i} \rightarrow \mathbb{R}^m$ . The function  $\mathbf{f}$  is assumed to be differentiable at least once and to have a nonsingular partial derivative with respect to  $\mathbf{x}$ . As a first step, the normal cone inclusion in (2.69) is reformulated as proximal point equation using (2.12). This yields

$$\boldsymbol{\lambda}_i = -\text{prox}_{\mathcal{C}_i}(-\boldsymbol{\lambda}_i + r_i \mathbf{g}_i(\mathbf{x})), \quad \mathbf{f}(\mathbf{x}, \boldsymbol{\lambda}) = 0, \quad i \in \mathcal{I}. \quad (2.70)$$

For a given approximation  $\boldsymbol{\lambda}^\nu$ , a better approximation for the vector  $\mathbf{x}^\nu$  can be found using Newton's method

$$\mathbf{x}^{\nu+1} = \mathbf{x}^\nu - \left( \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right)^{-1} \mathbf{f}(\mathbf{x}^\nu, \boldsymbol{\lambda}^\nu). \quad (2.71)$$

A detailed discussion of Newton's method can be found in [40]. Coupling the Newton iteration (2.71) with a proximal point iteration yields the projected Newton iteration

$$\begin{cases} \mathbf{x}^{\nu+1} = \mathbf{x}^\nu - \left( \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right)^{-1} \mathbf{f}(\mathbf{x}^\nu, \boldsymbol{\lambda}^\nu) \\ \boldsymbol{\lambda}^{\nu+1} = -\text{prox}_{\mathcal{C}_i}(-\boldsymbol{\lambda}_i^\nu + r_i \mathbf{g}_i(\mathbf{x}^{\nu+1})) \end{cases}. \quad (2.72)$$

For the factor  $r_i$  used in the proximal point iteration, the values

$$r_i = \frac{\alpha}{\max_k \left[ \frac{\partial \mathbf{g}_i}{\partial \mathbf{x}} \left( \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right)^{-1} \frac{\partial \mathbf{f}}{\partial \boldsymbol{\lambda}_i} \right]_{kk}} \quad (2.73)$$

obtained by similar arguments as for the projected Jacobi iteration can be used. Of course, a converged result of (2.72) is also a solution of the original problem (2.69). In contrast to the original Newton method, the convergence is not expected to be quadratic due to the coupling with a proximal point iteration. Therefore it can be a good idea to evaluate the inverse of the Jacobian  $\partial \mathbf{f} / \partial \mathbf{x}$  as well as the factors  $r_i$  only once using the initial values of the iteration, or to use only approximations for the partial derivatives to reduce the numerical costs per iteration. Note that in case of the linear functions

$$\mathbf{f}_i(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{x}_i - \sum_{j \in \mathcal{I}} \mathbf{G}_{ij} \boldsymbol{\lambda}_j - \mathbf{c}_i, \quad \mathbf{g}_i(\mathbf{x}) = \mathbf{x}_i, \quad i \in \mathcal{I} \quad (2.74)$$

the projected Newton iteration reduces exactly to the projected Jacobi iteration. The Newton method has only local convergence, that means in general one gets a convergent

iteration only when starting close enough to the solution. Giving any useful convergence results for the projected Newton iteration is very difficult at best. Nevertheless, for problems close enough to the linear problem discussed in Section 2.5.1 one would expect properties similar to the projected Jacobi iteration. In the projected Newton iteration (2.72), the Newton method is applied only to the differentiable nonlinear equations and not to the proximal point equations. This approach has been chosen, because applying the Newton method to all equations does not yield a good iteration method in general (see also [55]).

## Equations of Motion

---

In this chapter, the building blocks for the formulation of the equations of motion of a non-smooth mechanical system are described. Section 3.1 gives a short summary on the principle of virtual work and the principle of virtual power. The principle of virtual work is then used in Section 3.2 to formulate the equations of motion for a finite-dimensional mechanical model in terms of generalized coordinates. The result are equations close to Lagrange's equations of the second kind, but formulated as variational equations and with an additional force term. For rigid body systems, also an approach based on the projected Newton-Euler equations could be used for formulating the equations of motion in generalized coordinates, but this is not described here. In Section 3.3, the principle of virtual power is used to replace the derivative of the generalized coordinates with a generalized velocity, while optionally adding additional bilateral kinematic constraints to the model. In Section 3.4, the partial derivatives of the kinetic energy occurring in the equations of motion are further evaluated under the assumption of a mass matrix depending only on the generalized coordinates. With the additional assumption of a constant mass matrix with respect to generalized velocities, the inertial terms in the equations of motion are further simplified in Section 3.5. The simplifications made in Section 3.4 and Section 3.5 are in preparation for the construction of energy consistent integrators in Chapter 5. Up to Section 3.6 only the inertial forces and forces from generalized bilateral constraints on displacement and velocity level are considered in detail. All other additional forces are kept in the form of a remaining generalized force. In order to specify any additional forces, different abstract classes of generalized force laws are discussed in Section 3.6. In Section 3.7, a few examples of set-valued force laws of normal cone type are summarized. These set-valued force laws can be used for example to model unilateral contacts or Coulomb friction. The process of equipping the set-valued force laws with an impact law is summarized in Section 3.8. The contribution of different force laws to the total energy balance of a mechanical system is discussed in Section 3.9.

### 3.1 Principles of Virtual Work and Virtual Power

In this section, a short summary of the principle of virtual work and the principle of virtual power is given. For more background on the subject the reader is referred to [69, 25, 15, 34]. The virtual work of an infinite-dimensional mechanical system is classically given by

$$\delta W := \int_{\mathcal{S}} \delta \mathbf{x}^\top (\ddot{\mathbf{x}} dm - d\mathbf{F}), \quad (3.1)$$

where  $dm$  is the mass distribution and  $\mathbf{x} \in \mathbb{R}^3$  describes the displacement at each point of the system  $\mathcal{S}$  as depicted in Figure 3.1. The force distribution acting on the system is described by  $d\mathbf{F} \in \mathbb{R}^3$ .

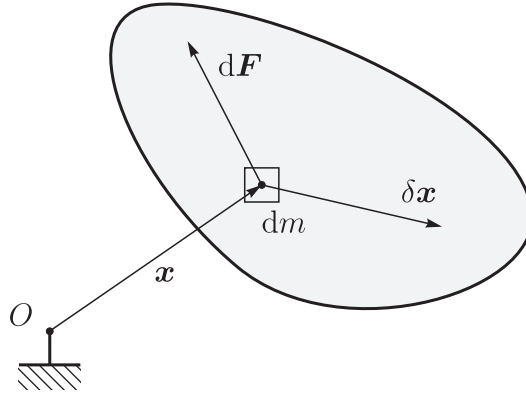


Figure 3.1: Mechanical System  $\mathcal{S}$ .

The virtual displacement  $\delta \mathbf{x}$  of each point is given by

$$\delta \mathbf{x}(t) := \frac{\partial \hat{\mathbf{x}}}{\partial \varepsilon}(t, \varepsilon_0) \delta \varepsilon \quad (3.2)$$

where  $\hat{\mathbf{x}}(t, \varepsilon)$  denotes a family of varied displacements parametrized by  $\varepsilon$ . The actual curve is obtained for  $\varepsilon = \varepsilon_0$ , i.e.

$$\mathbf{x}(t) = \hat{\mathbf{x}}(t, \varepsilon_0). \quad (3.3)$$

Note that the variation in (3.2) is done at fixed time. The equations of motion of the mechanical system  $\mathcal{S}$  can be described with the principle of virtual work

$$\delta W = 0, \quad \forall \delta \mathbf{x}, \quad (3.4)$$

which states that the virtual work  $\delta W$  has to vanish for all virtual displacements  $\delta \mathbf{x}$  at all times. The symbol  $\delta P$  is introduced for the time derivative of the virtual work,

$$\delta P := \frac{d}{dt}(\delta W) = \int_{\mathcal{S}} \delta \dot{\mathbf{x}}^\top (\ddot{\mathbf{x}} dm - d\mathbf{F}) + \delta \mathbf{x}^\top \frac{d}{dt}(\ddot{\mathbf{x}} dm - d\mathbf{F}). \quad (3.5)$$

The equivalent of the equations of motion (3.4) can be formulated by setting  $\delta P = 0$  for all  $\delta \dot{\mathbf{x}}$  with vanishing virtual displacements  $\delta \mathbf{x}$ , i. e.

$$\delta P = 0, \quad \forall \delta \dot{\mathbf{x}} \mid \delta \mathbf{x} = 0. \quad (3.6)$$

One might conclude incorrectly that from  $\delta \mathbf{x} = 0$  also  $\delta \dot{\mathbf{x}} = 0$  has to follow. But all the variations are done at a fixed time, so for each time a different family of varied displacements can be chosen where  $\delta \mathbf{x} = 0$  and  $\delta \dot{\mathbf{x}} \neq 0$ . To simplify (3.6), classically the virtual velocity  $\underline{\delta \dot{\mathbf{x}}}$  is defined as

$$\underline{\delta \dot{\mathbf{x}}} := \delta \dot{\mathbf{x}} |_{\delta \mathbf{x}=0}. \quad (3.7)$$

The virtual velocity  $\underline{\delta \dot{\mathbf{x}}}$  considers only variations of  $\dot{\mathbf{x}}$  where the virtual displacements  $\delta \mathbf{x}$  are still zero. Accordingly the virtual power  $\underline{\delta P}$  of a mechanical system is defined as

$$\underline{\delta P} := \delta P |_{\delta \mathbf{x}=0} = \int_S \underline{\delta \dot{\mathbf{x}}}^\top (\ddot{\mathbf{x}} \, dm - d\mathbf{F}). \quad (3.8)$$

The classical principle of virtual power then reads

$$\underline{\delta P} = 0, \quad \forall \underline{\delta \dot{\mathbf{x}}}, \quad (3.9)$$

which is equivalent to the principle of virtual work (3.4).

## 3.2 Generalized Coordinates

In this section, the equations of motion of a finite-dimensional mechanical model are derived in terms of generalized coordinates. The finite-dimensional model is obtained by imposing perfect bilateral constraints on the infinite-dimensional model of the continuum. The equations of motion of the continuum are given by the principle of virtual work as discussed in Section 3.1:

$$\delta W = \int_S \delta \mathbf{x}^\top (\ddot{\mathbf{x}} \, dm - d\mathbf{F}) = 0, \quad \forall \delta \mathbf{x}. \quad (3.10)$$

The force distribution  $d\mathbf{F}$  is composed of the constraint forces  $d\mathbf{F}_z$  and some remaining forces  $d\mathbf{F}_q$ , i.e.

$$d\mathbf{F} = d\mathbf{F}_z + d\mathbf{F}_q. \quad (3.11)$$

All perfect bilateral geometric constraints acting on a mechanical system may implicitly be taken into account by a minimal parametrization of the  $f$  surviving degrees of freedom via generalized coordinates  $\mathbf{q} \in \mathbb{R}^f$ . If the constraints are bilateral and perfect, they fulfill by definition the principle of d'Alembert-Lagrange which reads

$$\mathbf{x} = \boldsymbol{\xi}(\boldsymbol{\rho}, \mathbf{q}, t), \quad \int_S \delta \mathbf{x}^\top d\mathbf{F}_z = 0, \quad \delta \mathbf{x} = \frac{\partial \boldsymbol{\xi}}{\partial \mathbf{q}} \delta \mathbf{q}, \quad \forall \delta \mathbf{q}. \quad (3.12)$$

The vector  $\boldsymbol{\rho} \in \mathbb{R}^3$  is used to address each point of the system by its position in a reference configuration. The function  $\boldsymbol{\xi}$  returns the actual position of a point addressed by  $\boldsymbol{\rho}$  when the system is displaced according to the generalized coordinates  $\mathbf{q}$  at time  $t$ . The principle of d'Alembert-Lagrange is the force law of the perfect bilateral constraints that restrict the motion to the remaining degrees of freedom taken into account by  $\mathbf{q}$ . It states that the virtual work of the constraint forces has to vanish for any virtual displacements compatible

with the constraint. If the variation  $\delta \mathbf{x}$  is restricted to variations induced by  $\delta \mathbf{q}$ , then the constraint forces  $d\mathbf{F}_z$  disappear in (3.10) and one gets the variational equations of motion

$$\delta \mathbf{q}^\top \int_S \left( \frac{\partial \boldsymbol{\xi}}{\partial \mathbf{q}} \right)^\top \ddot{\boldsymbol{\xi}} dm - \delta \mathbf{q}^\top \underbrace{\int_S \left( \frac{\partial \boldsymbol{\xi}}{\partial \mathbf{q}} \right)^\top d\mathbf{F}_q}_{=: \mathbf{f}_q} = 0, \quad \forall \delta \mathbf{q} \quad (3.13)$$

for the system in terms of the coordinates  $\mathbf{q}$ . Classically, the acceleration terms in (3.13) are reformulated as a difference

$$\left( \frac{\partial \boldsymbol{\xi}}{\partial \mathbf{q}} \right)^\top \ddot{\boldsymbol{\xi}} = \frac{d}{dt} \left[ \left( \frac{\partial \boldsymbol{\xi}}{\partial \mathbf{q}} \right)^\top \dot{\boldsymbol{\xi}} \right] - \frac{d}{dt} \left( \frac{\partial \boldsymbol{\xi}}{\partial \mathbf{q}} \right)^\top \dot{\boldsymbol{\xi}} \quad (3.14)$$

by using Leibniz's law. The partial derivatives of  $\boldsymbol{\xi}$  in (3.14) can be expressed as partial derivative of the absolute velocity,

$$\frac{\partial \boldsymbol{\xi}}{\partial \mathbf{q}} = \frac{\partial \dot{\boldsymbol{\xi}}}{\partial \dot{\mathbf{q}}}, \quad \frac{d}{dt} \left( \frac{\partial \boldsymbol{\xi}}{\partial \mathbf{q}} \right) = \frac{\partial \dot{\boldsymbol{\xi}}}{\partial \mathbf{q}}, \quad (3.15)$$

where the absolute velocity function  $\dot{\boldsymbol{\xi}}$  is given by

$$\dot{\boldsymbol{\xi}}(\boldsymbol{\rho}, \dot{\mathbf{q}}, \mathbf{q}, t) := \frac{\partial \boldsymbol{\xi}}{\partial \dot{\mathbf{q}}} \dot{\mathbf{q}} + \frac{\partial \boldsymbol{\xi}}{\partial t}. \quad (3.16)$$

The left equation in (3.15) follows directly from (3.16). The right equation in (3.15) can be verified by using index notation,

$$\begin{aligned} \left[ \frac{d}{dt} \left( \frac{\partial \boldsymbol{\xi}}{\partial \mathbf{q}} \right) \right]_i &= \frac{d}{dt} \left( \frac{\partial \boldsymbol{\xi}}{\partial q_i} \right) \\ &= \frac{\partial^2 \boldsymbol{\xi}}{\partial q_i \partial q_j} \dot{q}_j + \frac{\partial^2 \boldsymbol{\xi}}{\partial q_i \partial t} \\ &= \frac{\partial}{\partial q_i} \left( \frac{\partial \boldsymbol{\xi}}{\partial q_j} \dot{q}_j + \frac{\partial \boldsymbol{\xi}}{\partial t} \right) \\ &= \frac{\partial \dot{\boldsymbol{\xi}}}{\partial q_i} = \left[ \frac{\partial \dot{\boldsymbol{\xi}}}{\partial \dot{\mathbf{q}}} \right]_i. \end{aligned} \quad (3.17)$$

Combining (3.13), (3.14) and (3.15) yields the variational equations of motion

$$\delta \mathbf{q}^\top \frac{d}{dt} \int_S \left( \frac{\partial \dot{\boldsymbol{\xi}}}{\partial \dot{\mathbf{q}}} \right)^\top \dot{\boldsymbol{\xi}} dm - \delta \mathbf{q}^\top \int_S \left( \frac{\partial \dot{\boldsymbol{\xi}}}{\partial \mathbf{q}} \right)^\top \dot{\boldsymbol{\xi}} dm - \delta \mathbf{q}^\top \mathbf{f}_q = 0, \quad \forall \delta \mathbf{q} \quad (3.18)$$

in terms of the absolute velocity  $\dot{\boldsymbol{\xi}}$  and its derivatives. Instead of using the absolute velocity  $\dot{\boldsymbol{\xi}}$  one can introduce the kinetic energy

$$T(\dot{\mathbf{q}}, \mathbf{q}, t) := \frac{1}{2} \int_S \dot{\boldsymbol{\xi}}^\top \dot{\boldsymbol{\xi}} dm, \quad (3.19)$$



and replace the partial derivatives in (3.18) by partial derivatives of the kinetic energy,

$$\delta \mathbf{q}^\top \frac{d}{dt} \left( \frac{\partial T}{\partial \dot{\mathbf{q}}} \right)^\top - \delta \mathbf{q}^\top \left( \frac{\partial T}{\partial \mathbf{q}} \right)^\top - \delta \mathbf{q}^\top \mathbf{f}_q = 0, \quad \forall \delta \mathbf{q}. \quad (3.20)$$

This variational equation of motion describes the dynamics of a mechanical system with kinematics  $\boldsymbol{\xi}(\boldsymbol{\rho}, \mathbf{q}, t)$  enforced by perfect bilateral constraints of d'Alembert-Lagrange type. The variational equations (3.13) and (3.18) describe exactly the same dynamics as equation (3.20) but in terms of the function  $\boldsymbol{\xi}$  and its derivatives. Equation (3.20) is very close to Lagrange's equations of the second kind, except that it is formulated as a variational equation and it still contains a general force term  $\mathbf{f}_q$ , in which any additional forces can be considered.

### 3.3 Generalized Velocities

In Section 3.2, the equations of motion have been formulated in terms of the generalized coordinates and its derivatives. In this section the classical transition to a formulation making use of a dedicated set of generalized velocities  $\mathbf{u}$  instead of the derivative  $\dot{\mathbf{q}}$  is described. As the introduction of additional perfect bilateral kinematic constraints during this transition is very natural, this more general case will be considered.

The equation of motion (3.20) can be written in the form

$$\delta \mathbf{q}^\top \hat{\mathbf{M}}(\mathbf{q}, t) \ddot{\mathbf{q}} - \delta \mathbf{q}^\top \mathbf{g}_q(\mathbf{q}, \dot{\mathbf{q}}, t) - \delta \mathbf{q}^\top \mathbf{f}_q = 0, \quad \forall \delta \mathbf{q}, \quad (3.21)$$

after evaluating the partial derivatives of the kinetic energy and grouping the resulting terms into the mass matrix  $\hat{\mathbf{M}}(\mathbf{q}, t) \in \mathbb{R}^{f \times f}$  and the vector  $\mathbf{g}_q(\mathbf{q}, \dot{\mathbf{q}}, t) \in \mathbb{R}^f$ . Taking the time derivative of (3.21) and introducing the generalized virtual velocity

$$\underline{\delta \dot{\mathbf{q}}} := \delta \dot{\mathbf{q}} |_{\delta \mathbf{q}=0} \quad (3.22)$$

yields the equations of motion in the form of the principle of virtual power

$$\underline{\delta \dot{\mathbf{q}}}^\top \hat{\mathbf{M}}(\mathbf{q}, t) \ddot{\mathbf{q}} - \underline{\delta \dot{\mathbf{q}}}^\top \mathbf{g}_q(\mathbf{q}, \dot{\mathbf{q}}, t) - \underline{\delta \dot{\mathbf{q}}}^\top \mathbf{f}_q = 0, \quad \forall \underline{\delta \dot{\mathbf{q}}} \quad (3.23)$$

as described in Section 3.1. As next step one defines the bilateral kinematic constraint

$$\dot{\mathbf{q}} = \mathbf{F}(\mathbf{q}, t) \mathbf{u} + \boldsymbol{\beta}(\mathbf{q}, t) \quad (3.24)$$

with a matrix  $\mathbf{F} \in \mathbb{R}^{f \times k}$  and a vector  $\boldsymbol{\beta} \in \mathbb{R}^f$ . The kinematic constraint (3.24) reduces the degrees of freedom on velocity level to the  $k$  degrees of freedom described by the generalized velocities  $\mathbf{u} \in \mathbb{R}^k$ . For the corresponding constraint forces the generalized force  $\mathbf{f}_k$  is introduced

$$\mathbf{f}_q = \mathbf{f}_k + \mathbf{f}_r \quad (3.25)$$

where the vector of remaining forces  $\mathbf{f}_r$  is kept for any additional forces one might like to consider in a later step. The force law for the perfect bilateral kinematic constraint is given by the principle of Jourdain

$$\int_S \underline{\delta \dot{\mathbf{x}}}^\top d\mathbf{F}_k = 0, \quad \forall \underline{\delta \dot{\mathbf{x}}} \text{ admissible}, \quad (3.26)$$

which states that the virtual power of the constraint forces has to vanish for any admissible virtual velocities. In generalized terms the principle of Jourdain reads

$$\underline{\delta \dot{\mathbf{q}}}^\top \mathbf{f}_k = 0, \quad \forall \underline{\delta \dot{\mathbf{q}}} \text{ admissible.} \quad (3.27)$$

The admissible generalized virtual velocities are obtained by evaluating (3.22) with (3.24), i.e.

$$\begin{aligned} \underline{\delta \dot{\mathbf{q}}} &= \delta \dot{\mathbf{q}} \big|_{\delta \mathbf{q}=0} \\ &= (\delta \mathbf{F}(\mathbf{q}, t) \mathbf{u} + \mathbf{F}(\mathbf{q}, t) \delta \mathbf{u} + \delta \boldsymbol{\beta}(\mathbf{q}, t)) \big|_{\delta \mathbf{q}=0} \\ &= \left( \frac{\partial \mathbf{F}}{\partial q_i}(\mathbf{q}, t) \delta q_i \mathbf{u} + \mathbf{F}(\mathbf{q}, t) \delta \mathbf{u} + \frac{\partial \boldsymbol{\beta}}{\partial \mathbf{q}} \delta \mathbf{q} \right) \bigg|_{\delta \mathbf{q}=0} \\ &= \mathbf{F}(\mathbf{q}, t) \delta \mathbf{u}. \end{aligned} \quad (3.28)$$

The complete kinematic constraint is now given by the constraint equation (3.24) and the force law

$$\underline{\delta \dot{\mathbf{q}}}^\top \mathbf{f}_k = 0, \quad \underline{\delta \dot{\mathbf{q}}} = \mathbf{F}(\mathbf{q}, t) \delta \mathbf{u}, \quad \forall \delta \mathbf{u}. \quad (3.29)$$

Reducing the variation in (3.23) to those induced by  $\delta \mathbf{u}$  and replacing all occurrences of  $\underline{\delta \dot{\mathbf{q}}}$  with  $\mathbf{u}$  yields the equation of motion

$$\begin{aligned} \delta \mathbf{u}^\top \underbrace{\mathbf{F}^\top \hat{\mathbf{M}} \mathbf{F}}_{=: \mathbf{M}(\mathbf{q}, t)} \dot{\mathbf{u}} - \delta \mathbf{u}^\top \underbrace{\mathbf{F}^\top \left( \mathbf{g}_q(\mathbf{q}, \mathbf{F} \mathbf{u} + \boldsymbol{\beta}, t) - \hat{\mathbf{M}} \dot{\mathbf{F}} \mathbf{u} - \hat{\mathbf{M}} \dot{\boldsymbol{\beta}} \right)}_{=: \mathbf{g}_u(\mathbf{q}, \mathbf{u}, t)} - \delta \mathbf{u}^\top \underbrace{\mathbf{F}^\top \mathbf{f}_r}_{=: \mathbf{f}_u} = 0, \quad \forall \delta \mathbf{u} \end{aligned} \quad (3.30)$$

for the system with kinematic constraints. Note that the constraint forces disappeared thanks to the principle of Jourdain. Regrouping all classical forces from  $\mathbf{f}_u$  and the gyroscopic accelerations  $\mathbf{g}_u$  into a function  $\mathbf{h}$  while keeping all other forces in  $\mathbf{f}_\lambda$  one gets the equations of motion in the form

$$\mathbf{M}(\mathbf{q}, t) \dot{\mathbf{u}} - \mathbf{h}(\mathbf{q}, \mathbf{u}, t) - \mathbf{f}_\lambda = 0, \quad \dot{\mathbf{q}} = \mathbf{F}(\mathbf{q}, t) \mathbf{u} + \boldsymbol{\beta}(\mathbf{q}, t). \quad (3.31)$$

Of course this formulation also includes the trivial case  $\mathbf{F} = \mathbf{I}$ ,  $\boldsymbol{\beta} = 0$  or the simple mapping with a quadratic and regular matrix  $\mathbf{F}$ , where no kinematic constraints are involved.

### 3.4 Coordinate Dependent Mass Matrix

The equations of motion of a mechanical system can be formulated in generalized coordinates using Lagrange's equations of the second kind with an additional generalized force term  $\mathbf{f}_q$  as described in Section 3.1, i. e.

$$\delta \mathbf{q}^\top \frac{d}{dt} \left( \frac{\partial T}{\partial \dot{\mathbf{q}}} \right)^\top - \delta \mathbf{q}^\top \left( \frac{\partial T}{\partial \mathbf{q}} \right)^\top - \delta \mathbf{q}^\top \mathbf{f}_q = 0 \quad \forall \delta \mathbf{q}. \quad (3.32)$$

To further evaluate the partial derivatives of the kinetic energy we assume in this section, that the kinetic energy is a quadratic form in  $\dot{\mathbf{q}}$  with a symmetric and positive definite

mass matrix  $\mathbf{M}(\mathbf{q}) \in \mathbb{R}^{f \times f}$  depending only on the generalized coordinates. So the kinetic energy function in this section is given by

$$T(\dot{\mathbf{q}}, \mathbf{q}) := \frac{1}{2} \dot{\mathbf{q}}^\top \mathbf{M}(\mathbf{q}) \dot{\mathbf{q}}. \quad (3.33)$$

The partial derivative

$$\left( \frac{\partial T}{\partial \dot{\mathbf{q}}} \right)^\top = \mathbf{M}(\mathbf{q}) \dot{\mathbf{q}} \quad (3.34)$$

can be obtained directly from the quadratic form of the kinetic energy. Inserting (3.34) into (3.32) and evaluating the time derivative yields

$$\delta \mathbf{q}^\top \left( \mathbf{M}(\mathbf{q}) \ddot{\mathbf{q}} + \dot{\mathbf{M}}(\mathbf{q}) \dot{\mathbf{q}} - \left( \frac{\partial T}{\partial \mathbf{q}} \right)^\top - \mathbf{f}_q \right) = 0, \quad \forall \delta \mathbf{q}. \quad (3.35)$$

The velocity dependent inertial forces can be split into a vector of gyroscopic forces and a remaining term expressed as partial derivative of the kinetic energy. Using index notation one gets

$$\begin{aligned} & \left[ \dot{\mathbf{M}}(\mathbf{q}) \dot{\mathbf{q}} - \left( \frac{\partial T}{\partial \mathbf{q}} \right)^\top \right]_k = \\ &= \frac{\partial M_{ki}}{\partial q_j} \dot{q}_j \dot{q}_i - \frac{1}{2} \frac{\partial M_{ji}}{\partial q_k} \dot{q}_j \dot{q}_i \\ &= \underbrace{\left( \frac{\partial M_{ki}}{\partial q_j} - \frac{\partial M_{ji}}{\partial q_k} \right)}_{=: G_{kj}(\dot{\mathbf{q}}, \mathbf{q})} \dot{q}_i \dot{q}_j + \frac{1}{2} \frac{\partial M_{ji}}{\partial q_k} \dot{q}_j \dot{q}_i \\ &= \left[ \mathbf{G}(\dot{\mathbf{q}}, \mathbf{q}) \dot{\mathbf{q}} + \left( \frac{\partial T}{\partial \mathbf{q}} \right)^\top \right]_k. \end{aligned} \quad (3.36)$$

The gyroscopic forces are expressed with the help of the gyro matrix  $\mathbf{G}(\dot{\mathbf{q}}, \mathbf{q}) \in \mathbb{R}^{f \times f}$ . The gyro matrix is skew symmetric,

$$\mathbf{G}(\dot{\mathbf{q}}, \mathbf{q}) = -\mathbf{G}^\top(\dot{\mathbf{q}}, \mathbf{q}), \quad (3.37)$$

as can be seen directly from its definition. This property will be important for the construction of an energy consistent integration scheme in Section 5.2.3. Inserting (3.36) into (3.35) finally yields the variational equation of motion

$$\delta \mathbf{q}^\top \left( \mathbf{M}(\mathbf{q}) \ddot{\mathbf{q}} + \mathbf{G}(\dot{\mathbf{q}}, \mathbf{q}) \dot{\mathbf{q}} + \left( \frac{\partial T}{\partial \mathbf{q}} \right)^\top - \mathbf{f}_q \right) = 0, \quad \forall \delta \mathbf{q} \quad (3.38)$$

for the case of a quadratic kinetic energy with a coordinate dependent mass matrix. After evaluating the variation in (3.38) one gets

$$\mathbf{M}(\mathbf{q}) \ddot{\mathbf{q}} + \mathbf{G}(\dot{\mathbf{q}}, \mathbf{q}) \dot{\mathbf{q}} + \left( \frac{\partial T}{\partial \mathbf{q}} \right)^\top - \mathbf{f}_q = 0 \quad (3.39)$$

as equation of motion.

### 3.5 Decomposed Mass Matrix

Under the assumptions made in Section 3.4 the resulting equations of motion still contain a partial derivative of the kinetic energy. In this section, in order to get an equation of motion without this partial derivative, a decomposition of the kinetic energy of the form

$$T(\dot{\mathbf{q}}, \mathbf{q}) := \frac{1}{2} \dot{\mathbf{q}}^\top \mathbf{Q}^\top(\mathbf{q}) \mathbf{M} \mathbf{Q}(\mathbf{q}) \dot{\mathbf{q}} \quad (3.40)$$

is used. This means, the kinetic energy is not only assumed to be a quadratic form in  $\dot{\mathbf{q}}$  with a symmetric and positive definite mass matrix, but also that a decomposition matrix  $\mathbf{Q}(\mathbf{q}) \in \mathbb{R}^{f \times f}$  is given in analytical form. The matrix  $\mathbf{Q}(\mathbf{q})$  is assumed to be regular while  $\mathbf{M} \in \mathbb{R}^{f \times f}$  is constant, symmetric and positive definite. The matrix  $\mathbf{M}$  is the mass matrix with respect to the generalized velocities  $\mathbf{u} \in \mathbb{R}^f$  defined as

$$\mathbf{u} := \mathbf{Q}(\mathbf{q}) \dot{\mathbf{q}}. \quad (3.41)$$

This becomes obvious if the kinetic energy is expressed in terms of the generalized velocities  $\mathbf{u}$ , i.e.

$$T(\dot{\mathbf{q}}, \mathbf{q}) = \frac{1}{2} \mathbf{u}^\top \mathbf{M} \mathbf{u}. \quad (3.42)$$

The starting point for the derivation of the equations of motion is again Lagrange's equations of the second kind with an additional generalized force term  $\mathbf{f}_q$  (see Section 3.2):

$$\delta \mathbf{q}^\top \frac{d}{dt} \left( \frac{\partial T}{\partial \dot{\mathbf{q}}} \right)^\top - \delta \mathbf{q}^\top \left( \frac{\partial T}{\partial \mathbf{q}} \right)^\top - \delta \mathbf{q}^\top \mathbf{f}_q = 0, \quad \forall \delta \mathbf{q}. \quad (3.43)$$

Evaluating the partial derivative of the kinetic energy with respect to  $\dot{\mathbf{q}}$  yields

$$\left( \frac{\partial T}{\partial \dot{\mathbf{q}}} \right)^\top = \mathbf{Q}^\top(\mathbf{q}) \mathbf{M} \mathbf{Q}(\mathbf{q}) \dot{\mathbf{q}} = \mathbf{Q}^\top(\mathbf{q}) \mathbf{M} \mathbf{u}, \quad (3.44)$$

and after applying the time derivative one gets

$$\frac{d}{dt} \left( \frac{\partial T}{\partial \dot{\mathbf{q}}} \right)^\top = \mathbf{Q}^\top(\mathbf{q}) \mathbf{M} \dot{\mathbf{u}} + \dot{\mathbf{Q}}^\top(\mathbf{q}) \mathbf{M} \mathbf{u} = \mathbf{Q}^\top(\mathbf{q}) \mathbf{M} \dot{\mathbf{u}} + \dot{\mathbf{Q}}^\top(\mathbf{q}) \mathbf{M} \mathbf{Q}(\mathbf{q}) \dot{\mathbf{q}}. \quad (3.45)$$

As a preparation for the next steps, the  $\mathbf{u}$  in the last term of (3.45) has been replaced with the expression in  $\dot{\mathbf{q}}$  once more. Inserting (3.45) into (3.43) results in the following variational equation of motion:

$$\delta \mathbf{q}^\top \left( \mathbf{Q}(\mathbf{q})^\top \mathbf{M} \dot{\mathbf{u}} + \dot{\mathbf{Q}}^\top(\mathbf{q}) \mathbf{M} \mathbf{Q}(\mathbf{q}) \dot{\mathbf{q}} - \left( \frac{\partial T}{\partial \mathbf{q}} \right)^\top - \mathbf{f}_q \right) = 0, \quad \forall \delta \mathbf{q}. \quad (3.46)$$

The velocity dependent inertial forces are now further simplified using index notation

$$\begin{aligned}
& \left[ \dot{\mathbf{Q}}^\top \mathbf{M} \mathbf{Q} \dot{\mathbf{q}} - \left( \frac{\partial T}{\partial \mathbf{q}} \right)^\top \right]_k = \\
& = \frac{\partial Q_{ik}}{\partial q_n} \dot{q}_n M_{ij} Q_{jm} \dot{q}_m - \frac{1}{2} \dot{q}_n \frac{\partial Q_{in}}{\partial q_k} M_{ij} Q_{jm} \dot{q}_m - \frac{1}{2} \dot{q}_m Q_{jm} M_{ji} \frac{\partial Q_{in}}{\partial q_k} \dot{q}_n \\
& = M_{ij} Q_{jm} \dot{q}_m \left( \frac{\partial Q_{ik}}{\partial q_n} - \frac{\partial Q_{in}}{\partial q_k} \right) \dot{q}_n \\
& = M_{ij} u_j \left( \frac{\partial Q_{ik}}{\partial q_n} - \frac{\partial Q_{in}}{\partial q_k} \right) \dot{q}_n \\
& = \underbrace{[\mathbf{M} \mathbf{u}]_i \left( \frac{\partial Q_{ik}}{\partial q_n} - \frac{\partial Q_{in}}{\partial q_k} \right)}_{=: H_{kn}(\mathbf{u}, \mathbf{q})} \dot{q}_n \\
& = [\mathbf{H}(\mathbf{u}, \mathbf{q}) \dot{\mathbf{q}}]_k
\end{aligned} \tag{3.47}$$

The result is a gyro matrix  $\mathbf{H}(\mathbf{u}, \mathbf{q}) \in \mathbb{R}^{f \times f}$  times the derivative  $\dot{\mathbf{q}}$  of the generalized coordinates. With the definition of the function

$$\mathbf{p}(\mathbf{u}, \mathbf{q}) := \mathbf{Q}(\mathbf{q})^\top \mathbf{M} \mathbf{u} \tag{3.48}$$

one can as well write the gyro matrix as difference of two partial derivatives

$$\mathbf{H}(\mathbf{u}, \mathbf{q}) = \left( \frac{\partial \mathbf{p}}{\partial \mathbf{q}} \right) - \left( \frac{\partial \mathbf{p}}{\partial \mathbf{q}} \right)^\top. \tag{3.49}$$

This can be verified by applying index notation

$$\begin{aligned}
H_{kn}(\mathbf{u}, \mathbf{q}) & = \frac{\partial p_k}{\partial q_n} - \frac{\partial p_n}{\partial q_k} \\
& = \frac{\partial Q_{ik}}{\partial q_n} [\mathbf{M} \mathbf{u}]_i - \frac{\partial Q_{in}}{\partial q_k} [\mathbf{M} \mathbf{u}]_i \\
& = [\mathbf{M} \mathbf{u}]_i \left( \frac{\partial Q_{ik}}{\partial q_n} - \frac{\partial Q_{in}}{\partial q_k} \right).
\end{aligned} \tag{3.50}$$

The gyro matrix is skew symmetric,

$$\mathbf{H}(\mathbf{u}, \mathbf{q}) = -\mathbf{H}(\mathbf{u}, \mathbf{q})^\top, \tag{3.51}$$

as can be seen from its definition. Note that the gyro matrix  $\mathbf{H}$  does not directly correspond to the matrix  $\mathbf{G}$  from Section 3.4, although the two objects are similar. In this section, some of the velocity dependent inertial forces are considered as part of the acceleration term already, while at the same time there is no partial derivative of the kinetic energy anymore. Inserting (3.47) into (3.46) yields the following variational equations of motion:

$$\delta \mathbf{q}^\top \left( \mathbf{Q}(\mathbf{q})^\top \mathbf{M} \dot{\mathbf{u}} + \mathbf{H}(\mathbf{u}, \mathbf{q}) \dot{\mathbf{q}} - \mathbf{f}_q \right) = 0, \quad \forall \delta \mathbf{q}, \quad -\mathbf{Q}(\mathbf{q}) \dot{\mathbf{q}} + \mathbf{u} = 0. \tag{3.52}$$

In this form the equation of motion is formulated based on projections associated with  $\mathbf{q}$ , while the acceleration term is formulated using generalized velocities  $\mathbf{u}$ . This results in the matrix  $\mathbf{Q}(\mathbf{q})^\top \mathbf{M}$  in front of the accelerations  $\dot{\mathbf{u}}$  which is not symmetric. Taking the time derivative of (3.52) and setting variations  $\delta\mathbf{q}$  equal to zero yields

$$\underline{\delta\dot{\mathbf{q}}}^\top \left( \mathbf{Q}(\mathbf{q})^\top \mathbf{M} \dot{\mathbf{u}} + \mathbf{H}(\mathbf{u}, \mathbf{q}) \dot{\mathbf{q}} - \mathbf{f}_q \right) = 0, \quad \forall \underline{\delta\dot{\mathbf{q}}}. \quad (3.53)$$

With the help of the inverse

$$\mathbf{F}(\mathbf{q}) := \mathbf{Q}^{-1}(\mathbf{q}) \quad (3.54)$$

of the matrix  $\mathbf{Q}$  one can express the variations  $\underline{\delta\dot{\mathbf{q}}}$  in terms of variations  $\delta\mathbf{u}$ ,

$$\underline{\delta\dot{\mathbf{q}}} = \mathbf{F}(\mathbf{q}) \delta\mathbf{u}. \quad (3.55)$$

Inserting this into (3.53) and replacing any occurrence of  $\dot{\mathbf{q}}$  with  $\mathbf{F}(\mathbf{q})\mathbf{u}$  yields

$$\delta\mathbf{u}^\top \left( \mathbf{M} \dot{\mathbf{u}} + \mathbf{F}^\top(\mathbf{q}) \mathbf{H}(\mathbf{u}, \mathbf{q}) \mathbf{F}(\mathbf{q}) \mathbf{u} - \mathbf{F}^\top(\mathbf{q}) \mathbf{f}_q \right) = 0, \quad \forall \delta\mathbf{u}, \quad -\dot{\mathbf{q}} + \mathbf{F}(\mathbf{q})\mathbf{u} = 0. \quad (3.56)$$

In this formulation based on projections associated with  $\mathbf{u}$  only the symmetric mass matrix  $\mathbf{M}$  is left in front of the accelerations  $\dot{\mathbf{u}}$ . Note that also the matrix  $\mathbf{F}^\top(\mathbf{q}) \mathbf{H}(\mathbf{u}, \mathbf{q}) \mathbf{F}(\mathbf{q})$  is skew symmetric.

## 3.6 Generalized Forces

In this section, three classes of abstract force laws and their integration into the equations of motion are described. The three classes are: Local force laws on displacement level, local force laws on velocity level and scalar potential force laws. For further details on the subject the reader is referred to [25].

### 3.6.1 Local Force Laws on Displacement Level

The class of local force laws on displacement level consists (by definition) of any force law that can be written in the form

$$\mathbf{g}(\mathbf{q}, t) \in \mathcal{D}(-\boldsymbol{\lambda}), \quad \delta W = \delta\mathbf{g}^\top \boldsymbol{\lambda}, \quad (3.57)$$

where  $\delta W$  is the contribution of the force to the virtual work of the system. The basis for the force law is a function  $\mathbf{g}(\mathbf{q}, t)$  of the generalized coordinates  $\mathbf{q}$  and time  $t$ , which measures a displacement in the mechanical model. The displacement has to be understood in a generalized sense, for example it can be a distance or an angle. The force  $\boldsymbol{\lambda}$  is acting in the direction associated with variations of the displacement  $\mathbf{g}$  by setting the virtual work  $\delta W$  of the force equal to the product of  $\boldsymbol{\lambda}$  and the variation  $\delta\mathbf{g}$  of the displacement. The (possibly set-valued) function  $\mathcal{D}$  determines the characteristics of the force law. The generalized force  $\mathbf{f}_q$  required for the integration of the force law into the equations of

motion can be obtained by setting the virtual work of  $\mathbf{f}_q$  equal to the virtual work of  $\boldsymbol{\lambda}$ . One gets

$$\delta W = \delta \mathbf{q}^\top \mathbf{f}_q = \delta \mathbf{g}^\top \boldsymbol{\lambda}, \quad \delta \mathbf{g} = \frac{\partial \mathbf{g}}{\partial \mathbf{q}} \delta \mathbf{q}, \quad \forall \delta \mathbf{q} \quad (3.58)$$

which can be simplified to

$$\delta \mathbf{q}^\top \left( \mathbf{f}_q - \left( \frac{\partial \mathbf{g}}{\partial \mathbf{q}} \right)^\top \boldsymbol{\lambda} \right) = 0, \quad \forall \delta \mathbf{q}. \quad (3.59)$$

Evaluating the variation finally yields the generalized force

$$\mathbf{f}_q = \left( \frac{\partial \mathbf{g}}{\partial \mathbf{q}} \right)^\top \boldsymbol{\lambda}. \quad (3.60)$$

The vector given by

$$\mathbf{w} := \left( \frac{\partial \mathbf{g}}{\partial \mathbf{q}} \right)^\top \quad (3.61)$$

is called the generalized force direction of the scalar force law (see [25]). Accordingly, a matrix of generalized force directions

$$\mathbf{W} := \left( \frac{\partial \mathbf{g}}{\partial \mathbf{q}} \right)^\top \quad (3.62)$$

can be introduced for the multidimensional case, which allows to write the generalized force as

$$\mathbf{f}_q = \mathbf{W} \boldsymbol{\lambda}. \quad (3.63)$$

An example of a force law of this class is show in Figure 3.2. The scalar displacement

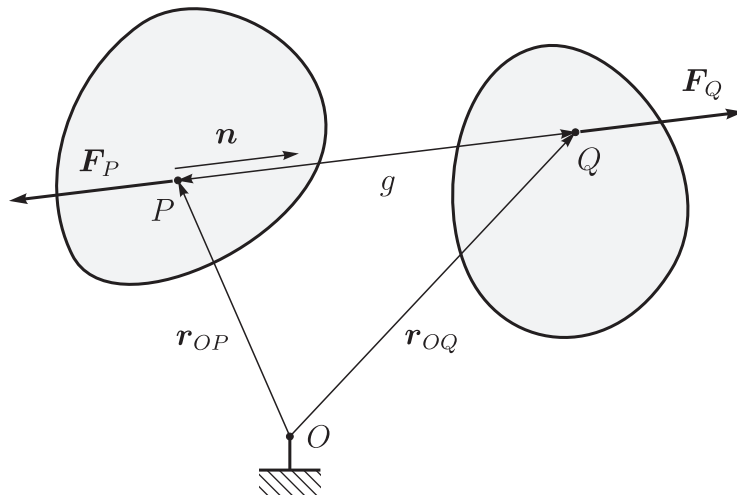


Figure 3.2: Scalar force law.

function  $g(\mathbf{q}, t)$  is the distance between the points  $P$  and  $Q$ . If  $\mathbf{n}(\mathbf{q}, t)$  is a unit vector in the direction of  $\mathbf{r}_{PQ}(\mathbf{q}, t)$  then  $g$  is given by

$$g = \mathbf{n}^\top (\mathbf{r}_{OQ} - \mathbf{r}_{OP}) \quad (3.64)$$

where  $\mathbf{r}_{OP}(\mathbf{q}, t)$  and  $\mathbf{r}_{OQ}(\mathbf{q}, t)$  are the position vectors of  $P$  and  $Q$  respectively. The forces in the points  $P$  and  $Q$  are given by

$$\mathbf{F}_P = -\mathbf{n}\lambda, \quad \mathbf{F}_Q = \mathbf{n}\lambda. \quad (3.65)$$

The force law is completed with a relation between the scalar displacement function and the scalar force

$$g(\mathbf{q}, t) \in \mathcal{D}(-\lambda). \quad (3.66)$$

The contribution of the force  $\lambda$  to the virtual work is equal to

$$\begin{aligned} \delta W &= \int_S \delta \mathbf{x}^\top d\mathbf{F} \\ &= \delta \mathbf{r}_{OQ}^\top \mathbf{n} \lambda - \delta \mathbf{r}_{OP}^\top \mathbf{n} \lambda \\ &= (\delta \mathbf{r}_{OQ} - \delta \mathbf{r}_{OP})^\top \mathbf{n} \lambda. \end{aligned} \quad (3.67)$$

To verify that the force law given by (3.64), (3.65) and (3.66) is of the type (3.57) we first calculate the variation  $\delta g$  of the distance, i. e.

$$\delta g = \frac{\partial g}{\partial \mathbf{q}} \delta \mathbf{q} = \delta \mathbf{n}^\top (\mathbf{r}_{OQ} - \mathbf{r}_{OP}) + \mathbf{n}^\top (\delta \mathbf{r}_{OQ} - \delta \mathbf{r}_{OP}). \quad (3.68)$$

Since  $\mathbf{n}$  is a unit vector and parallel to  $\mathbf{r}_{PQ}$ ,

$$\left. \begin{array}{l} \mathbf{n}^\top \mathbf{n} = 1 \Rightarrow \delta \mathbf{n}^\top \mathbf{n} = 0 \\ \mathbf{n} \parallel \mathbf{r}_{PQ} = \mathbf{r}_{OQ} - \mathbf{r}_{OP} \end{array} \right\} \Rightarrow \delta \mathbf{n}^\top (\mathbf{r}_{OQ} - \mathbf{r}_{OP}) = 0, \quad (3.69)$$

the first summand in (3.68) is equal to zero. For the product  $\delta g \lambda$  we get

$$\delta g \lambda = \mathbf{n}^\top (\delta \mathbf{r}_{OQ} - \delta \mathbf{r}_{OP}) \lambda \quad (3.70)$$

which is exactly the same as we obtained in (3.67). This means, the force law formulated with the displacement (3.64), the forces (3.65) and the inclusion (3.66) is a local force law on displacement level as defined in (3.57).

### 3.6.2 Local Force Laws on Velocity Level

The basis for the local force laws on velocity level is a function  $\dot{\mathbf{g}}(\dot{\mathbf{q}}, \mathbf{q}, t)$  which measures a velocity in the mechanical model. The velocity  $\dot{\mathbf{g}}$  can be a time derivative of some displacement  $\mathbf{g}$ , but an underlying displacement  $\mathbf{g}$  is not a requirement. All force laws that can be written in the form

$$\dot{\mathbf{g}}(\dot{\mathbf{q}}, \mathbf{q}, t) \in \mathcal{D}(-\lambda), \quad \underline{\delta} P = \underline{\delta} \dot{\mathbf{g}}^\top \lambda \quad (3.71)$$



then define the class of local force laws on velocity level. With  $\underline{\delta}P$  the contribution of the force to the virtual power of the system is denoted. The force  $\boldsymbol{\lambda}$  is acting in the direction associated with variations  $\underline{\delta}\dot{\mathbf{g}}$  of the velocity by setting the virtual power  $\underline{\delta}P$  of the force equal to the product of  $\boldsymbol{\lambda}$  and the variation  $\underline{\delta}\dot{\mathbf{g}}$ . The function  $\mathcal{D}$ , which can be set-valued, determines the characteristics of the force law. If a (possibly reduced) vector of generalized velocities  $\mathbf{u}$  is used,

$$\dot{\mathbf{q}} = \mathbf{F}(\mathbf{q}, t)\mathbf{u} + \boldsymbol{\beta}(\mathbf{q}, t), \quad (3.72)$$

then the local velocity function  $\dot{\mathbf{g}}$  can be replaced with the function

$$\boldsymbol{\gamma}(\mathbf{u}, \mathbf{q}, t) = \dot{\mathbf{g}}(\mathbf{F}(\mathbf{q}, t)\mathbf{u} + \boldsymbol{\beta}(\mathbf{q}, t), \mathbf{q}, t) \quad (3.73)$$

which is using  $\mathbf{u}$  as argument. The local force law on velocity level than has the following form:

$$\boldsymbol{\gamma}(\mathbf{u}, \mathbf{q}, t) \in \mathcal{D}(-\boldsymbol{\lambda}), \quad \underline{\delta}P = \underline{\delta}\boldsymbol{\gamma}^\top \boldsymbol{\lambda}. \quad (3.74)$$

To get the generalized force  $\mathbf{f}_q$  required for the integration into the equations of motion, one can set the virtual power contribution of the generalized force  $\mathbf{f}_q$  equal to the virtual power contribution of the force  $\boldsymbol{\lambda}$ , i.e.

$$\underline{\delta}P = \underline{\delta}\dot{\mathbf{q}}^\top \mathbf{f}_q = \underline{\delta}\dot{\mathbf{g}}^\top \boldsymbol{\lambda}, \quad \underline{\delta}\dot{\mathbf{g}} = \frac{\partial \dot{\mathbf{g}}}{\partial \dot{\mathbf{q}}} \underline{\delta}\dot{\mathbf{q}}, \quad \forall \underline{\delta}\dot{\mathbf{q}}. \quad (3.75)$$

Eliminating  $\underline{\delta}\dot{\mathbf{g}}$  from (3.75) yields

$$\underline{\delta}\dot{\mathbf{q}}^\top \left( \mathbf{f}_q - \left( \frac{\partial \dot{\mathbf{g}}}{\partial \dot{\mathbf{q}}} \right)^\top \boldsymbol{\lambda} \right) = 0, \quad \forall \underline{\delta}\dot{\mathbf{q}}, \quad (3.76)$$

from which one gets the generalized force

$$\mathbf{f}_q = \left( \frac{\partial \dot{\mathbf{g}}}{\partial \dot{\mathbf{q}}} \right)^\top \boldsymbol{\lambda} \quad (3.77)$$

after evaluating the variation. For the corresponding matrix of generalized force directions the abbreviation  $\mathbf{W}$  is introduced,

$$\mathbf{W} := \left( \frac{\partial \dot{\mathbf{g}}}{\partial \dot{\mathbf{q}}} \right)^\top. \quad (3.78)$$

which allows to write the generalized force as

$$\mathbf{f}_q = \mathbf{W}\boldsymbol{\lambda}. \quad (3.79)$$

The matrix  $\mathbf{W}$  is the same as the one given in (3.62) in case the function  $\dot{\mathbf{g}}$  is obtained by differentiating a displacement function  $\mathbf{g}$  with respect to time. If the equations of motion are formulated based on  $\delta\mathbf{u}$ , then a corresponding generalized force  $\mathbf{f}_u$  is required for the

integration into the equations of motion. As with  $\mathbf{f}_q$  the contribution of  $\mathbf{f}_u$  to the virtual power is set equal to the contribution of  $\boldsymbol{\lambda}$ ,

$$\underline{\delta}P = \delta\mathbf{u}^\top \mathbf{f}_u = \underline{\delta}\boldsymbol{\gamma}^\top \boldsymbol{\lambda}, \quad \underline{\delta}\boldsymbol{\gamma} = \frac{\partial \boldsymbol{\gamma}}{\partial \mathbf{u}} \delta\mathbf{u}, \quad \forall \delta\mathbf{u}, \quad (3.80)$$

but now by using variations  $\underline{\delta}\boldsymbol{\gamma}$ . After simplifying to

$$\delta\mathbf{u}^\top \left( \mathbf{f}_u - \left( \frac{\partial \boldsymbol{\gamma}}{\partial \mathbf{u}} \right)^\top \boldsymbol{\lambda} \right) = 0, \quad \forall \delta\mathbf{u}, \quad (3.81)$$

and evaluating the variation one gets the generalized force  $\mathbf{f}_u$  based on  $\delta\mathbf{u}$ :

$$\mathbf{f}_u = \left( \frac{\partial \boldsymbol{\gamma}}{\partial \mathbf{u}} \right)^\top \boldsymbol{\lambda} = \mathbf{F}^\top \left( \frac{\partial \dot{\boldsymbol{g}}}{\partial \dot{\mathbf{q}}} \right)^\top \boldsymbol{\lambda} = \mathbf{F}^\top \mathbf{W} \boldsymbol{\lambda}. \quad (3.82)$$

Of course one can replace  $\boldsymbol{\gamma}$  as well with  $\dot{\boldsymbol{g}}$  when calculating the generalized force  $\mathbf{f}_u$  as has been done in the above equation. For the matrix of generalized force directions associated with  $\mathbf{f}_u$  the abbreviation  $\hat{\mathbf{W}}$  is introduced,

$$\hat{\mathbf{W}} := \left( \frac{\partial \boldsymbol{\gamma}}{\partial \mathbf{u}} \right)^\top, \quad (3.83)$$

which allows to write the generalized force as

$$\mathbf{f}_u = \hat{\mathbf{W}} \boldsymbol{\lambda}. \quad (3.84)$$

Note that, in general, the generalized force directions based on  $\mathbf{u}$  are different from those based on  $\dot{\mathbf{q}}$ . If the matrix  $\mathbf{F}$  is square and regular, then the function  $\dot{\boldsymbol{g}}$  can be expressed in terms of  $\boldsymbol{\gamma}$ , i.e.

$$\dot{\boldsymbol{g}}(\dot{\mathbf{q}}, \mathbf{q}, t) = \boldsymbol{\gamma}(\mathbf{F}^{-1}(\dot{\mathbf{q}} - \boldsymbol{\beta}), t). \quad (3.85)$$

This can be used to express the generalized force  $\mathbf{f}_q$  in terms of the function  $\boldsymbol{\gamma}$ . Inserting the function (3.85) into (3.77) yields

$$\mathbf{f}_q = \mathbf{F}^{-\top} \left( \frac{\partial \boldsymbol{\gamma}}{\partial \mathbf{u}} \right)^\top \boldsymbol{\lambda}. \quad (3.86)$$

Usually, the velocity functions  $\dot{\boldsymbol{g}}$  and  $\boldsymbol{\gamma}$  are linear in  $\dot{\mathbf{q}}$  and  $\mathbf{u}$  respectively. This means they can be written in the form

$$\begin{aligned} \dot{\boldsymbol{g}}(\dot{\mathbf{q}}, \mathbf{q}, t) &= \mathbf{W}^\top(\mathbf{q}, t) \dot{\mathbf{q}} + \boldsymbol{\chi}(\mathbf{q}, t), \\ \boldsymbol{\gamma}(\mathbf{u}, \mathbf{q}, t) &= \hat{\mathbf{W}}^\top(\mathbf{q}, t) \mathbf{u} + \hat{\boldsymbol{\chi}}(\mathbf{q}, t), \end{aligned} \quad (3.87)$$

using the definitions (3.78) and (3.83).

### 3.6.3 Potential Force Laws

The class of potential force laws consists of any force law that can be written in the form

$$\mathbf{f}_q = - \left( \frac{\partial V}{\partial \mathbf{q}} \right)^\top, \quad (3.88)$$

where  $\mathbf{f}_q$  is the associated generalized force and  $V(\mathbf{q}, t)$  the scalar potential function. An example for a potential force law is the gravitational force

$$d\mathbf{F} = \mathbf{g} dm \quad (3.89)$$

where  $\mathbf{g}$  is the gravitational acceleration and  $dm$  the mass distribution. An illustration of the force law is shown in Figure 3.3. Using the definition (3.13), the generalized force

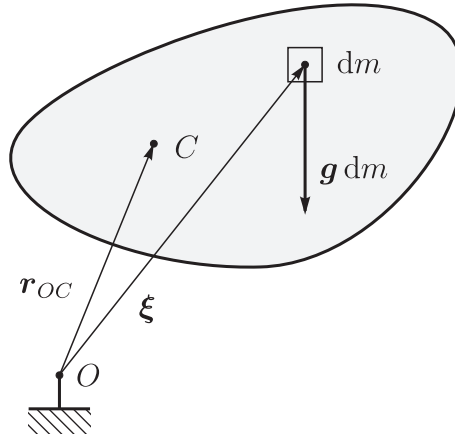


Figure 3.3: Gravitation.

becomes

$$\mathbf{f}_q = \int_S \left( \frac{\partial \boldsymbol{\xi}}{\partial \mathbf{q}} \right)^\top \mathbf{g} dm. \quad (3.90)$$

We can verify that this is a potential force law with the potential function

$$V(\mathbf{q}, t) = -\mathbf{g}^\top \int_S \boldsymbol{\xi} dm. \quad (3.91)$$

Inserting (3.91) into (3.88) yields exactly the same generalized force as calculated in (3.90). The potential  $V$  associated with the gravitational force can be written also in terms of the position  $\mathbf{r}_{OC}$  of the center of mass  $C$  and the mass  $m$  of the system,

$$\mathbf{r}_{OC}(\mathbf{q}, t) = \frac{1}{m} \int_S \boldsymbol{\xi} dm, \quad m = \int_S dm. \quad (3.92)$$

The potential function for the gravity force law then reads

$$V(\mathbf{q}, t) = -m \mathbf{g}^\top \mathbf{r}_{OC}(\mathbf{q}, t). \quad (3.93)$$

### 3.7 Set-Valued Force Laws of Normal Cone Type

Set-valued force laws of normal cone type can be used to describe many of the important non-smooth interactions like unilateral contact, one-dimensional Coulomb friction or spatial Coulomb friction [25, 29]. At the same time it can describe also classical interactions like perfect bilateral constraints. A set-valued force law of normal cone type on displacement level has the form of the normal cone inclusion

$$\mathbf{g} \in \mathcal{N}_{\mathcal{C}}(-\boldsymbol{\lambda}), \quad \delta W = \delta \mathbf{g}^T \boldsymbol{\lambda}, \quad (3.94)$$

where  $\mathbf{g}$  is a displacement function and  $\boldsymbol{\lambda}$  its associated force. It is a special case of the force laws on displacement level (3.57). A set-valued force law of normal cone type on velocity level makes use of a velocity function  $\boldsymbol{\gamma}$  and has the form

$$\boldsymbol{\gamma} \in \mathcal{N}_{\mathcal{C}}(-\boldsymbol{\lambda}), \quad \underline{\delta} P = \underline{\delta} \boldsymbol{\gamma}^T \boldsymbol{\lambda}. \quad (3.95)$$

This is a special case of the force laws on velocity level (3.74). The characteristics of a force law of normal cone type is given by its convex set  $\mathcal{C}$ . The set  $\mathcal{C}$  is to be seen as a force reservoir containing all values that  $-\boldsymbol{\lambda}$  is allowed to assume. When  $-\boldsymbol{\lambda}$  is in the interior of  $\mathcal{C}$  then  $\mathbf{g}$  or  $\boldsymbol{\gamma}$  respectively is equal to zero and the force law acts as an ideal constraint. When  $-\boldsymbol{\lambda}$  reaches the boundary of  $\mathcal{C}$  then  $\mathbf{g}$  or  $\boldsymbol{\gamma}$  respectively is allowed to assume a value different from zero, such that  $-\boldsymbol{\lambda}$  can remain in the set  $\mathcal{C}$ .

In Section 3.7.1 the formulation of a bilateral constraint as a normal cone inclusion is shown. The set-valued force law for unilateral contacts is discussed in Section 3.7.2. In Section 3.7.3, Section 3.7.4 and Section 3.7.5 one- and multidimensional force laws for Coulomb friction are described.

#### 3.7.1 Bilateral Constraint

The force law for perfect bilateral constraints can be expressed as normal cone inclusion. If the convex set  $\mathcal{C}$  forming the force reservoir is given by all real numbers

$$\mathcal{C} = \mathbb{R}, \quad (3.96)$$

then the normal cone  $\mathcal{N}_{\mathcal{C}}$  reduces to the set containing just zero,

$$\mathcal{N}_{\mathcal{C}} = \mathcal{N}_{\mathbb{R}} = \{0\}. \quad (3.97)$$

This can be used to write the bilateral constraint on displacement level as a normal cone inclusion

$$g = 0, \quad \lambda \in \mathbb{R} \quad \Leftrightarrow \quad g \in \mathcal{N}_{\mathbb{R}}(-\lambda), \quad (3.98)$$

where  $g$  is the constraint function. Of course the same can be done on velocity level as well

$$\boldsymbol{\gamma} = 0, \quad \boldsymbol{\lambda} \in \mathbb{R} \quad \Leftrightarrow \quad \boldsymbol{\gamma} \in \mathcal{N}_{\mathbb{R}}(-\boldsymbol{\lambda}). \quad (3.99)$$

The kinematic constraint is described by the function  $\boldsymbol{\gamma}$  in this case. The virtual work as given in (3.94) is equal to zero for any admissible virtual displacements, which is exactly the principle of d'Alembert-Lagrange. Any admissible virtual velocity yields a virtual power equal to zero in (3.95), which is the principle of Jourdain.

### 3.7.2 Unilateral Contact

The force law of a perfect unilateral contact can be formulated as complementarity

$$g_N \geq 0, \quad \lambda_N \geq 0, \quad g_N \lambda_N = 0 \quad (3.100)$$

between the scalar gap function  $g_N$  and the scalar contact force  $\lambda_N$ . The gap function  $g_N$  measures the distance between the two objects that are potentially in contact. The gap function  $g_N$  and the contact force  $\lambda_N$  can be introduced as discussed in Section 3.6.1. An example of a contact between two bodies is shown in Figure 3.4.

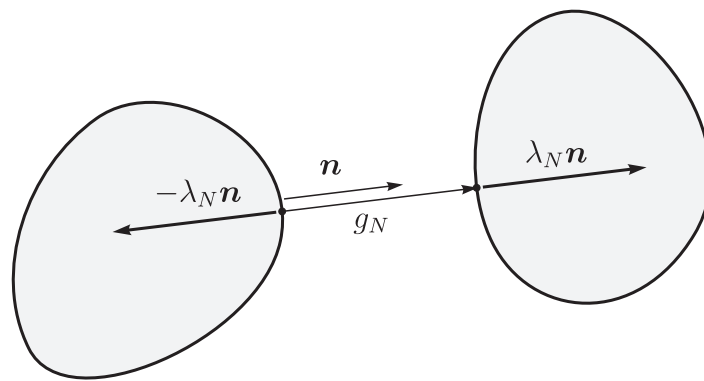


Figure 3.4: Contact between two bodies.

A closed contact has a gap function  $g_N$  equal to zero, while the contact force  $\lambda_N$  can assume any positive value. The force  $\lambda_N$  of a closed contact prevents the gap function from decreasing to values below zero. Opening the contact to  $g_N > 0$  results in a contact force equal to zero. This set-valued force law is illustrated in Figure 3.5.

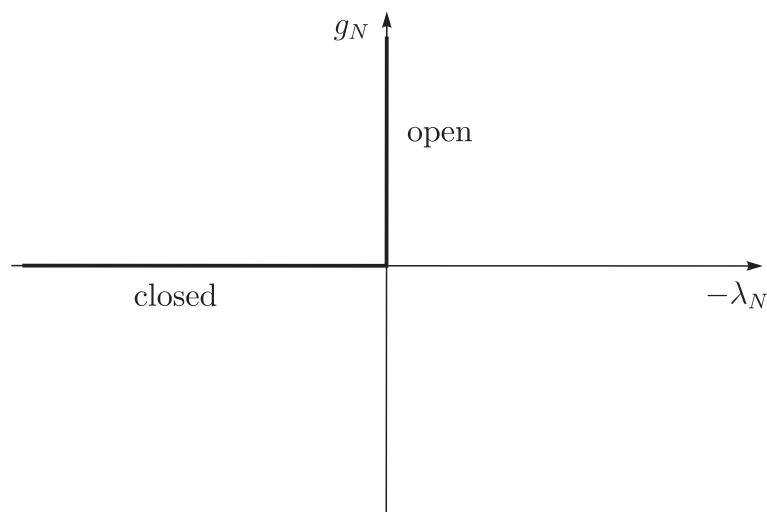


Figure 3.5: Unilateral contact.

For a normal cone based formulation of the force law the convex set

$$\mathcal{C}_N = \mathbb{R}_0^- \quad (3.101)$$

containing all non-positive real numbers is used. Note that the force reservoir  $\mathcal{C}_N$  contains all possible values of  $-\lambda_N$ . Since  $\lambda_N$  is always positive, the set of all  $-\lambda_N$  is  $\mathbb{R}_0^-$ . With the set  $\mathcal{C}_N$  we can now formulate the force law of a unilateral contact as normal cone inclusion

$$g_N \in \mathcal{N}_{\mathcal{C}_N}(-\lambda_N), \quad (3.102)$$

which is identical to the complementarity formulation (3.100). The force law (3.102) can be expressed also with a normal cone inclusion on velocity level

$$\begin{aligned} g_N > 0 : \lambda_N &= 0 \\ g_N = 0 : \gamma_N &\in \mathcal{N}_{\mathcal{C}_N}(-\lambda_N) \end{aligned} \quad (3.103)$$

as derived in [25]. Note that one still needs to distinguish whether the contact is open on displacement level. A force law which contains only the inclusion on velocity level from (3.103) would correspond to the force law of a sprag clutch.

### 3.7.3 Coulomb Friction

One dimensional Coulomb friction can be described using a scalar tangential relative velocity  $\gamma_T$  and a scalar tangential friction force  $\lambda_T$ . With a scalar normal force  $\lambda_N$  and a friction coefficient  $\mu$  the force reservoir for  $-\lambda_T$  is given by

$$\mathcal{C}_T = [-\mu \lambda_N, \mu \lambda_N]. \quad (3.104)$$

If the friction force  $\lambda_T$  is in the inner of the set  $\mathcal{C}_T$  then the relative velocity is zero (sticking). Sliding occurs once the force  $\lambda_T$  reaches the upper or lower boundary of the set  $\mathcal{C}_T$ . With a value  $\lambda_T = \mu \lambda_N$  one gets sliding with a negative relative velocity, while sliding with a positive relative velocity is obtained for  $\lambda_T = -\mu \lambda_N$ . This Coulomb friction law can be described with the force law

$$\gamma_T \in \mathcal{N}_{\mathcal{C}_T}(-\lambda_T) \quad (3.105)$$

of normal cone type. The graph of this force law is depicted in Figure 3.6. The normal force  $\lambda_N$  used in the Coulomb friction law can be a constant normal force or the normal force associated with a unilateral contact (see Section 3.7.2). The combined force law for the unilateral contact with friction then becomes

$$g_N \in \mathcal{N}_{\mathcal{C}_N}(-\lambda_N), \quad \gamma_T \in \mathcal{N}_{[-\mu\lambda_N, \mu\lambda_N]}(-\lambda_T). \quad (3.106)$$

If the unilateral contact is formulated on velocity level one gets

$$\begin{aligned} g_N > 0 : \lambda_N &= 0, \quad \lambda_T = 0 \\ g_N = 0 : \gamma_N &\in \mathcal{N}_{\mathcal{C}_N}(-\lambda_N), \quad \gamma_T \in \mathcal{N}_{\mathcal{C}_T(\lambda_N)}(-\lambda_T). \end{aligned} \quad (3.107)$$

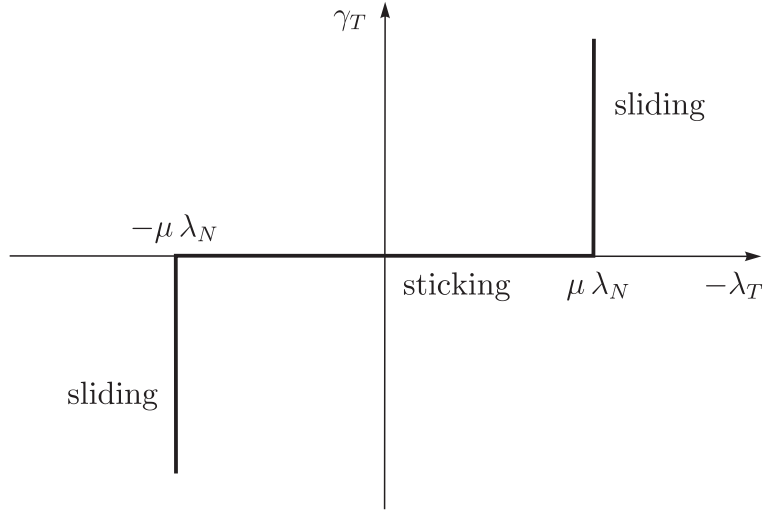


Figure 3.6: Coulomb friction.

Note that when combining the force reservoirs in normal and tangential direction into the convex set

$$\mathcal{C}_{NT} = \{(\lambda_N, \lambda_T)^\top \mid \lambda_N \geq 0, |\lambda_T| \leq \mu \lambda_N\} \quad (3.108)$$

and formulating a force law with the corresponding normal cone inclusion

$$\begin{aligned} g_N > 0 : (\lambda_N, \lambda_T)^\top &= 0 \\ g_N = 0 : (\gamma_N, \gamma_T)^\top &\in \mathcal{N}_{\mathcal{C}_{NT}}(-(\lambda_N, \lambda_T)^\top) \end{aligned} \quad (3.109)$$

is not identical with (3.107). To get a correct formulation based on the set  $\mathcal{C}_{NT}$ , a nonlinear transformation has to be added to the relative velocities. As derived in [20], the force law

$$\begin{aligned} g_N > 0 : (\lambda_N, \lambda_T)^\top &= 0 \\ g_N = 0 : (\gamma_N + \mu |\gamma_T|, \gamma_T)^\top &\in \mathcal{N}_{\mathcal{C}_{NT}}(-(\lambda_N, \lambda_T)^\top) \end{aligned} \quad (3.110)$$

is identical with (3.107) and correctly describes a normal contact with Coulomb friction. In the formulation (3.110) a constant convex set  $\mathcal{C}_{NT}$  and a nonlinear transformation of the relative velocity is used. In contrast, the formulation (3.107) uses directly the relative velocity, but the set  $\mathcal{C}_T(\lambda_N)$  is not constant. An illustration of the formulations (3.107) and (3.110) is shown in Figure 3.7.

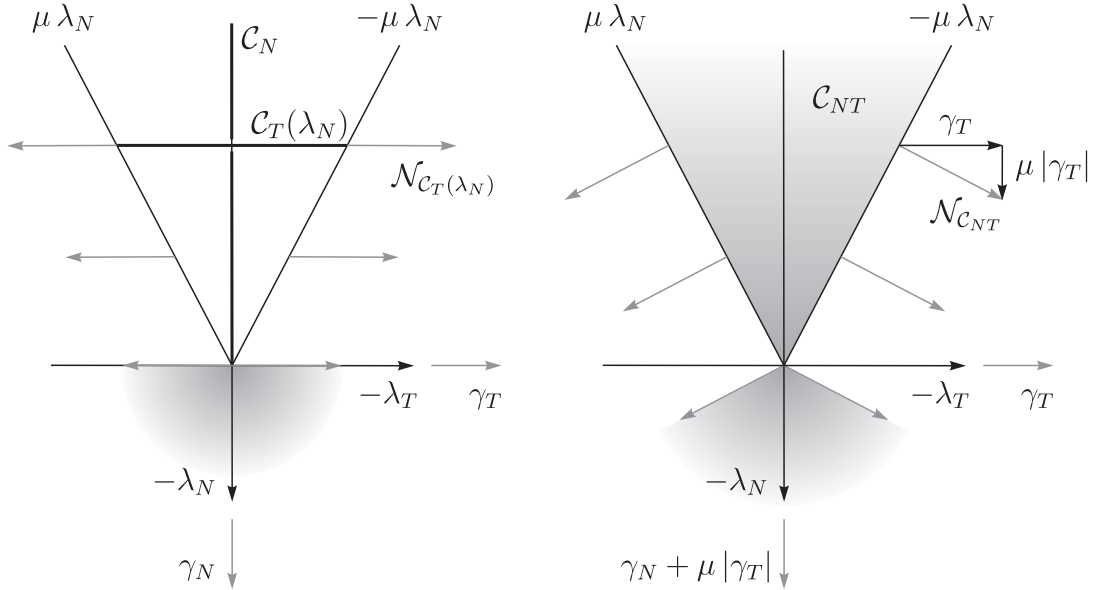


Figure 3.7: Formulations of the force law for normal contact with Coulomb friction.

### 3.7.4 Spatial Coulomb Friction

The force law for spatial Coulomb friction can be used to model the frictional contact of a point on a plane when there is no prescribed direction for the frictional interaction. With the one dimensional friction law described in section 3.7.3 the direction of the friction is known a priori and is not determined by the force law itself. Since the direction is not known, two orthogonal relative velocities  $\gamma_{T1}$  and  $\gamma_{T2}$  are introduced, measuring the two components of the relative velocity in the tangent plane. And for each relative velocity there is an associated tangential force  $\lambda_{T1}$  and  $\lambda_{T2}$  respectively. Both components are collected into two-dimensional vectors,

$$\boldsymbol{\lambda}_T := \begin{pmatrix} \lambda_{T1} \\ \lambda_{T2} \end{pmatrix}, \quad \boldsymbol{\gamma}_T := \begin{pmatrix} \gamma_{T1} \\ \gamma_{T2} \end{pmatrix}. \quad (3.111)$$

The force reservoir for the spatial Coulomb friction law is given by the circular disk with radius  $\mu \lambda_N$ ,

$$\mathcal{C}_T = \{ \boldsymbol{\lambda}_T \in \mathbb{R}^2 \mid \| \boldsymbol{\lambda}_T \| \leq \mu \lambda_N \}, \quad (3.112)$$

where  $\lambda_N$  is the normal force and  $\mu$  the friction coefficient. The contact is sticking if the norm of the force vector  $\boldsymbol{\lambda}_T$  is smaller than  $\mu \lambda_N$ . If the force is on the boundary of  $\mathcal{C}_T$  then the contact is sliding into the direction of  $-\boldsymbol{\lambda}_T$ . This behavior can be described with the following normal cone inclusion:

$$\boldsymbol{\gamma}_T \in \mathcal{N}_{\mathcal{C}_T}(-\boldsymbol{\lambda}_T). \quad (3.113)$$

An illustration of the spatial Coulomb friction law is shown in Figure 3.8.



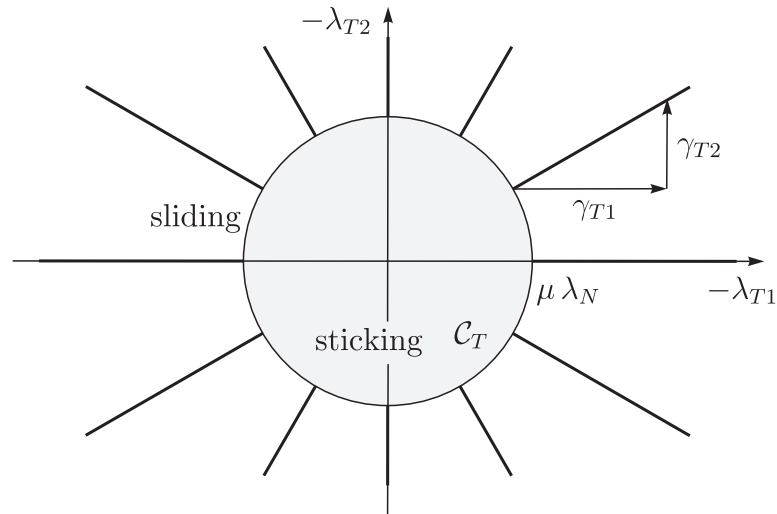


Figure 3.8: Spatial Coulomb friction.

For an anisotropic Coulomb friction law, the friction force reservoir (3.112) can be replaced with an ellipse

$$C_T = \left\{ \begin{pmatrix} \lambda_{T1} \\ \lambda_{T2} \end{pmatrix} \mid \left( \frac{\lambda_{T1}}{\mu_1 \lambda_N} \right)^2 + \left( \frac{\lambda_{T2}}{\mu_2 \lambda_N} \right)^2 \leq 1 \right\} \quad (3.114)$$

with the semi-axes  $\mu_1 \lambda_N$  and  $\mu_2 \lambda_N$ . In this case one gets a friction coefficient  $\mu_1$  in the direction of  $\gamma_{T1}$  and a friction coefficient  $\mu_2$  in the direction of  $\gamma_{T2}$ . Note that in this case, the sliding direction is not always identical with the direction of  $-\lambda_T$ . The force law for anisotropic Coulomb friction is depicted in Figure 3.9.

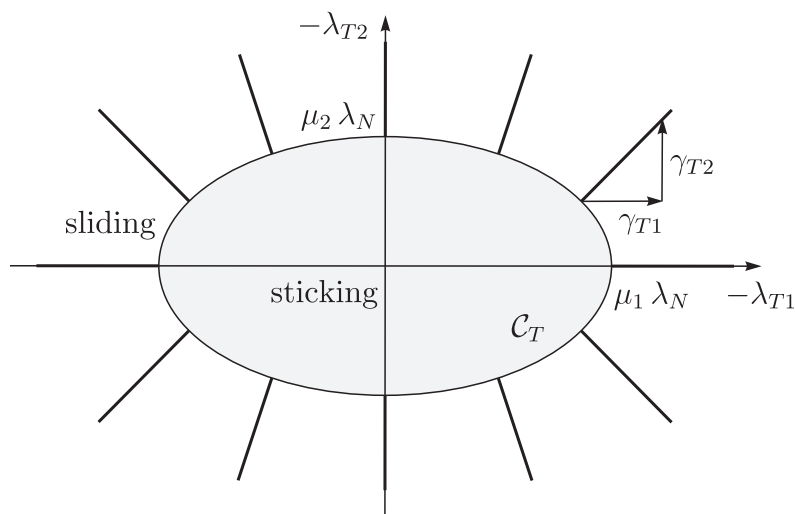


Figure 3.9: Anisotropic spatial Coulomb friction.

### 3.7.5 Coulomb-Contensou Friction

The set-valued force law for Coulomb-Contensou friction can be used to model a contact with Coulomb friction and drilling friction torque [47]. The drilling friction force is the result of an area with distributed Coulomb friction. The force law is formulated by using two translational scalar friction forces  $\lambda_{T1}$ ,  $\lambda_{T2}$  and a scalar friction moment  $\lambda_\tau$  as shown in Figure 3.10. Each of the scalar forces  $\lambda_{T1}$ ,  $\lambda_{T2}$  and  $\lambda_\tau$  has its corresponding relative

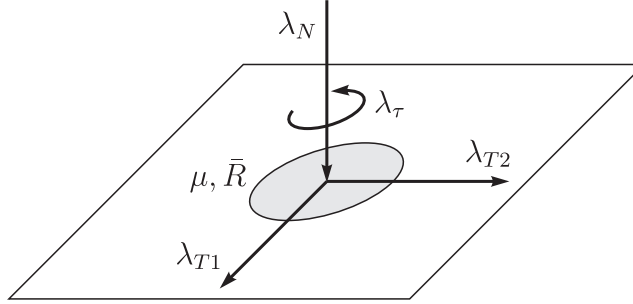


Figure 3.10: Coulomb-Contensou friction forces.

velocity  $\gamma_{T1}$ ,  $\gamma_{T2}$  and  $\gamma_\tau$ . The translational forces and velocities are grouped into vectors to simplify the notation,

$$\boldsymbol{\lambda}_T := \begin{pmatrix} \lambda_{T1} \\ \lambda_{T2} \end{pmatrix}, \quad \boldsymbol{\gamma}_T := \begin{pmatrix} \gamma_{T1} \\ \gamma_{T2} \end{pmatrix}. \quad (3.115)$$

The scalar force  $\lambda_N$  is provided by the associated normal force law. As parameters, the Coulomb friction coefficient  $\mu$  and an average friction radius  $\bar{R}$  are used. For pure translational Coulomb friction, this results in a maximal friction force of  $\mu\lambda_N$ , while pure rotational friction gives a maximum friction moment of  $\bar{R}\mu\lambda_N$ . In general, the three friction forces and relative velocities are not independent, especially the drilling friction interacts with the translational friction. In [47] a force law of the form

$$\begin{pmatrix} \boldsymbol{\gamma}_T \\ \gamma_\tau \end{pmatrix} \in \mathcal{N}_{\mathcal{B}_F} \left( - \begin{pmatrix} \boldsymbol{\lambda}_T \\ \lambda_\tau \end{pmatrix} \right) \quad (3.116)$$

has been derived for a contact with a circular contact area on which a Coulomb friction interaction is formulated in each point and the normal force  $\lambda_N$  is parabolically distributed (Hertz contact model). A more general approach to the reduction of distributed set-valued force laws and their application to Coulomb-Contensou friction can be found in [28]. In the case of parabolic normal force distribution, the average friction radius  $\bar{R}$  is

$$\bar{R} = \frac{3\pi}{16} R \approx 0.589R, \quad (3.117)$$

where  $R$  is the radius of the assumed circular contact area. The functions

$$b_T(u) := \begin{cases} \frac{3\pi\mu\lambda_N}{32}(-u^3 + 4u), & u \leq 1 \\ \frac{3\mu\lambda_N}{16} \left( (-u^3 + 4u) \arcsin\left(\frac{1}{u}\right) + \frac{1}{u}(u^2 + 2)\sqrt{u^2 - 1} \right), & u > 1 \end{cases} \quad (3.118)$$

and

$$b_\tau(u) := \begin{cases} \frac{\bar{R}\mu\lambda_N}{8}(3u^4 - 8u^2 + 8), & u \leq 1 \\ \frac{\bar{R}\mu\lambda_N}{4\pi} \left( (3u^4 - 8u^2 + 8) \arcsin\left(\frac{1}{u}\right) + (-3u^2 + 6)\sqrt{u^2 - 1} \right), & u > 1 \end{cases} \quad (3.119)$$

give a parametric description of the boundary of the set  $\mathcal{B}_F$  used in the force law (3.116). The set  $\mathcal{B}_F$  is then described by the implicit formulation

$$\mathcal{B}_F := \left\{ \begin{pmatrix} \boldsymbol{\lambda}_T \\ \lambda_\tau \end{pmatrix} \mid \|\boldsymbol{\lambda}_T\| \leq b_T(u), |\lambda_\tau| \leq b_\tau(u), u \in [0, \infty) \right\}. \quad (3.120)$$

The set is rotationally symmetric around the  $\lambda_\tau$ -axis and has the cross-section and shape shown in Figure 3.11. The normal cone inclusion (3.116) together with the set  $\mathcal{B}_F$  gives the exact condensed force law for a contact with perfectly circular contact area, parabolic Hertz normal pressure distribution and a distributed Coulomb friction law. Unfortunately, calculating the proximal point to this set in the Euclidean norm is complicated and time-consuming. The simplest approximation of the set  $\mathcal{B}_F$  would be a cylindrical set  $\mathcal{Z}_F$  with the  $\lambda_\tau$ -axis as axis of symmetry, a radius equal to  $\mu\lambda_N$  and a height of  $2\bar{R}\mu\lambda_N$

$$\mathcal{Z}_F := \left\{ \begin{pmatrix} \boldsymbol{\lambda}_T \\ \lambda_\tau \end{pmatrix} \mid \|\boldsymbol{\lambda}_T\| \leq \mu\lambda_N, |\lambda_\tau| \leq \bar{R}\mu\lambda_N \right\}. \quad (3.121)$$

For the set  $\mathcal{Z}_F$ , the proximal point function  $\text{prox}_{\mathcal{Z}_F}$  is very simple and the projection can even be decoupled in  $\lambda_\tau$  and  $\boldsymbol{\lambda}_T$  direction. But exactly this decoupling is not desired, because the interaction between drilling friction and the translational friction is an important property of the model. A better approximation is an ellipsoidal set  $\mathcal{E}_F$  with one semi-axis of  $\bar{R}\mu\lambda_N$  and the other two of  $\mu\lambda_N$  (cf. Figure 3.11),

$$\mathcal{E}_F := \left\{ \begin{pmatrix} \boldsymbol{\lambda}_T \\ \lambda_\tau \end{pmatrix} \mid \left( \frac{\|\boldsymbol{\lambda}_T\|}{\mu\lambda_N} \right)^2 + \left( \frac{\lambda_\tau}{\bar{R}\mu\lambda_N} \right)^2 \leq 1 \right\}. \quad (3.122)$$

Further approximations for Coulomb-Contensou friction and a comparison with experimental results can be found in [37]. The elliptical set (3.122) can be transformed to a sphere  $\mathcal{S}_F$  with radius  $\mu\lambda_N$  by introducing the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \bar{R} \end{pmatrix}, \quad (3.123)$$

and applying the inverse mapping

$$\mathbf{A}^{-1}\mathcal{E}_F = \mathcal{S}_F. \quad (3.124)$$

The resulting set

$$\mathcal{S}_F := \left\{ \begin{pmatrix} \boldsymbol{\lambda}_T \\ \lambda_\tau \end{pmatrix} \mid \|\boldsymbol{\lambda}_T\|^2 + \lambda_\tau^2 \leq \mu^2\lambda_N^2 \right\} \quad (3.125)$$

allows then for an explicit closed form description of the proximal point function which can be evaluated very efficiently (see Section 2.3).

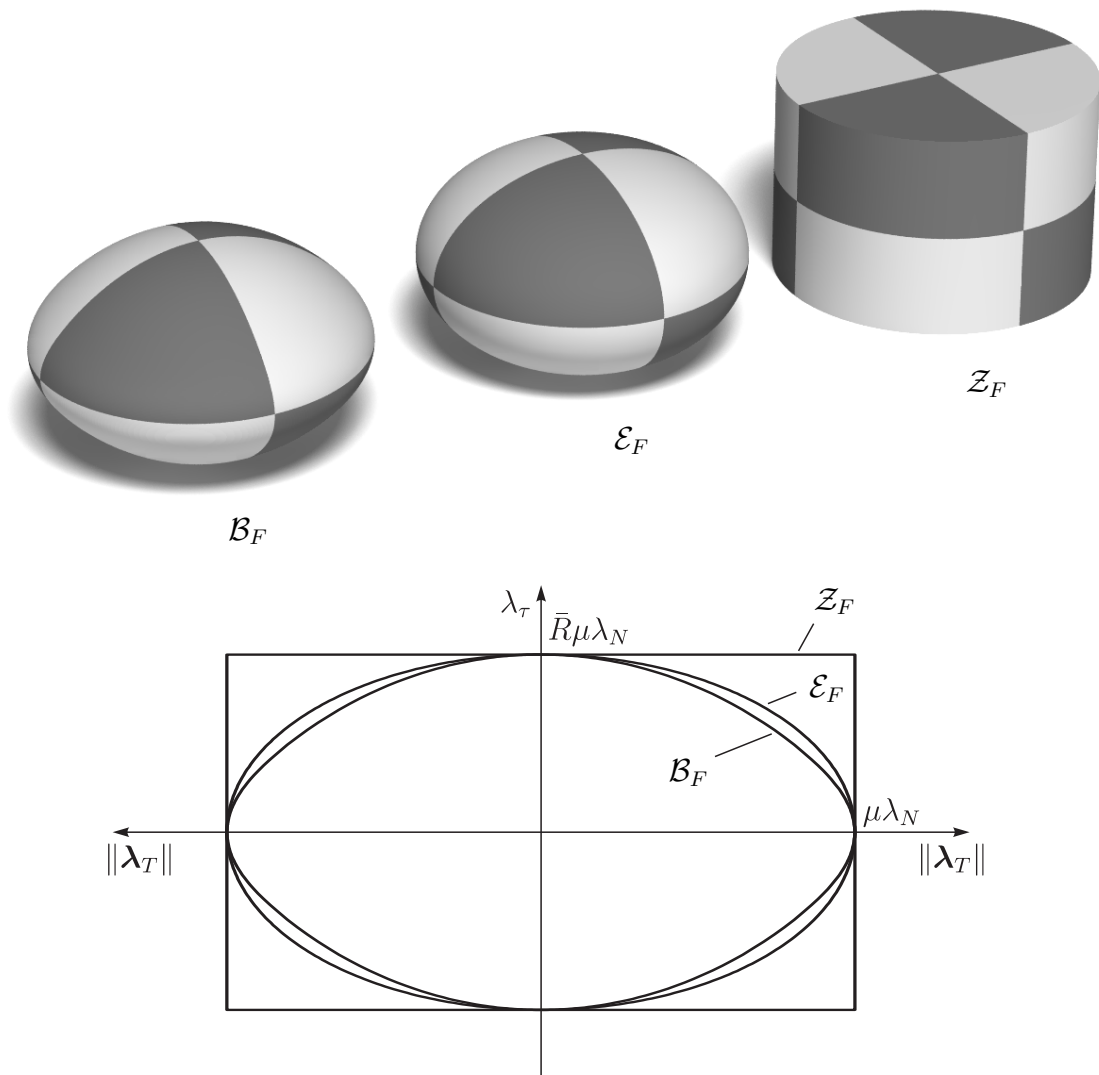


Figure 3.11: Approximations of Coulomb-Contensou friction.

## 3.8 Impacts

In a non-smooth mechanical model with unilateral contacts and friction, the velocities can get discontinuous. A discontinuity in the velocity can occur for example when there is a collision in a unilateral contact. Since the velocities remain finite, the generalized coordinates will always be continuous. To describe the discontinuities in the velocities, the non-smooth set-valued force laws have to be equipped with an impact law and the equations of motion have to be completed with impact equations. A detailed description of this process can be found in [23, 27, 29]. In the following only a short summary is presented.

The equations of motion of a non-smooth mechanical model can be obtained by extending (3.31) with forces  $\lambda_i$  of set-valued force laws from Section 3.7. This yields an equation of motion of the form

$$\mathbf{M}(\mathbf{q}, t)\dot{\mathbf{u}} - \mathbf{h}(\mathbf{q}, \mathbf{u}, t) - \sum_{i \in \mathcal{I}(\mathbf{q}, t)} \mathbf{W}_i(\mathbf{q}, t)\lambda_i = 0, \quad \dot{\mathbf{q}} = \mathbf{F}(\mathbf{q}, t)\mathbf{u} + \boldsymbol{\beta}(\mathbf{q}, t), \quad (3.126)$$

together with set valued force laws

$$\gamma_i \in \mathcal{N}_{C_i}(-\lambda_i), \quad \gamma_i = \mathbf{W}_i^T(\mathbf{q}, t)\mathbf{u} + \chi_i(\mathbf{q}, t), \quad i \in \mathcal{I}(\mathbf{q}, t) \quad (3.127)$$

formulated on velocity level. For unilateral contacts formulated on displacement level and transferred to velocity level as in (3.103), the set  $\mathcal{I}$  contains only the indices  $i$  of the contacts that are closed on displacement level. The impact equations can be obtained by integrating the equations of motion over a singleton in time. This yields the impact equations

$$\mathbf{M}(\mathbf{q}, t)(\mathbf{u}^+ - \mathbf{u}^-) - \sum_{i \in \mathcal{I}(\mathbf{q}, t)} \mathbf{W}_i(\mathbf{q}, t)\Lambda_i = 0 \quad (3.128)$$

where  $\mathbf{u}^-$  and  $\mathbf{u}^+$  are the pre- and post-impact velocities respectively. The pre- and post-impact velocities are the right and left limits

$$\mathbf{u}^-(t) := \lim_{\tau \uparrow 0} \mathbf{u}(t + \tau), \quad \mathbf{u}^+(t) := \lim_{\tau \downarrow 0} \mathbf{u}(t + \tau) \quad (3.129)$$

at the impact time  $t$ . Note that the vector of classical forces  $\mathbf{h}$  disappears in the impact equations, as it is assumed to be free from impulsive forces. The impact-impulses  $\Lambda_i$  have to be understood as time integral of the forces  $\lambda_i$  over the impact time. Since the generalized coordinates are continuous, no distinction between pre- or post-impact coordinates is made. While the impact equations can be derived directly from the equations of motion, the formulation of an impact law involves making additional modeling assumptions. An impact law describes the relation between the impact-impulse, the local pre-impact velocity and the local post-impact velocity. The local pre- and post-impact velocity can be obtained from the local velocities

$$\gamma_i = \mathbf{W}_i^T(\mathbf{q}, t)\mathbf{u} + \chi_i(\mathbf{q}, t), \quad (3.130)$$

by taking the right and the left limit

$$\boldsymbol{\gamma}_i^\pm = \mathbf{W}_i^\top(\mathbf{q}, t)\mathbf{u}^\pm + \boldsymbol{\chi}_i(\mathbf{q}, t). \quad (3.131)$$

If the force law is given by

$$\boldsymbol{\gamma}_i \in \mathcal{N}_{\mathcal{C}_i}(-\boldsymbol{\lambda}_i), \quad (3.132)$$

then one can formulate the associated impact law as

$$\boldsymbol{\gamma}_i^+ + \varepsilon_i \boldsymbol{\gamma}_i^- \in \mathcal{N}_{\mathcal{D}_i}(-\boldsymbol{\Lambda}_i), \quad (3.133)$$

when Newton's impact law in inclusion form is assumed. In the impact law the set  $\mathcal{D}_i$  is the reservoir of all impact-impulses compatible with the force law. The scalar parameter  $\varepsilon_i \in [0, 1]$  is the coefficient of restitution. For  $\varepsilon_i = 0$  one gets a maximal dissipative impact behavior, while  $\varepsilon_i = 1$  yields an elastic impact. Instead of Newton's impact law one can use also Poisson's impact law in inclusion form [23, 77]. For a detailed discussion of impact laws the reader is referred to [5, 23, 27, 70, 77].

### 3.9 Total Energy

In this section, the total energy of a mechanical system is discussed for a few special cases. In the first subsection the development of the total energy during impact-free motion is analyzed, and in the second subsection the changes of the kinetic energy during an impact is discussed. In this section, only the time-autonomous case is considered.

#### Impact-Free Motion

We assume that the kinetic energy of the mechanical system is a quadratic form in  $\dot{\mathbf{q}}$  with a symmetric and positive definite mass matrix  $\mathbf{M}(\mathbf{q})$  depending only on the generalized coordinates,

$$T(\dot{\mathbf{q}}, \mathbf{q}) = \frac{1}{2} \dot{\mathbf{q}}^\top \mathbf{M}(\mathbf{q}) \dot{\mathbf{q}}. \quad (3.134)$$

Beside the inertial forces described by  $T(\dot{\mathbf{q}}, \mathbf{q})$ , we allow potential forces created by the potential  $V(\mathbf{q})$ . In this case the total energy is equal to the sum of the kinetic and the potential energy,

$$E(\dot{\mathbf{q}}, \mathbf{q}) := T(\dot{\mathbf{q}}, \mathbf{q}) + V(\mathbf{q}). \quad (3.135)$$

As described in Section 3.2, the equations of motion of the system are given by

$$\frac{d}{dt} \left( \frac{\partial T}{\partial \dot{\mathbf{q}}} \right)^\top - \left( \frac{\partial T}{\partial \mathbf{q}} \right)^\top + \left( \frac{\partial V}{\partial \mathbf{q}} \right)^\top - \mathbf{f}_q = 0 \quad (3.136)$$

where the generalized force  $\mathbf{f}_q$  can be used to add further force laws. Multiplying (3.136) from the left with  $\dot{\mathbf{q}}^\top$  yields

$$\frac{d}{dt} \left( \frac{\partial T}{\partial \dot{\mathbf{q}}} \right) \dot{\mathbf{q}} - \frac{\partial T}{\partial \mathbf{q}} \dot{\mathbf{q}} + \frac{\partial V}{\partial \mathbf{q}} \dot{\mathbf{q}} - \mathbf{f}_q^\top \dot{\mathbf{q}} = 0. \quad (3.137)$$

The expression with the partial derivative of the potential  $V$  is obviously just the time derivative of the potential

$$\frac{\partial V}{\partial \mathbf{q}} \dot{\mathbf{q}} = \frac{d}{dt} V(\mathbf{q}). \quad (3.138)$$

All terms in which the kinetic energy is used can be simplified as follows:

$$\begin{aligned} \frac{d}{dt} \left( \frac{\partial T}{\partial \dot{\mathbf{q}}} \right) \dot{\mathbf{q}} - \frac{\partial T}{\partial \mathbf{q}} \dot{\mathbf{q}} &= \frac{d}{dt} \left( \frac{\partial T}{\partial \dot{\mathbf{q}}} \dot{\mathbf{q}} \right) - \frac{\partial T}{\partial \dot{\mathbf{q}}} \ddot{\mathbf{q}} - \frac{\partial T}{\partial \mathbf{q}} \dot{\mathbf{q}} \\ &= \frac{d}{dt} \left( \frac{\partial T}{\partial \dot{\mathbf{q}}} \dot{\mathbf{q}} \right) - \frac{d}{dt} T(\dot{\mathbf{q}}, \mathbf{q}) \\ &= \frac{d}{dt} \left( \frac{\partial T}{\partial \dot{\mathbf{q}}} \dot{\mathbf{q}} - T(\dot{\mathbf{q}}, \mathbf{q}) \right) \\ &= \frac{d}{dt} T(\dot{\mathbf{q}}, \mathbf{q}) \end{aligned} \quad (3.139)$$

In the last line of (3.139) the simplification

$$\frac{\partial T}{\partial \dot{\mathbf{q}}} = \dot{\mathbf{q}}^\top \mathbf{M}(\mathbf{q}) \quad \Rightarrow \quad \frac{\partial T}{\partial \dot{\mathbf{q}}} \dot{\mathbf{q}} = 2T(\dot{\mathbf{q}}, \mathbf{q}) \quad (3.140)$$

has been used. Inserting (3.138) and (3.139) into (3.137) yields

$$\frac{d}{dt} (T(\dot{\mathbf{q}}, \mathbf{q}) + V(\mathbf{q})) - \mathbf{f}_q^\top \dot{\mathbf{q}} = 0. \quad (3.141)$$

This means that the time derivative of the total energy is given by

$$\dot{E}(\dot{\mathbf{q}}, \mathbf{q}) = \mathbf{f}_q^\top \dot{\mathbf{q}}. \quad (3.142)$$

If there are no additional forces then the time derivative of the total energy is zero. In this case the total energy is conserved. Considering additional forces from geometric, scleronomic and perfect bilateral constraints of the form

$$g(\mathbf{q}) = 0, \quad \mathbf{f}_q = \left( \frac{\partial g}{\partial \mathbf{q}} \right)^\top \lambda \quad (3.143)$$

does not provide any additional contribution to  $\mathbf{f}_q^\top \dot{\mathbf{q}}$ . The total energy is still preserved as follows directly from the force law

$$g(\mathbf{q}) = 0 \quad \Rightarrow \quad \frac{\partial g}{\partial \mathbf{q}} \dot{\mathbf{q}} = 0 \quad \Rightarrow \quad \mathbf{f}_q^\top \dot{\mathbf{q}} = 0. \quad (3.144)$$

A perfect unilateral constraint that depends only on the generalized coordinates  $\mathbf{q}$  is given by the force law

$$g(\mathbf{q}) \in \mathcal{N}_{\mathbb{R}_0^-}(-\lambda), \quad \mathbf{f}_q = \left( \frac{\partial g}{\partial \mathbf{q}} \right)^\top \lambda. \quad (3.145)$$

If the contact is open, then the force of the unilateral contact is zero

$$g(\mathbf{q}) > 0 \quad \Rightarrow \quad \lambda = 0 \quad \Rightarrow \quad \mathbf{f}_q^\top \dot{\mathbf{q}} = 0 \quad (3.146)$$

and the total kinetic energy remains unchanged. For situations where the contact is closed we get the same situation as for the perfect bilateral constraint, i.e.

$$g(\mathbf{q}) = 0 \quad \Rightarrow \quad \frac{\partial g}{\partial \mathbf{q}} \dot{\mathbf{q}} = 0 \quad \Rightarrow \quad \mathbf{f}_q^\top \dot{\mathbf{q}} = 0. \quad (3.147)$$

This means the total energy is preserved in systems with perfect unilateral constraints of the above type as well. A simple kinematic force law of normal cone type has the form

$$\mathbf{W}^\top(\mathbf{q}, t) \dot{\mathbf{q}} \in \mathcal{N}_C(-\boldsymbol{\lambda}), \quad \mathbf{f}_q = \mathbf{W}(\mathbf{q}, t) \boldsymbol{\lambda}. \quad (3.148)$$

For convex sets  $\mathcal{C}$  that contain the zero element, we get a (not strictly) dissipative system. This can be verified using the definition (2.4) of the normal cone

$$0 \in \mathcal{C}, \quad \mathbf{W}^\top(\mathbf{q}, t) \dot{\mathbf{q}} \in \mathcal{N}_C(-\boldsymbol{\lambda}) \quad \Rightarrow \quad \boldsymbol{\lambda}^\top \mathbf{W}^\top(\mathbf{q}, t) \dot{\mathbf{q}} \leq 0 \quad \Rightarrow \quad \mathbf{f}_q^\top \dot{\mathbf{q}} \leq 0. \quad (3.149)$$

A force law of the form (3.148) is for example the Coulomb friction law discussed in Section 3.7.3.

## Impacts

In this subsection the relation between pre- and post-impact kinetic energy of a non-smooth mechanical system is discussed for a few special cases. For more details on the subject the reader is referred to [5, 27, 46, 49, 70, 77]. We assume a system with an impact equation and impact laws of the form

$$\begin{aligned} \mathbf{M}(\mathbf{u}^+ - \mathbf{u}^-) - \sum_{i \in \mathcal{I}} \mathbf{W}_i \boldsymbol{\Lambda}_i &= 0 \\ i \in \mathcal{I}: \mathbf{W}_i^\top(\mathbf{u}^+ + \varepsilon_i \mathbf{u}^-) &\in \mathcal{N}_{C_i}(-\boldsymbol{\Lambda}_i). \end{aligned} \quad (3.150)$$

The dependencies of the mass matrix  $\mathbf{M}$  and the matrices of generalized force directions  $\mathbf{W}_i$  on time  $t$  and generalized coordinates  $\mathbf{q}$  are neglected here, as both are constant for an impact. The kinetic energy of the system (3.150) is given by

$$T(\mathbf{u}) = \frac{1}{2} \mathbf{u}^\top \mathbf{M} \mathbf{u}. \quad (3.151)$$

The difference of the pre- and post-impact kinetic energy can be reformulated using the impact equation (first equation in (3.150)). One obtains the relation

$$\begin{aligned} T(\mathbf{u}^+) - T(\mathbf{u}^-) &= \frac{1}{2} \mathbf{u}^{+\top} \mathbf{M} \mathbf{u}^+ - \frac{1}{2} \mathbf{u}^{-\top} \mathbf{M} \mathbf{u}^- \\ &= \frac{1}{2} (\mathbf{u}^+ - \mathbf{u}^-)^\top \mathbf{M} (\mathbf{u}^+ + \mathbf{u}^-) \\ &= \frac{1}{2} \sum_{i \in \mathcal{I}} \boldsymbol{\Lambda}_i^\top \mathbf{W}_i^\top (\mathbf{u}^+ + \mathbf{u}^-) \end{aligned} \quad (3.152)$$



which shows that the difference in kinetic energy is the product of the average local velocity  $\mathbf{W}_i^\top(\mathbf{u}^+ + \mathbf{u}^-)/2$  and the impact-impulses  $\Lambda_i$ . The impact laws can be rewritten to use only the sum and the difference of the pre- and post-impact generalized velocities

$$\begin{aligned}
& \mathbf{W}_i^\top(\mathbf{u}^+ + \varepsilon_i \mathbf{u}^-) \in \mathcal{N}_{\mathcal{C}_i}(-\Lambda_i) \\
& \Leftrightarrow \frac{1}{2} \mathbf{W}_i^\top \left( (1 + \varepsilon_i)(\mathbf{u}^+ + \mathbf{u}^-) + (1 - \varepsilon_i)(\mathbf{u}^+ - \mathbf{u}^-) \right) \in \mathcal{N}_{\mathcal{C}_i}(-\Lambda_i) \\
& \Leftrightarrow \underbrace{\mathbf{W}_i^\top \left( (\mathbf{u}^+ + \mathbf{u}^-) + \frac{1 - \varepsilon_i}{1 + \varepsilon_i} (\mathbf{u}^+ - \mathbf{u}^-) \right)}_{=: \boldsymbol{\xi}_i} \in \mathcal{N}_{\mathcal{C}_i}(-\Lambda_i) \\
& \Leftrightarrow \boldsymbol{\xi}_i \in \mathcal{N}_{\mathcal{C}_i}(-\Lambda_i).
\end{aligned} \tag{3.153}$$

In the third line of (3.153) the invariance of a cone to scaling with a positive number has been used. The product of the vector  $\boldsymbol{\xi}_i$  introduced in (3.153) with the impact-impulses  $\Lambda_i$  is always negative if the convex set  $\mathcal{C}_i$  contains the zero element. This follows directly

$$0 \in \mathcal{C}_i, \quad \boldsymbol{\xi}_i \in \mathcal{N}_{\mathcal{C}_i}(-\Lambda_i) \quad \Rightarrow \quad \boldsymbol{\xi}_i^\top \Lambda_i \leq 0. \tag{3.154}$$

from the definition of the normal cone (2.4). To exploit the inequality (3.154) in the further analysis of (3.152), we will rewrite the sum  $\mathbf{W}_i^\top(\mathbf{u}^+ + \mathbf{u}^-)$  of local velocities using  $\boldsymbol{\xi}_i$ . One gets

$$\begin{aligned}
\mathbf{W}_i^\top(\mathbf{u}^+ + \mathbf{u}^-) &= \boldsymbol{\xi}_i - \delta_i \mathbf{W}_i^\top(\mathbf{u}^+ - \mathbf{u}^-) \\
&= \boldsymbol{\xi}_i - \delta_i \mathbf{W}_i^\top \mathbf{M}^{-1} \sum_{j \in \mathcal{I}} \mathbf{W}_j \Lambda_j
\end{aligned} \tag{3.155}$$

where the definition of the dissipation index

$$\delta_i := \frac{1 - \varepsilon_i}{1 + \varepsilon_i} \tag{3.156}$$

has been used to simplify the notation. Inserting (3.155) into (3.152) yields

$$\begin{aligned}
T(\mathbf{u}^+) - T(\mathbf{u}^-) &= \frac{1}{2} \sum_{i \in \mathcal{I}} \left( \Lambda_i^\top \boldsymbol{\xi}_i - \delta_i \Lambda_i^\top \mathbf{W}_i^\top \mathbf{M}^{-1} \sum_{j \in \mathcal{I}} \mathbf{W}_j \Lambda_j \right) \\
&= \frac{1}{2} \sum_{i \in \mathcal{I}} \Lambda_i^\top \boldsymbol{\xi}_i - \frac{1}{2} \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{I}} \delta_i \Lambda_i^\top \mathbf{W}_i^\top \mathbf{M}^{-1} \mathbf{W}_j \Lambda_j.
\end{aligned} \tag{3.157}$$

As next step, the vectors  $\boldsymbol{\xi}_i$  and  $\Lambda_i$  are assembled into global vectors  $\boldsymbol{\xi}$  and  $\Lambda$ , and the products  $\mathbf{W}_i^\top \mathbf{M}^{-1} \mathbf{W}_j$  are combined into one large Delassus matrix  $\mathbf{G}$ , where the  $ij$ -th sub-block is given by

$$\mathbf{G}_{ij} = \mathbf{W}_i^\top \mathbf{M}^{-1} \mathbf{W}_j. \tag{3.158}$$

The resulting Delassus matrix  $\mathbf{G}$  is symmetric and at least positive semidefinite. For the dissipation indices a matrix

$$\Delta_{ij} = \begin{cases} \delta_i \mathbf{I}, & i = j \\ 0, & i \neq j \end{cases} \tag{3.159}$$

is introduced. Using these assembled objects, the kinetic energy difference can be written as

$$T(\mathbf{u}^+) - T(\mathbf{u}^-) = \frac{1}{2} \mathbf{\Lambda}^\top \boldsymbol{\xi} - \frac{1}{2} \mathbf{\Lambda}^\top \mathbf{\Delta} \mathbf{G} \mathbf{\Lambda}. \quad (3.160)$$

The first term  $\mathbf{\Lambda}^\top \boldsymbol{\xi}$  is lower or equal to zero as shown in (3.154). If the matrix  $\mathbf{\Delta} \mathbf{G}$  is positive semidefinite, then the second term in (3.160) is always lower or equal to zero as well. In that case the inequality  $T(\mathbf{u}^+) \leq T(\mathbf{u}^-)$  holds. Whether  $\mathbf{\Delta} \mathbf{G}$  is at least positive semidefinite depends on the coefficients of restitution and the Delassus matrix  $\mathbf{G}$ . A discussion of sufficient conditions for positive definite matrix  $\mathbf{\Delta} \mathbf{G}$  can be found in [46]. A simple case is using equal coefficients of restitution  $\varepsilon \in [0, 1]$  for all impact laws. This simplifies the matrix of dissipation indices to

$$\mathbf{\Delta} = \frac{1 - \varepsilon}{1 + \varepsilon} \mathbf{I}, \quad (3.161)$$

which yields a positive semidefinite product  $\mathbf{\Delta} \mathbf{G}$ , since  $\mathbf{G}$  is at least positive semidefinite. From (3.160) then follows the inequality

$$T(\mathbf{u}^+) \leq T(\mathbf{u}^-) \quad (3.162)$$

guaranteeing that the kinetic energy after an impact is never larger than before it, if all coefficients of restitution are equal.

In the remaining part of this section we will discuss a simple example where an impact could lead to an increase in total energy if the coefficients of restitution are not chosen carefully. We consider the mechanical system shown in Figure 3.12, consisting of a wall, two masses, a sprag clutch, and a unilateral contact.

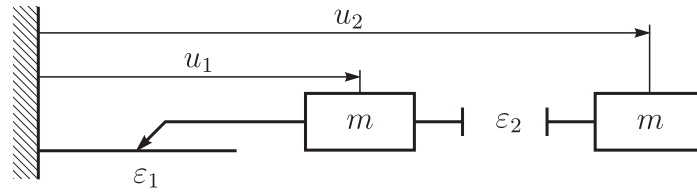


Figure 3.12: Sprag clutch and unilateral contact.

We assume a situation where the unilateral contact is closed on displacement level. The impact equations for the system are then given by

$$\begin{pmatrix} m & 0 \\ 0 & m \end{pmatrix} \begin{pmatrix} u_1^+ - u_1^- \\ u_2^+ - u_2^- \end{pmatrix} - \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Lambda_1 \\ \Lambda_2 \end{pmatrix} = 0. \quad (3.163)$$

For both elements an impact law of Newton type in inclusion form is assumed

$$\begin{aligned} u_1^+ + \varepsilon_1 u_1^- &\in \mathcal{N}_{\mathbb{R}_0^-}(-\Lambda_1) \\ (u_2^+ - u_1^+) + \varepsilon_2 (u_2^- - u_1^-) &\in \mathcal{N}_{\mathbb{R}_0^-}(-\Lambda_2). \end{aligned} \quad (3.164)$$

The coefficient of restitution  $\varepsilon_1$  of the sprag clutch has to be zero, because any value larger than zero could lead to kinematically inconsistent post impact velocities. For the unilateral contact we assume a perfect elastic impact behavior, i.e.

$$\varepsilon_1 = 0, \quad \varepsilon_2 = 1. \quad (3.165)$$

We consider a situation where the left mass is moving right and the right mass is moving left before the impact

$$u_1^- > 0, \quad u_2^- < 0. \quad (3.166)$$

As can be easily verified, the solution to the impact problem is given by

$$u_1^+ = 0, \quad u_2^+ = u_1^- - u_2^-, \quad \Lambda_1 = -2m u_2^-, \quad \Lambda_2 = m(u_1^- - 2u_2^-). \quad (3.167)$$

After the impact, the left mass is resting and the right mass moves to the right. For the change in kinetic energy caused by the impact we get

$$T^+ - T^- = \frac{m}{2}(u_1^- - u_2^-)^2 - \frac{m}{2}(u_1^{-2} + u_2^{-2}) = -m u_1^- u_2^- > 0. \quad (3.168)$$

Obviously the kinetic energy has increased, while the system itself is passive. The problem can be solved either by using a different impact law (e.g. Poisson's impact law), or by making  $\varepsilon_2$  smaller. For  $\varepsilon_2 = 0$  both coefficient of restitution are identical and the kinetic energy does not increase during an impact, as shown above.



## Rigid Body Formulations

---

In this chapter, different formulations for the equations of motion of a rigid body are described. These formulations can then be used with the energy consistent integration schemes derived in Chapter 5. Beside this, the aim of this chapter is also to provide some mechanical interpretation for these rigid body formulations.

The equations of motion of a rigid body can be formulated as a set of ordinary differential equations (ODE) in the form of the Newton-Euler equations. For a complete description, these are equipped with a parametrization of finite rotations in the form of a differential equation linking the angular velocity to the derivative of the rotation parameters. Well known concepts for the parametrization of rotations are Euler and Kardan angles, unit quaternions (also known as Euler parameters), Rodrigues parameters, rotation vectors and components of the rotation matrix. The Euler and Kardan angles use three parameters for the three rotational degrees of freedom of a rigid body. The mapping of the derivatives of the rotation parameters to the angular velocity of the body can become singular for three parameter formulations. In these cases the derivative of the parameters can not be determined from the angular velocity, which is of course a disadvantage.

The four parameter formulations like the unit quaternions, Rodrigues parameters or rotation vectors, as well as the six or nine parameter formulations based on the components of the rotation matrix don't have the singularity problem. Since more parameters than degrees of freedom are used, the parameters are subject to additional constraints. The additional constraints, which are enforced only on velocity level, can lead to drift problems in an ODE formulation. To prevent drift in numerical simulations, the rotation parameters have therefore to be projected back to pure rotations after each integration step (e.g. as in [45] for quaternions) or, alternatively, a discretization scheme which preserves the pure rotation constraints has to be employed. One way to achieve the latter is to formulate the equation of motion of a rigid body as a set of differential-algebraic equations (DAE). In the DAE formulation the algebraic equations restrict the rotation parameters to the subset that describe pure rotation.

In this chapter, DAE formulations for the dynamics of a rigid body are derived using the following approach: First, the infinite dimensional dynamics of the underlying continuum is reduced to a body with the full degrees of freedom created by the rotation parameters.

When the components of the rotation matrix are used as rotation parameters, then the finite-dimensional body is the affine body discussed in Section 4.2. The scalable body described in Section 4.3 is obtained when a quaternion is used as rotation parameter. To the model of a body with the full degrees of freedom of the rotation parameters, a set of perfect bilateral constraints is added, constraining the rotation parameters to pure rotation. This then yields a DAE formulation for the dynamics of a rigid body where the Lagrange multipliers have the meaning of constraint forces, and where the mass matrix is nonsingular. As a preparation for the affine body discussed in Section 4.2 and the scalable body discussed in Section 4.3, some definitions to characterize the mass distribution in a body are given in Section 4.1.

## 4.1 Mass Distribution

In this section, the definition for the total mass, the classical inertia tensor and Binet's inertia tensor are given. These integrals capture the most important properties of a mass distribution on a body. The mass distribution  $dm$  shall be given as a function of the location  $\boldsymbol{\varrho}$  of each point of the body  $\mathcal{B}$  (cf. Figure 4.1). The location  $\boldsymbol{\varrho}$  is measured

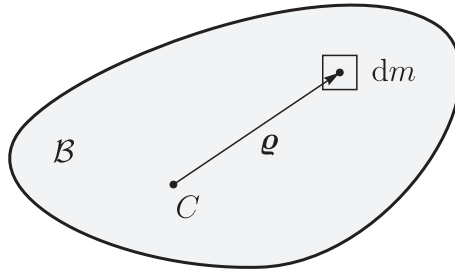


Figure 4.1: Mass distribution.

starting from the point  $C$ , which is chosen to be identical with the center of mass of the mass distribution. This means that the integral

$$\int_{\mathcal{B}} \boldsymbol{\varrho} \, dm = 0 \quad (4.1)$$

is always zero. Note that the point  $C$  remains the center of mass, even if the body is deformed linearly

$$\int_{\mathcal{B}} \mathbf{A} \boldsymbol{\varrho} \, dm = \mathbf{A} \int_{\mathcal{B}} \boldsymbol{\varrho} \, dm = 0. \quad (4.2)$$

The total mass  $m$  of the body is obtained by integrating  $dm$  over the complete body, i.e.

$$m := \int_{\mathcal{B}} dm. \quad (4.3)$$

The definition for the classical inertia tensor with respect to point  $C$  is given by

$$\boldsymbol{\Theta} := \int_{\mathcal{B}} -\tilde{\boldsymbol{\varrho}}^2 \, dm \quad (4.4)$$

where  $\tilde{\boldsymbol{\rho}}$  denotes the skew-symmetric matrix associated with the cross product. For a vector  $\mathbf{a} = (a_1 \ a_2 \ a_3)^\top$ , the skew-symmetric matrix  $\tilde{\mathbf{a}} \in \mathbb{R}^{3 \times 3}$  associated with the cross product is given by

$$\tilde{\mathbf{a}} = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix}. \quad (4.5)$$

It holds that  $\tilde{\mathbf{a}}\mathbf{b} = \mathbf{a} \times \mathbf{b}$  for any  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ . Binet's inertia tensor, also known as Euler tensor, can be defined as

$$\mathbf{E} := \int_{\mathcal{B}} \boldsymbol{\rho} \boldsymbol{\rho}^\top dm. \quad (4.6)$$

The classical inertia tensor and Binet's inertia tensor capture the same information about the mass distribution on a body. Both tensors are formulated with respect to the center of gravity and both are symmetric. With the identity

$$-\tilde{\boldsymbol{\rho}}^2 \equiv \boldsymbol{\rho}^\top \boldsymbol{\rho} \mathbf{I} - \boldsymbol{\rho} \boldsymbol{\rho}^\top, \quad (4.7)$$

which follows from the vector triple product expansion (Lagrange's formula) and the relation

$$\boldsymbol{\rho}^\top \boldsymbol{\rho} = \text{Tr}(\boldsymbol{\rho} \boldsymbol{\rho}^\top) = \frac{1}{2} \text{Tr}(-\tilde{\boldsymbol{\rho}}^2) \quad (4.8)$$

one can derive directly the formulas

$$\boldsymbol{\Theta} = \text{Tr}(\mathbf{E})\mathbf{I} - \mathbf{E}, \quad \mathbf{E} = \frac{1}{2} \text{Tr}(\boldsymbol{\Theta})\mathbf{I} - \boldsymbol{\Theta} \quad (4.9)$$

to calculate the classical inertia tensor from Binet's inertia tensor and vice versa.

## 4.2 Affine Body

The mechanical model of an affine body has twelve degrees of freedom. The kinematics of an affine body can be described with an affine mapping in three dimensions as discussed in Section 4.2.1. The mass matrix of an affine body is constant when the generalized coordinates are formed by a translation vector and the components of a  $3 \times 3$  transformation matrix. The resulting equations of motion of an affine body have a very simple form as derived in Section 4.2.2. In Section 4.2.3, additional perfect bilateral constraints are added to the model in order to make the body rigid. The resulting simple DAE description for the dynamics of a rigid body is useful for the construction of energy consistent integration schemes as discussed in Chapter 5. In the literature this rigid body DAE formulation can be found for example in [13].

### 4.2.1 Kinematics

The affine body model has three translational, three rotational and six linear deformational degrees of freedom. A vector  $\boldsymbol{\rho} \in \mathbb{R}^3$  ( $\boldsymbol{\rho} = \text{const.}$ ) is used to address each point  $P'$  of

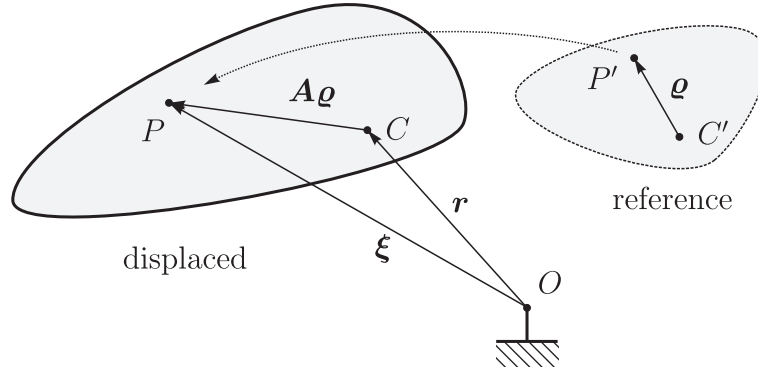


Figure 4.2: Affine body kinematics.

the affine body in a reference configuration (cf. Figure 4.2), where  $\boldsymbol{\rho}$  starts at a reference point  $C'$ . The actual position  $P$  of a point in the displaced configuration is described with the vector  $\boldsymbol{\xi} \in \mathbb{R}^3$  relative to the inertial point  $O$ . The actual position  $C$  of the reference point is identical to the displacement  $\boldsymbol{r} \in \mathbb{R}^3$ . The vector  $\boldsymbol{\xi}$  can be obtained by applying a linear and nonsingular transformation  $\boldsymbol{A} \in \mathbb{R}^{3 \times 3}$  to the vector  $\boldsymbol{\rho}$  and adding it to the displacement  $\boldsymbol{r}$ . This yields the kinematic relation

$$\boldsymbol{\xi} = \boldsymbol{A}\boldsymbol{\rho} + \boldsymbol{r} \quad (4.10)$$

for an affine body. The linear transformation  $\boldsymbol{A}$  is parametrized with three direction vectors  $\boldsymbol{d}_1, \boldsymbol{d}_2$  and  $\boldsymbol{d}_3 \in \mathbb{R}^3$ ,

$$\boldsymbol{A} = \boldsymbol{d}_1 \boldsymbol{e}_1^\top + \boldsymbol{d}_2 \boldsymbol{e}_2^\top + \boldsymbol{d}_3 \boldsymbol{e}_3^\top, \quad (4.11)$$

and the help of the constant and orthonormal basis  $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ . The three direction vectors together with the displacement  $\boldsymbol{r}$  are grouped into a generalized coordinates vector

$$\boldsymbol{q} := \begin{pmatrix} \boldsymbol{r} \\ \boldsymbol{d}_1 \\ \boldsymbol{d}_2 \\ \boldsymbol{d}_3 \end{pmatrix}. \quad (4.12)$$

Writing (4.10) in terms of the generalized coordinates  $\boldsymbol{q}$  yields

$$\begin{aligned} \boldsymbol{\xi} &= \boldsymbol{d}_1 \boldsymbol{e}_1^\top \boldsymbol{\rho} + \boldsymbol{d}_2 \boldsymbol{e}_2^\top \boldsymbol{\rho} + \boldsymbol{d}_3 \boldsymbol{e}_3^\top \boldsymbol{\rho} + \boldsymbol{r} \\ &= (\boldsymbol{I} \quad \boldsymbol{\rho}^\top \boldsymbol{e}_1 \boldsymbol{I} \quad \boldsymbol{\rho}^\top \boldsymbol{e}_2 \boldsymbol{I} \quad \boldsymbol{\rho}^\top \boldsymbol{e}_3 \boldsymbol{I}) \boldsymbol{q} \end{aligned} \quad (4.13)$$

The absolute velocity  $\dot{\boldsymbol{\xi}}$  of a point of the body is obtained by differentiating (4.13) with respect to time. This yields

$$\dot{\boldsymbol{\xi}} = (\boldsymbol{I} \quad \boldsymbol{\rho}^\top \boldsymbol{e}_1 \boldsymbol{I} \quad \boldsymbol{\rho}^\top \boldsymbol{e}_2 \boldsymbol{I} \quad \boldsymbol{\rho}^\top \boldsymbol{e}_3 \boldsymbol{I}) \dot{\boldsymbol{q}} \quad (4.14)$$

for the absolute velocity  $\dot{\boldsymbol{\xi}}$  in terms of the derivative of the generalized coordinates  $\dot{\boldsymbol{q}}$ .



### 4.2.2 Equations of Motion

The equations of motion of an affine body can be obtained using the variational formulation (3.20) of Lagrange's equations of the second kind. With the help of the kinematics from Section 4.2.1 one can evaluate the kinetic energy. This yields

$$\begin{aligned}
 T &= \frac{1}{2} \int_{\mathcal{B}} \dot{\boldsymbol{\xi}}^\top \dot{\boldsymbol{\xi}} \, dm \\
 &= \frac{1}{2} \dot{\mathbf{q}}^\top \int_{\mathcal{B}} \underbrace{\begin{pmatrix} \mathbf{I} & \boldsymbol{\rho}^\top \mathbf{e}_1 \mathbf{I} & \boldsymbol{\rho}^\top \mathbf{e}_2 \mathbf{I} & \boldsymbol{\rho}^\top \mathbf{e}_3 \mathbf{I} \\ \boldsymbol{\rho}^\top \mathbf{e}_1 \mathbf{I} & \mathbf{e}_1^\top \boldsymbol{\rho} \boldsymbol{\rho}^\top \mathbf{e}_1 \mathbf{I} & \mathbf{e}_1^\top \boldsymbol{\rho} \boldsymbol{\rho}^\top \mathbf{e}_2 \mathbf{I} & \mathbf{e}_1^\top \boldsymbol{\rho} \boldsymbol{\rho}^\top \mathbf{e}_3 \mathbf{I} \\ \boldsymbol{\rho}^\top \mathbf{e}_2 \mathbf{I} & \mathbf{e}_2^\top \boldsymbol{\rho} \boldsymbol{\rho}^\top \mathbf{e}_1 \mathbf{I} & \mathbf{e}_2^\top \boldsymbol{\rho} \boldsymbol{\rho}^\top \mathbf{e}_2 \mathbf{I} & \mathbf{e}_2^\top \boldsymbol{\rho} \boldsymbol{\rho}^\top \mathbf{e}_3 \mathbf{I} \\ \boldsymbol{\rho}^\top \mathbf{e}_3 \mathbf{I} & \mathbf{e}_3^\top \boldsymbol{\rho} \boldsymbol{\rho}^\top \mathbf{e}_1 \mathbf{I} & \mathbf{e}_3^\top \boldsymbol{\rho} \boldsymbol{\rho}^\top \mathbf{e}_2 \mathbf{I} & \mathbf{e}_3^\top \boldsymbol{\rho} \boldsymbol{\rho}^\top \mathbf{e}_3 \mathbf{I} \end{pmatrix}}_{=: \mathbf{M}} \, dm \, \dot{\mathbf{q}} \\
 &= \frac{1}{2} \dot{\mathbf{q}}^\top \mathbf{M} \dot{\mathbf{q}},
 \end{aligned} \tag{4.15}$$

where  $\mathbf{M}$  is the mass matrix of the affine body. Choosing the reference point  $C'$  identical with the center of mass of the mass distribution in the reference configuration yields an integral

$$\int_{\mathcal{B}} \boldsymbol{\rho} \, dm = 0 \tag{4.16}$$

which is always zero. The two abbreviations for the mass and Binet's inertia tensor

$$m = \int_{\mathcal{B}} dm, \quad \mathbf{E} = \int_{\mathcal{B}} \boldsymbol{\rho} \boldsymbol{\rho}^\top \, dm \tag{4.17}$$

introduced in Section 4.1 represent the mass distribution in the body. For the components of Binet's inertia tensor we introduce the following abbreviation

$$E_{ij} := \mathbf{e}_i^\top \mathbf{E} \mathbf{e}_j. \tag{4.18}$$

Using these abbreviations one gets the mass matrix

$$\mathbf{M} = \begin{pmatrix} m\mathbf{I} & 0 & 0 & 0 \\ 0 & E_{11}\mathbf{I} & E_{12}\mathbf{I} & E_{13}\mathbf{I} \\ 0 & E_{21}\mathbf{I} & E_{22}\mathbf{I} & E_{23}\mathbf{I} \\ 0 & E_{31}\mathbf{I} & E_{32}\mathbf{I} & E_{33}\mathbf{I} \end{pmatrix}, \tag{4.19}$$

for an affine body. With the constant mass matrix  $\mathbf{M}$ , the evaluation of (3.20) yields

$$\mathbf{M} \ddot{\mathbf{q}} - \mathbf{f}_q = 0 \tag{4.20}$$

as equation of motion of the affine body.

### 4.2.3 Rigidity Constraints

This section is used to add perfect bilateral constraints to the affine body model, in order to get a rigid body model. First the remaining generalized forces  $\mathbf{f}_q$  are split up

$$\mathbf{f}_q = \mathbf{f}_g + \mathbf{f}_e \quad (4.21)$$

into the generalized constraint forces  $\mathbf{f}_g$  formulated in this section, and the additional generalized external forces  $\mathbf{f}_e$ . To obtain a rigid body, the six constraints

$$\mathbf{g}(\mathbf{q}) = \begin{pmatrix} \mathbf{d}_1^\top \mathbf{d}_1 - 1 \\ \mathbf{d}_2^\top \mathbf{d}_2 - 1 \\ \mathbf{d}_3^\top \mathbf{d}_3 - 1 \\ \mathbf{d}_2^\top \mathbf{d}_3 \\ \mathbf{d}_1^\top \mathbf{d}_3 \\ \mathbf{d}_1^\top \mathbf{d}_2 \end{pmatrix} = 0 \quad (4.22)$$

are imposed. The first three enforce a unit length of the direction vectors, while the last three guarantee that the direction vectors are pairwise orthogonal. The force law is completed by requiring that the virtual work of the constraint forces vanishes for any virtual displacements induced by  $\delta \mathbf{q}$  that are compatible with the constraints, i.e.

$$\delta \mathbf{q}^\top \mathbf{f}_g = 0 \quad \forall \delta \mathbf{q} \mid \delta \mathbf{g} = 0. \quad (4.23)$$

This is a perfect bilateral constraint of d'Alembert-Lagrange type. The variations  $\delta \mathbf{g}$  of the constraints are related to the variations  $\delta \mathbf{q}$  of the generalized coordinates by

$$\delta \mathbf{g} = \frac{\partial \mathbf{g}}{\partial \mathbf{q}} \delta \mathbf{q}. \quad (4.24)$$

Inserting this into the variational formulation of the force law yields

$$\delta \mathbf{q}^\top \mathbf{f}_g = 0 \quad \forall \delta \mathbf{q} \mid \delta \mathbf{q}^\top \left( \frac{\partial \mathbf{g}}{\partial \mathbf{q}} \right)^\top = 0. \quad (4.25)$$

From (4.25) one gets the generalized constraint forces  $\mathbf{f}_g$  by evaluating the variation. The generalized constraint forces lie in the linear subspace spanned by the columns of  $(\partial \mathbf{g} / \partial \mathbf{q})^\top$ , i.e.

$$\mathbf{f}_g = \left( \frac{\partial \mathbf{g}}{\partial \mathbf{q}} \right)^\top \boldsymbol{\lambda}, \quad \boldsymbol{\lambda} \in \mathbb{R}^6. \quad (4.26)$$

The partial derivative  $\partial \mathbf{g} / \partial \mathbf{q}$  yields

$$\frac{\partial \mathbf{g}}{\partial \mathbf{q}} = \begin{pmatrix} 0 & 2\mathbf{d}_1^\top & 0 & 0 \\ 0 & 0 & 2\mathbf{d}_2^\top & 0 \\ 0 & 0 & 0 & 2\mathbf{d}_3^\top \\ 0 & 0 & \mathbf{d}_3^\top & \mathbf{d}_2^\top \\ 0 & \mathbf{d}_3^\top & 0 & \mathbf{d}_1^\top \\ 0 & \mathbf{d}_2^\top & \mathbf{d}_1^\top & 0 \end{pmatrix}. \quad (4.27)$$

Inserting the generalized constraint force  $\mathbf{f}_g$  into the equations of motion of the affine body and completing the set of equations with the constraint equations yields the DAE description

$$\begin{aligned} M\ddot{\mathbf{q}} - \left(\frac{\partial \mathbf{g}}{\partial \dot{\mathbf{q}}}\right)^\top \boldsymbol{\lambda} - \mathbf{f}_e &= 0 \\ \mathbf{g}(\mathbf{q}) &= 0 \end{aligned} \quad (4.28)$$

of the dynamics of a rigid body. Note that in this formulation all gyroscopic forces of a rigid body are generated by the forces  $\boldsymbol{\lambda}$  of the rigidity constraints.

### 4.3 Scalable Body

The mechanical model of a scalable body has three translational, three rotational and one uniform scaling degree of freedom. The displacement of the center of mass and an unconstrained quaternion are used as generalized coordinates. By introducing an additional perfect bilateral constraint one can force the scalable body to become a rigid body, i.e. the scaling degree of freedom is suppressed by this constraint. Without reducing the set of coordinates, this yields naturally a DAE description of the dynamics of a rigid body where the  $7 \times 7$  mass matrix is positive definite and the unit length restriction is enforced by a mechanical constraint. This section is based on the article to be published as [58].

In [65, 66] the Newton-Euler equations of a rigid body in terms of the translational and angular velocity is used as a starting point for the derivation of a quaternion based rigid body DAE formulation. Subsequently, the derivative of the unit quaternion is related to the angular velocity, and the equations of motion are extended to a DAE. Unfortunately, this approach yields a cumbersome formulation for which the Lagrange multiplier and the equation enforcing the unit length of the quaternion have no direct physical meaning. Additionally, the resulting mass matrix is either singular or uses an arbitrarily chosen mass. Similarly, in [54, 62] a singular or in parts arbitrary mass matrix is obtained by starting as well with a unit quaternion and the angular velocity of a rigid body.

In contrast to previous works [12, 54, 62, 65, 84], the quaternion in this work is not assumed to be of unit length while deriving the equations of motion. The unit length restriction is added only in a last step in the form of a perfect bilateral constraint to reduce the scalable body to a rigid body. Besides deriving the non-singular  $7 \times 7$  mass matrix for a quaternion based rigid body formulation we also discuss the associated mechanical model in the form of the unconstrained scalable body.

In the recent work by Betsch and Siebert [12] the assumption of a unit length quaternion is introduced right at the beginning, but still a non-singular  $7 \times 7$  mass matrix is obtained. This is achieved by using a director-based formulation for the kinetic energy of a rigid body [13], which is identical to the kinetic energy of an affine body with all twelve affine degrees of freedom. This kinetic energy still contains the contributions from the scaling degree of freedom, contrary to the kinetic energy based on the angular velocity. By reducing the directors based generalized coordinates to a unit quaternion – without using the unit length property in a harmful way – a non-singular  $7 \times 7$  quaternion based rigid

body mass matrix is obtained. This mass matrix is identical to the one associated with the full unrestricted quaternion degrees of freedom of a scalable body, but the link to the mechanical model of a scalable body is not shown in [12]. When formulating the inverse mass matrix, [12] uses a simplification valid only for unit quaternions. This leads to a DAE formulation where the Lagrange multiplier is not the mechanical scaling constraint force: Setting the multiplier to zero does not recover the full dynamics of a scalable body. Beside this, in [12] the Lagrange multiplier is always zero for a free rigid body, even when the body is rotating. For a rotating rigid body one would expect a non-zero scaling constraint force preventing the body from getting larger.

In the work by Vadali [84], the quaternion unit length restriction is not imposed everywhere in the derivation of the equations of motion, but still the kinetic energy of the rigid body is used as a starting point. This kinetic energy based on the angular velocity does not contain the contributions from the scaling degree of freedom anymore, thus it is valid only under the quaternion unit length assumption. The result is a singular  $4 \times 4$  mass matrix for the rotational dynamics of a rigid body and a Lagrange multiplier which is always zero.

This section is organized as follows: In section 4.3.1, an overview on quaternions and their matrix and vector representation is given. The parametrization of rotations and uniform scaling with quaternions is described in section 4.3.2. The kinematics of the scalable body on displacement and velocity level is derived in section 4.3.3. In section 4.3.4 the variational equations of motion from section 3.2 are evaluated with the kinematics of the scalable body from section 4.2.1. To the resulting equations of motion, a perfect bilateral constraint is added in section 4.3.5, yielding the DAE formulation of a rigid body. In section 4.3.6, the derivative of the quaternion is replaced by a scaling velocity and a generalized angular velocity, giving the DAE a form which can be directly linked to the Newton-Euler equations of motion of a rigid body.

### 4.3.1 Quaternions

In this section, a brief introduction to quaternions is given. For more background on quaternions the reader is referred to [35, 7, 42, 85].

A quaternion  $A \in \mathbb{H}$  is a hypercomplex number with one real and three imaginary parts. The imaginary parts are formed with three real coefficients and three imaginary units  $i, j, k$ , i.e.

$$A = a_0 + a_1i + a_2j + a_3k \in \mathbb{H}, \quad a_i \in \mathbb{R}. \quad (4.29)$$

In this work, the notation  $A = (a_0, \mathbf{a})$  for a quaternion is used, where  $a_0$  is the real part and  $\mathbf{a} = (a_1 \ a_2 \ a_3)^\top \in \mathbb{R}^3$  is a vector consisting of the three coefficients of the imaginary part. The conjugate  $A^*$  of a quaternion is defined as

$$A^* = (a_0, -\mathbf{a}). \quad (4.30)$$

The addition of two quaternions is associative and is done component-by-component. The real part  $\text{Re}(A)$  and the imaginary part  $\text{Im}(A)$  of a quaternion is given by

$$\text{Re}(A) = (a_0, 0) = \frac{1}{2}(A + A^*) = a_0, \quad \text{Im}(A) = (0, \mathbf{a}) = \frac{1}{2}(A - A^*). \quad (4.31)$$

All possible products of the imaginary units can be determined from the definition

$$i^2 = j^2 = k^2 = ijk = -1 \quad (4.32)$$

as formulated in [35]. This gives the rule

$$AB = (a_0, \mathbf{a})(b_0, \mathbf{b}) = (a_0 b_0 - \mathbf{a}^\top \mathbf{b}, a_0 \mathbf{b} + b_0 \mathbf{a} + \tilde{\mathbf{a}} \mathbf{b}) \quad (4.33)$$

for the multiplication of two quaternions, where  $\tilde{\mathbf{a}} \in \mathbb{R}^{3 \times 3}$  is the real skew-symmetric matrix associated with the cross product, so that  $\tilde{\mathbf{a}} \mathbf{b} = \mathbf{a} \times \mathbf{b}$  for any  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ . The quaternion multiplication is not commutative in general. The product of a real number and a quaternion is commutative and yields a quaternion scaled component-by-component

$$\alpha A = A \alpha = (\alpha, 0) A = A (\alpha, 0) = (\alpha a_0, \alpha \mathbf{a}). \quad (4.34)$$

For the addition and multiplication of quaternions the distributive law holds. The conjugate of a quaternion product is the product of the conjugates in inverse order, i.e.

$$(AB)^* = B^* A^*. \quad (4.35)$$

The norm of a quaternion is defined by

$$|A| = \sqrt{a_0^2 + \mathbf{a}^\top \mathbf{a}}. \quad (4.36)$$

The norm of a quaternions and its conjugate are identical,

$$|A| = |A^*|, \quad (4.37)$$

as is obvious from the definition. The norm of the product of two quaternions equals the product of the norm of each, i.e.

$$|AB| = |A||B|. \quad (4.38)$$

The product of a quaternion and its conjugate is equal to the square of the norm of the quaternion, i.e.

$$AA^* = (|A|^2, 0) = |A|^2. \quad (4.39)$$

This can be used to form the inverse of a quaternion as

$$A^{-1} = \frac{A^*}{|A|^2} \quad (4.40)$$

for any non-zero quaternion. A quaternion can be mapped to a real  $4 \times 4$  matrix with the function

$$\varphi : \mathbb{H} \rightarrow \mathbb{R}^{4 \times 4}, \quad \varphi((a_0, \mathbf{a})) = \begin{pmatrix} a_0 & -\mathbf{a}^\top \\ \mathbf{a} & a_0 \mathbf{I} + \tilde{\mathbf{a}} \end{pmatrix}. \quad (4.41)$$

Note that the matrix of a quaternion is equal to the sum of the diagonal matrix  $a_0 \mathbf{I}$  and a skew symmetric matrix formed with the imaginary part of the quaternion, i.e.

$$\varphi((a_0, \mathbf{a})) = a_0 \mathbf{I} + \begin{pmatrix} 0 & -\mathbf{a}^\top \\ \mathbf{a} & \tilde{\mathbf{a}} \end{pmatrix}. \quad (4.42)$$

From this splitting one can directly see, that the matrix of a conjugated quaternion is the transposed of the matrix of a quaternion, i.e.

$$\varphi(A^*) = \varphi^\top(A). \quad (4.43)$$

The matrix of a product of two quaternions is equal to the product of the matrices of the quaternions

$$\varphi(AB) = \varphi(A)\varphi(B). \quad (4.44)$$

The matrix of a quaternion sum is equal to the sum of the matrices of each quaternion summand

$$\varphi(A + B) = \varphi(A) + \varphi(B), \quad (4.45)$$

because the matrix of a quaternion is linear in the coefficients of the quaternion. The inverse of the matrix of a quaternion can be obtained by mapping the inverse of the quaternion to a matrix or by normalizing the transposed of the matrix

$$\varphi^{-1}(A) = \varphi(A^{-1}) = \frac{1}{|A|^2} \varphi^\top(A). \quad (4.46)$$

Sometimes it is useful to interpret a quaternion as a real 4-dimensional vector, for which the function

$$\psi : \mathbb{H} \rightarrow \mathbb{R}^4, \quad \psi((a_0, \mathbf{a})) = \begin{pmatrix} a_0 \\ \mathbf{a} \end{pmatrix} \quad (4.47)$$

is introduced. The vector representation of the product of two quaternions is then equal to the product of the matrix of the first quaternion and the vector of the second quaternion, i.e.

$$\psi(AB) = \varphi(A)\psi(B). \quad (4.48)$$

Obviously the vector of a quaternion sum is equal to the sum of the vectors of each quaternion summand

$$\psi(A + B) = \psi(A) + \psi(B). \quad (4.49)$$

The vector of a conjugated quaternion can be obtained by multiplying the vector of the quaternion with the matrix

$$\mathbf{T} := \begin{pmatrix} 1 & 0 \\ 0 & -\mathbf{I} \end{pmatrix} \quad (4.50)$$

which yields

$$\psi(A^*) = \mathbf{T} \psi(A). \quad (4.51)$$

### 4.3.2 Rotation and Scaling

In this section, the parameterization of rotations and uniform scaling by using quaternions is shown. As a starting point we note, that the product of a quaternion  $A \in \mathbb{H}$ , a purely imaginary quaternion  $(0, \mathbf{x})$  generated by a vector  $\mathbf{x} \in \mathbb{R}^3$ , and the conjugate  $A^*$  of the

first quaternion always evaluates to a purely imaginary quaternion. This property can be formulated as

$$(0, \mathbf{y}) = A(0, \mathbf{x}) A^* \quad (4.52)$$

and follows directly with the help of equation (4.31):

$$\operatorname{Re}(A(0, \mathbf{x}) A^*) = \frac{1}{2}(A(0, \mathbf{x}) A^* + A(0, -\mathbf{x}) A^*) = \frac{1}{2}A((0, \mathbf{x}) + (0, -\mathbf{x}))A^* = 0. \quad (4.53)$$

The product in equation (4.52) can be rewritten by using equation (4.35),

$$(0, \mathbf{y}) = A(A(0, \mathbf{x})^*)^*, \quad (4.54)$$

and after applying (4.47), (4.48) and (4.51) one obtains the linear relation

$$\begin{pmatrix} 0 \\ \mathbf{y} \end{pmatrix} = \varphi(A) \mathbf{T} \varphi(A) \mathbf{T} \begin{pmatrix} 0 \\ \mathbf{x} \end{pmatrix} \quad (4.55)$$

in vector notation. Multiplication of (4.55) by  $\begin{pmatrix} 0 & \mathbf{I} \end{pmatrix}$  from the left yields

$$\mathbf{y} = \underbrace{\begin{pmatrix} 0 & \mathbf{I} \end{pmatrix} \varphi(A)}_{\begin{pmatrix} \mathbf{a} & a_0 \mathbf{I} + \tilde{\mathbf{a}} \end{pmatrix}} \underbrace{\mathbf{T} \varphi(A) \mathbf{T} \begin{pmatrix} 0 & \mathbf{I} \end{pmatrix}^\top}_{\begin{pmatrix} \mathbf{a}^\top \\ a_0 \mathbf{I} + \tilde{\mathbf{a}} \end{pmatrix}} \mathbf{x}. \quad (4.56)$$

By using the abbreviations

$$\mathbf{R} := \frac{1}{|A|^2} \begin{pmatrix} \mathbf{a} & a_0 \mathbf{I} + \tilde{\mathbf{a}} \end{pmatrix} \begin{pmatrix} \mathbf{a}^\top \\ a_0 \mathbf{I} + \tilde{\mathbf{a}} \end{pmatrix}, \quad s := |A|^2, \quad |A| \neq 0 \quad (4.57)$$

equation (4.56) finally becomes

$$\mathbf{y} = s \mathbf{R} \mathbf{x}. \quad (4.58)$$

The matrix  $\mathbf{R}$  is a rotation matrix as it has the properties

$$\mathbf{R} \mathbf{R}^\top = \mathbf{R}^\top \mathbf{R} = \mathbf{I}, \quad \operatorname{Det}(\mathbf{R}) = 1, \quad (4.59)$$

which can be verified by using the definition (4.57) and the identity

$$\tilde{\mathbf{x}} \tilde{\mathbf{y}} \equiv \mathbf{y} \mathbf{x}^\top - \mathbf{x}^\top \mathbf{y} \mathbf{I}. \quad (4.60)$$

The identity (4.60) is the vector triple product expansion, also known as Lagrange's formula, written in matrix notation. It follows that the product  $A(0, \mathbf{x}) A^*$  evaluates to a quaternion  $(0, \mathbf{y})$  for which the vector  $\mathbf{y}$  is the result of rotating and scaling the vector  $\mathbf{x}$  by  $\mathbf{R}$  and  $s$ , respectively. As a consequence, for any quaternion  $A \in \mathbb{H}$ ,  $A \neq 0$  there exists a rotation matrix  $\mathbf{R} \in \operatorname{SO}(3)$  and a scaling factor  $s \in \mathbb{R}^+$  such that

$$A(0, \mathbf{x}) A^* = (0, s \mathbf{R} \mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^3. \quad (4.61)$$

To deduce the inverse, we take an arbitrary rotation matrix  $\mathbf{R} \in \text{SO}(3)$  and show the existence of a quaternion  $A \in \mathbb{H}$  such that (4.61) holds. Any rotation matrix  $\mathbf{R} \in \text{SO}(3)$  may be represented as

$$\mathbf{R} = \mathbf{I} + \tilde{\mathbf{n}} \sin \varphi + \tilde{\mathbf{n}}^2 (1 - \cos \varphi) \quad (4.62)$$

with  $\mathbf{n} \in \mathbb{R}^3$  being the axis of rotation ( $|\mathbf{n}| = 1$ ) and  $\varphi$  the rotation angle. Since

$$(0, s\mathbf{R}) = (0, s(\mathbf{I} + \tilde{\mathbf{n}} \sin \varphi + \tilde{\mathbf{n}}^2 (1 - \cos \varphi))) = A(0, \mathbf{x})A^*, \quad \forall \mathbf{x} \in \mathbb{R}^3 \quad (4.63)$$

is fulfilled for any of the two quaternions

$$A = \pm \left( \sqrt{s} \cos \frac{\varphi}{2}, \mathbf{n} \sqrt{s} \sin \frac{\varphi}{2} \right), \quad (4.64)$$

one already has proven this assertion, i.e. that for any rotation matrix  $\mathbf{R} \in \text{SO}(3)$  and scaling factor  $s \in \mathbb{R}^+$  there exists a quaternion  $A \in \mathbb{H}$  such that

$$(0, s\mathbf{R}\mathbf{x}) = A(0, \mathbf{x})A^*, \quad \forall \mathbf{x} \in \mathbb{R}^3. \quad (4.65)$$

It has to be noted that the associated mapping  $s\mathbf{R} \rightarrow A$  is not unique, because the two quaternions  $A$  and  $-A$  yield the same rotation and scaling.

### 4.3.3 Kinematics

In this section, the kinematics of the scalable body with three translational, three rotational and one uniform scaling degree of freedom is described. Every point  $P'$  of the scalable body in a reference configuration can be addressed by a fixed vector  $\boldsymbol{\rho} \in \mathbb{R}^3$  ( $\boldsymbol{\rho} = \text{const.}$ ) starting from a reference point  $C'$  (cf. Figure 4.3). The actual position  $P$  of a

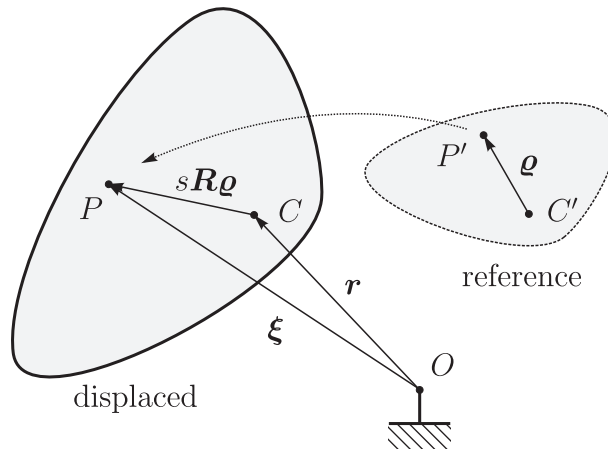


Figure 4.3: Scalable body kinematics.

point in a displaced configuration is described by the vector  $\boldsymbol{\xi} \in \mathbb{R}^3$  starting at the inertial point  $O$ . The vector  $\boldsymbol{\xi}$  can be obtained by applying a rotation  $\mathbf{R} \in \text{SO}(3)$  and a



scaling  $s \in \mathbb{R}^+$  on the vector  $\boldsymbol{\rho}$  and adding a displacement  $\boldsymbol{r} \in \mathbb{R}^3$  for the translational degrees of freedom. This yields the kinematic relation

$$\boldsymbol{\xi} = s\mathbf{R}\boldsymbol{\rho} + \boldsymbol{r} \quad (4.66)$$

for the scalable body. To parametrize the rotation  $\mathbf{R}$  and the scaling  $s$  a quaternion

$$A = (a_0, \mathbf{a}) \in \mathbb{H}, \quad |A| \neq 0 \quad (4.67)$$

is used, which allows to reformulate the kinematic relation (4.66) in quaternion notation

$$(0, \boldsymbol{\xi}) = A(0, \boldsymbol{\rho})A^* + (0, \boldsymbol{r}) \quad (4.68)$$

by using the results from Section 4.3.2. The displacement  $\boldsymbol{r}$  and the components of the quaternion  $A$  are grouped into a generalized coordinates vector

$$\mathbf{q} := \begin{pmatrix} \boldsymbol{r} \\ a_0 \\ \mathbf{a} \end{pmatrix}. \quad (4.69)$$

The absolute velocity  $\dot{\boldsymbol{\xi}}$  of a point on the body is obtained by differentiating equation (4.68) with respect to time,

$$\begin{aligned} (0, \dot{\boldsymbol{\xi}}) &= \dot{A}(0, \boldsymbol{\rho})A^* + A(0, \boldsymbol{\rho})\dot{A}^* + (0, \dot{\boldsymbol{r}}) \\ &= \frac{AA^*}{|A|^2} \dot{A}(0, \boldsymbol{\rho})A^* + A(0, \boldsymbol{\rho})\dot{A}^* \frac{AA^*}{|A|^2} + (0, \dot{\boldsymbol{r}}) \\ &= \frac{A}{|A|} \left( (0, \boldsymbol{\rho})\dot{A}^*A - A^*\dot{A}(0, -\boldsymbol{\rho}) \right) \frac{A^*}{|A|} + (0, \dot{\boldsymbol{r}}) \\ &= \frac{A}{|A|} \operatorname{Im} \left( (0, \boldsymbol{\rho})(2A^*\dot{A})^* \right) \frac{A^*}{|A|} + (0, \dot{\boldsymbol{r}}). \end{aligned} \quad (4.70)$$

Here it is useful to introduce new variables

$$\mathbf{v} := \dot{\boldsymbol{r}}, \quad (\nu, \boldsymbol{\omega}) := 2A^*\dot{A} \quad (4.71)$$

for the terms containing the derivatives of the generalized coordinates. A geometric interpretation of  $\mathbf{v}$ ,  $\nu$  and  $\boldsymbol{\omega}$  is given at the end of this section. The new variables  $\mathbf{v}$ ,  $\nu$  and  $\boldsymbol{\omega}$  can then be taken to define the vector of generalized velocities as

$$\mathbf{u} := \begin{pmatrix} \mathbf{v} \\ \nu \\ \boldsymbol{\omega} \end{pmatrix} = \underbrace{\begin{pmatrix} \mathbf{I} & 0 \\ 0 & 2\boldsymbol{\varphi}(A^*) \end{pmatrix}}_{=: \mathbf{Q}} \dot{\mathbf{q}} \quad (4.72)$$

and relate them to the derivative of the generalized coordinates as shown. The matrix  $\mathbf{Q} \in \mathbb{R}^{7 \times 7}$  is regular for any  $A \neq 0$  and can be obtained by rewriting (4.71) with the

help of (4.48). In terms of the generalized velocities  $\mathbf{u}$ , equation (4.70) becomes

$$\begin{aligned}
(0, \dot{\xi}) &= \frac{A}{|A|} \operatorname{Im}((0, \boldsymbol{\rho})(\nu, -\boldsymbol{\omega})) \frac{A^*}{|A|} + (0, \mathbf{v}) \\
&= \frac{A}{|A|} \operatorname{Im}((\boldsymbol{\rho}^\top \boldsymbol{\omega}, \boldsymbol{\rho}\nu - \tilde{\boldsymbol{\rho}}\boldsymbol{\omega})) \frac{A^*}{|A|} + (0, \mathbf{v}) \\
&= \frac{A}{|A|} (0, \boldsymbol{\rho}\nu - \tilde{\boldsymbol{\rho}}\boldsymbol{\omega}) \frac{A^*}{|A|} + (0, \mathbf{v}) \\
&= (0, \mathbf{R}\boldsymbol{\rho}\nu - \mathbf{R}\tilde{\boldsymbol{\rho}}\boldsymbol{\omega}) + (0, \mathbf{v}) \\
&= (0, (\mathbf{I} \quad \mathbf{R}\boldsymbol{\rho} \quad -\mathbf{R}\tilde{\boldsymbol{\rho}})\mathbf{u}),
\end{aligned} \tag{4.73}$$

where the quaternion product has been removed with the help of (4.61). Two purely imaginary quaternions are equal when their imaginary components are equal. This yields the equation

$$\dot{\xi} = (\mathbf{I} \quad \mathbf{R}\boldsymbol{\rho} \quad -\mathbf{R}\tilde{\boldsymbol{\rho}})\mathbf{Q}\dot{\mathbf{q}} \tag{4.74}$$

relating the absolute velocity  $\dot{\xi}$  of a point of the scalable body to the derivative of the generalized velocities  $\dot{\mathbf{q}}$ .

The generalized velocity defined in (4.72) consists of the vector  $\mathbf{v}$ , the scalar  $\nu$  and the vector  $\boldsymbol{\omega}$ . The velocity  $\mathbf{v}$  is the absolute velocity of the point  $C$ , which can be seen directly from the definition. To get an interpretation of  $\nu$  and  $\boldsymbol{\omega}$ , their definition (4.71) can be reformulated in terms of  $\dot{s}$  and  $\dot{\mathbf{R}}$ . First, the absolute velocity  $\dot{\xi}$  expressed in  $\dot{s}$  and  $\dot{\mathbf{R}}$  is obtained by differentiating (4.66) with respect to time, yielding

$$\begin{aligned}
\dot{\xi} &= \dot{\mathbf{r}} + \dot{s}\mathbf{R}\boldsymbol{\rho} + s\dot{\mathbf{R}}\boldsymbol{\rho} \\
&= \dot{\mathbf{r}} + \mathbf{R}(\dot{s}\boldsymbol{\rho} + s\mathbf{R}^\top \dot{\mathbf{R}}\boldsymbol{\rho}).
\end{aligned} \tag{4.75}$$

On the other hand, the absolute velocity  $\dot{\xi}$  in terms of  $\nu$  and  $\boldsymbol{\omega}$  can be obtained from the last line of (4.73),

$$\dot{\xi} = \mathbf{v} + \mathbf{R}(\nu\boldsymbol{\rho} + \tilde{\boldsymbol{\omega}}\boldsymbol{\rho}). \tag{4.76}$$

Of course, both representations of the velocity field have to be equal for any point of the body, i.e.

$$\mathbf{v} + \mathbf{R}\nu\boldsymbol{\rho} + \mathbf{R}\tilde{\boldsymbol{\omega}}\boldsymbol{\rho} = \dot{\mathbf{r}} + \mathbf{R}\dot{s}\boldsymbol{\rho} + \mathbf{R}s\mathbf{R}^\top \dot{\mathbf{R}}\boldsymbol{\rho}, \quad \forall \boldsymbol{\rho} \in \mathbb{R}^3. \tag{4.77}$$

Solving this variational equation for  $\boldsymbol{\rho} = 0$  and  $\boldsymbol{\rho} \neq 0$  yields

$$\mathbf{v} = \dot{\mathbf{r}}, \quad \nu = \dot{s}, \quad \tilde{\boldsymbol{\omega}} = s\mathbf{R}^\top \dot{\mathbf{R}}. \tag{4.78}$$

Obviously  $\nu$  is the scalar scaling velocity associated with the scaling factor  $s$ . The absolute angular velocity  $\boldsymbol{\Omega}$  associated with a rotation  $\mathbf{R}$  is given by

$$\tilde{\boldsymbol{\Omega}} := \dot{\mathbf{R}}\mathbf{R}^\top. \tag{4.79}$$

Expressing  $\tilde{\boldsymbol{\omega}}$  in terms of  $\tilde{\boldsymbol{\Omega}}$ ,

$$\tilde{\boldsymbol{\omega}} = s\mathbf{R}^\top \dot{\mathbf{R}} = s\mathbf{R}^\top \dot{\mathbf{R}}\mathbf{R}^\top \mathbf{R} = s\mathbf{R}^\top \tilde{\boldsymbol{\Omega}}\mathbf{R}, \tag{4.80}$$

together with the rotational invariance of the cross product

$$(\mathbf{R}\mathbf{x})^\sim = \mathbf{R}\tilde{\mathbf{x}}\mathbf{R}^\top, \quad \forall \mathbf{x} \in \mathbb{R}^3, \mathbf{R} \in \text{SO}(3) \quad (4.81)$$

yields the matrix relation

$$\tilde{\boldsymbol{\omega}} = s(\mathbf{R}^\top \boldsymbol{\Omega})^\sim. \quad (4.82)$$

Removing the cross product operator one gets the equation

$$\boldsymbol{\omega} = s\mathbf{R}^\top \boldsymbol{\Omega}. \quad (4.83)$$

This means the vector  $\boldsymbol{\omega}$  is the angular velocity  $\boldsymbol{\Omega}$  associated with the rotation  $\mathbf{R}$ , scaled by the factor  $s$  and rotated with  $\mathbf{R}^\top$  from the displaced configuration back to the reference configuration.

### 4.3.4 Equations of Motion

The variational equations of motion described in Section 3.2 can be combined with the kinematics from Section 4.3.3 to obtain the equations of motion of the scalable body. In the following the variational formulation (3.20) based on the kinetic energy will be used to derive the equations of motion. Alternatively one could also use equation (3.13) or equation (3.18) directly. As a first step the definition of the kinetic energy (3.19) is evaluated with the absolute velocity (4.74), which yields

$$\begin{aligned} T &= \frac{1}{2} \int_{\mathcal{B}} \dot{\boldsymbol{\xi}}^\top \dot{\boldsymbol{\xi}} dm \\ &= \frac{1}{2} \dot{\mathbf{q}}^\top \mathbf{Q}^\top \underbrace{\int_{\mathcal{B}} \begin{pmatrix} \mathbf{I} & \mathbf{R}\boldsymbol{\rho} & -\mathbf{R}\tilde{\boldsymbol{\rho}} \\ \boldsymbol{\rho}^\top \mathbf{R}^\top & \boldsymbol{\rho}^\top \boldsymbol{\rho} & 0 \\ \tilde{\boldsymbol{\rho}} \mathbf{R}^\top & 0 & -\tilde{\boldsymbol{\rho}}^2 \end{pmatrix} dm}_{=: \mathbf{M}} \mathbf{Q} \dot{\mathbf{q}} \\ &= \frac{1}{2} \dot{\mathbf{q}}^\top \mathbf{Q}^\top \mathbf{M} \mathbf{Q} \dot{\mathbf{q}}. \end{aligned} \quad (4.84)$$

The arbitrary reference point  $C'$  introduced in Section 4.3.3 is chosen now to be identical with the center of mass of the mass distribution in the reference configuration. This means that the integral

$$\int_{\mathcal{B}} \boldsymbol{\rho} dm = 0 \quad (4.85)$$

is always zero. The two abbreviations for the mass and classical inertia tensor

$$m = \int_{\mathcal{B}} dm, \quad \boldsymbol{\Theta} = \int_{\mathcal{B}} -\tilde{\boldsymbol{\rho}}^2 dm \quad (4.86)$$

introduced in Section 4.1 represent the mass distribution in the body. The remaining integral

$$\int_{\mathcal{B}} \boldsymbol{\rho}^\top \boldsymbol{\rho} dm = \frac{1}{2} \text{Tr } \boldsymbol{\Theta} \quad (4.87)$$

is half the trace of the classical inertia tensor  $\Theta$  as can be verified easily. This finally yields the constant and symmetric mass matrix

$$\mathbf{M} = \begin{pmatrix} m\mathbf{I} & 0 & 0 \\ 0 & \frac{1}{2} \text{Tr } \Theta & 0 \\ 0 & 0 & \Theta \end{pmatrix}. \quad (4.88)$$

If the classical inertia tensor  $\Theta$  of the body is positive definite, then the mass matrix  $\mathbf{M}$  is positive definite as well. For the lower right submatrix of  $\mathbf{M}$  the abbreviation

$$\hat{\Theta} := \begin{pmatrix} \frac{1}{2} \text{Tr } \Theta & 0 \\ 0 & \Theta \end{pmatrix} \quad (4.89)$$

is introduced. The kinetic energy  $T$  as given in (4.84) is a quadratic form in  $\dot{\mathbf{q}}$ , from which the partial derivative  $\partial T / \partial \dot{\mathbf{q}}$  can be obtained directly. To calculate the partial derivative  $\partial T / \partial \mathbf{q}$  it is useful to note that in this case the kinetic energy can be formulated as the sum of a quadratic form in  $\mathbf{q}$  plus a term not depending on  $\mathbf{q}$ . One gets

$$\begin{aligned} T &= \frac{1}{2} \dot{\mathbf{q}}^\top \mathbf{Q}^\top \mathbf{M} \mathbf{Q} \dot{\mathbf{q}} \\ &= \frac{m}{2} \dot{\mathbf{r}}^\top \dot{\mathbf{r}} + 2 \psi^\top(\dot{A}) \varphi(A) \hat{\Theta} \varphi^\top(A) \psi(\dot{A}) \\ &= \frac{m}{2} \dot{\mathbf{r}}^\top \dot{\mathbf{r}} + 2 \psi^\top(A) \varphi(\dot{A}) \mathbf{T}^\top \hat{\Theta} \mathbf{T} \varphi^\top(\dot{A}) \psi(A) \\ &= \frac{m}{2} \dot{\mathbf{r}}^\top \dot{\mathbf{r}} + 2 \psi^\top(A) \varphi(\dot{A}) \hat{\Theta} \varphi^\top(\dot{A}) \psi(A) \\ &= \frac{m}{2} \dot{\mathbf{r}}^\top \dot{\mathbf{r}} + \frac{1}{2} \mathbf{q}^\top \dot{\mathbf{Q}}^\top \mathbf{M} \dot{\mathbf{Q}} \mathbf{q} \end{aligned} \quad (4.90)$$

where the identity

$$\varphi^\top(A) \psi(B) = \psi(A^*B) = \psi((B^*A)^*) = \mathbf{T} \psi(B^*A) = \mathbf{T} \varphi^\top(B) \psi(A) \quad (4.91)$$

has been used. Using these two representations of the kinetic energy one obtains

$$\left( \frac{\partial T}{\partial \dot{\mathbf{q}}} \right)^\top = \mathbf{Q}^\top \mathbf{M} \mathbf{Q} \dot{\mathbf{q}}, \quad \left( \frac{\partial T}{\partial \mathbf{q}} \right)^\top = \dot{\mathbf{Q}}^\top \mathbf{M} \dot{\mathbf{Q}} \mathbf{q} \quad (4.92)$$

for the partial derivatives. Inserting them into the variational equations of motion (3.20) yields

$$\delta \mathbf{q}^\top \left( \frac{d}{dt} (\mathbf{Q}^\top \mathbf{M} \mathbf{Q} \dot{\mathbf{q}}) - \dot{\mathbf{Q}}^\top \mathbf{M} \dot{\mathbf{Q}} \mathbf{q} - \mathbf{f}_q \right) = 0 \quad \forall \delta \mathbf{q}. \quad (4.93)$$

By evaluating the time derivative and the variation one gets the equations of motion

$$\mathbf{Q}^\top \mathbf{M} \mathbf{Q} \ddot{\mathbf{q}} + \dot{\mathbf{Q}}^\top \mathbf{M} \dot{\mathbf{Q}} \dot{\mathbf{q}} + \dot{\mathbf{Q}}^\top \mathbf{M} (\mathbf{Q} \dot{\mathbf{q}} - \dot{\mathbf{Q}} \mathbf{q}) - \mathbf{f}_q = 0 \quad (4.94)$$

of the scalable body. The only thing that remains to do, is to specify the force distribution  $d\mathbf{F}_q$  and to calculate from it the associated generalized force  $\mathbf{f}_q$ ,

$$\mathbf{f}_q = \int_B \left( \frac{\partial \xi}{\partial \mathbf{q}} \right)^\top d\mathbf{F}_q. \quad (4.95)$$

Before doing this, the force distribution is split up once more

$$d\mathbf{F}_q = d\mathbf{F}_e + d\mathbf{F}_g \quad (4.96)$$

into a portion  $d\mathbf{F}_e$  and a force distribution  $d\mathbf{F}_g$  which will be used in Section 4.3.5 to realize an additional perfect bilateral constraint that makes the scalable body rigid. According to (4.95), the generalized force associated with  $d\mathbf{F}_g$  is denoted by

$$\mathbf{f}_g := \int_{\mathcal{B}} \left( \frac{\partial \boldsymbol{\xi}}{\partial \mathbf{q}} \right)^\top d\mathbf{F}_g. \quad (4.97)$$

The partial derivative occurring in (4.95) can be obtained via relation (3.15) and the absolute velocity (4.74),

$$\frac{\partial \boldsymbol{\xi}}{\partial \mathbf{q}} = \frac{\partial \dot{\boldsymbol{\xi}}}{\partial \dot{\mathbf{q}}} = \begin{pmatrix} \mathbf{I} & \mathbf{R}\boldsymbol{\rho} & -\mathbf{R}\tilde{\boldsymbol{\rho}} \end{pmatrix} \mathbf{Q}. \quad (4.98)$$

The three integrals that result for  $d\mathbf{F}_e$  when putting (4.98) into (4.95) are

$$\mathbf{F} := \int_{\mathcal{B}} d\mathbf{F}_e, \quad S_C := \int_{\mathcal{B}} \boldsymbol{\rho}^\top \mathbf{R}^\top d\mathbf{F}_e, \quad \mathbf{M}_C := \int_{\mathcal{B}} \tilde{\boldsymbol{\rho}} \mathbf{R}^\top d\mathbf{F}_e \quad (4.99)$$

and have the following meaning: The vector  $\mathbf{F}$  is the resultant external force and the vector  $\mathbf{M}_C$  is the resultant moment with respect to the point  $C$ . The scalar  $S_C$  is the resultant scaling force with respect to point  $C$ . The resultant scaling force and the resultant moment are formed by rotating the external forces with  $\mathbf{R}^\top$  back from the displaced configuration to the reference configuration. With these abbreviations one gets the complete generalized force (4.95) as

$$\mathbf{f}_q = \mathbf{Q}^\top \begin{pmatrix} \mathbf{F} \\ S_C \\ \mathbf{M}_C \end{pmatrix} + \mathbf{f}_g. \quad (4.100)$$

Setting the resultant forces and the additional generalized force  $\mathbf{f}_g$  equal to zero results in a generalized force  $\mathbf{f}_q$  which is equal to zero as well. In this case the equations of motion (4.94) would describe the dynamics of a free scalable body.

### 4.3.5 Scaling Constraint

In this section, an additional perfect bilateral constraint is applied on the scalable body in order to make it rigid. The scaling is the only additional degree of freedom that makes the scalable body different from a rigid body. For a rigid body the scaling  $s$  as introduced in Section 4.3.3 is always equal to one. The corresponding constraint equation can be written by (4.57) as

$$g(\mathbf{q}) = |A|^2 - 1 = 0. \quad (4.101)$$

A force law of d'Alembert-Lagrange type

$$\delta \mathbf{q}^\top \mathbf{f}_g = 0 \quad \forall \delta \mathbf{q} \mid \delta g = 0 \quad (4.102)$$

will now be formulated to complete the description of the perfect bilateral constraint. The product  $\delta \mathbf{q}^\top \mathbf{f}_g$  is the virtual work done by the constraint force (cf. Section 3.2). The virtual work has to vanish for any virtual displacements induced by  $\delta \mathbf{q}$  that are compatible with the constraint (i.e.  $\delta g = 0$ ). A simplified illustration of this situation is shown in Figure 4.4. The relation between variations  $\delta g$  of the constraint and variations  $\delta \mathbf{q}$  of the

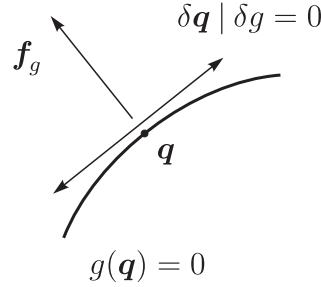


Figure 4.4: Constraint.

generalized coordinates is classically given by

$$\delta g = \frac{\partial g}{\partial \mathbf{q}} \delta \mathbf{q}. \quad (4.103)$$

Combining this with the force law (4.102) yields

$$\delta \mathbf{q}^\top \mathbf{f}_g = 0 \quad \forall \delta \mathbf{q} \mid \delta \mathbf{q}^\top \left( \frac{\partial g}{\partial \mathbf{q}} \right)^\top = 0. \quad (4.104)$$

Evaluating the variation reveals that the generalized constraint force  $\mathbf{f}_g$  lies in the linear subspace spanned by the vector  $(\partial g / \partial \mathbf{q})^\top$ . This can be formulated with (4.69) and (4.47) as

$$\mathbf{f}_g = \left( \frac{\partial g}{\partial \mathbf{q}} \right)^\top \lambda = \begin{pmatrix} 0 \\ 2\psi(A) \end{pmatrix} \lambda, \quad \lambda \in \mathbb{R}, \quad (4.105)$$

where  $\lambda$  is the scalar constraint force associated with the constraint. Inserting the generalized constraint force  $\mathbf{f}_g$  into the equation (4.94) finally yields with the help of (4.100) the DAE description

$$\mathbf{Q}^\top \mathbf{M} \mathbf{Q} \ddot{\mathbf{q}} + \mathbf{Q}^\top \mathbf{M} \dot{\mathbf{Q}} \dot{\mathbf{q}} + \dot{\mathbf{Q}}^\top \mathbf{M} (\mathbf{Q} \dot{\mathbf{q}} - \dot{\mathbf{Q}} \mathbf{q}) - \mathbf{Q}^\top \begin{pmatrix} \mathbf{F} \\ S_C + \lambda \\ \mathbf{M}_C \end{pmatrix} = 0, \quad |A|^2 = 1 \quad (4.106)$$

of the dynamics of a rigid body. In this formulation the unit length restriction of the quaternion is explicitly contained as algebraic constraint. The associated constraint force  $\lambda$  is mechanically consistent, as setting it to zero and dropping the constraint equation restores the unrestricted dynamics of the scalable body.

### 4.3.6 Generalized Velocities

The DAE formulation of the dynamics of a rigid body obtained in the last section can be further simplified by replacing the derivative of the generalized coordinates  $\dot{\mathbf{q}}$  with the generalized velocities  $\mathbf{u}$  as introduced in Section 4.3.3. As a first step the equations of motion (4.93) based on the principle of virtual work are replaced by an equivalent formulation based on the principle of virtual power

$$\underline{\delta \mathbf{q}}^\top \left( \frac{d}{dt} (\mathbf{Q}^\top \mathbf{M} \mathbf{Q} \dot{\mathbf{q}}) - \dot{\mathbf{Q}}^\top \mathbf{M} \dot{\mathbf{Q}} \mathbf{q} - \mathbf{f}_q \right) = 0 \quad \forall \underline{\delta \mathbf{q}}. \quad (4.107)$$

Next, the kinematic relation (4.72) is solved for the derivative of the generalized coordinates

$$\dot{\mathbf{q}} = \mathbf{Q}^{-1} \mathbf{u}. \quad (4.108)$$

The inverse  $\mathbf{Q}^{-1}$  can be obtained by inverting each block on the diagonal of  $\mathbf{Q}$ . With the help of relation (4.46) one gets

$$\mathbf{Q}^{-1} = \begin{pmatrix} \mathbf{I} & 0 \\ 0 & \frac{1}{2|A|^2} \boldsymbol{\varphi}(A) \end{pmatrix}. \quad (4.109)$$

Equation (4.108) applies in the same form for the virtual velocities,

$$\underline{\delta \mathbf{q}} = \mathbf{Q}^{-1} \delta \mathbf{u} \quad (4.110)$$

as follows directly from (3.22). Inserting (4.108) and (4.110) into (4.107) yields

$$\delta \mathbf{u}^\top \mathbf{Q}^{-\top} \frac{d}{dt} (\mathbf{Q}^\top \mathbf{M} \mathbf{u}) - \delta \mathbf{u}^\top \mathbf{Q}^{-\top} \dot{\mathbf{Q}}^\top \mathbf{M} \dot{\mathbf{Q}} \mathbf{q} - \delta \mathbf{u}^\top \mathbf{Q}^{-\top} \mathbf{f}_q = 0 \quad \forall \delta \mathbf{u}. \quad (4.111)$$

After having carried out the derivatives with respect to time one gets

$$\delta \mathbf{u}^\top \mathbf{M} \dot{\mathbf{u}} + \delta \mathbf{u}^\top \mathbf{Q}^{-\top} \dot{\mathbf{Q}}^\top \mathbf{M} (\mathbf{u} - \dot{\mathbf{Q}} \mathbf{q}) - \delta \mathbf{u}^\top \mathbf{Q}^{-\top} \mathbf{f}_q = 0 \quad \forall \delta \mathbf{u}. \quad (4.112)$$

To further simplify this equation, the product  $\mathbf{Q}^{-\top} \dot{\mathbf{Q}}^\top$  is evaluated in terms of the generalized velocities  $\mathbf{u}$  and the generalized coordinates  $\mathbf{q}$ . The matrix  $\mathbf{Q}^{-\top}$  can be obtained by transposing (4.109) and applying relation (4.43). Transposing the time derivative of (4.72) yields the matrix  $\dot{\mathbf{Q}}^\top$ . For the product one obtains

$$\begin{aligned} \mathbf{Q}^{-\top} \dot{\mathbf{Q}}^\top &= \begin{pmatrix} \mathbf{I} & 0 \\ 0 & \frac{1}{2|A|^2} \boldsymbol{\varphi}(A^*) \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 2\boldsymbol{\varphi}(\dot{A}) \end{pmatrix} = \frac{1}{2|A|^2} \begin{pmatrix} 0 & 0 \\ 0 & \boldsymbol{\varphi}(2A^* \dot{A}) \end{pmatrix} \\ &= \frac{1}{2|A|^2} \begin{pmatrix} 0 & 0 \\ 0 & \boldsymbol{\varphi}((\nu, \boldsymbol{\omega})) \end{pmatrix} = \frac{1}{2|A|^2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & \nu & -\boldsymbol{\omega}^\top \\ 0 & \boldsymbol{\omega} & \nu \mathbf{I} + \tilde{\boldsymbol{\omega}} \end{pmatrix} \end{aligned} \quad (4.113)$$

by using relations (4.41) and (4.44) as well as the kinematic relation (4.71) in quaternion notation. Similarly the product  $\dot{\mathbf{Q}} \mathbf{q}$  can be evaluated

$$\dot{\mathbf{Q}} \mathbf{q} = \begin{pmatrix} 0 & 0 \\ 0 & 2\boldsymbol{\varphi}(\dot{A}^*) \end{pmatrix} \begin{pmatrix} \mathbf{r} \\ \boldsymbol{\psi}(A) \end{pmatrix} = \begin{pmatrix} 0 \\ \boldsymbol{\psi}(2\dot{A}^* A) \end{pmatrix} = \begin{pmatrix} 0 \\ \boldsymbol{\psi}((\nu, \boldsymbol{\omega})^*) \end{pmatrix} = \begin{pmatrix} 0 \\ \nu \\ -\boldsymbol{\omega} \end{pmatrix}. \quad (4.114)$$

Combining the results from (4.113) and (4.114), one gets the simplification

$$\begin{aligned}
\mathbf{Q}^{-\top} \dot{\mathbf{Q}}^\top \mathbf{M}(\mathbf{u} - \dot{\mathbf{Q}}\mathbf{q}) &= \\
&= \frac{1}{2|A|^2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & \nu & -\boldsymbol{\omega}^\top \\ 0 & \boldsymbol{\omega} & \nu \mathbf{I} + \tilde{\boldsymbol{\omega}} \end{pmatrix} \begin{pmatrix} m\mathbf{I} & 0 & 0 \\ 0 & \frac{1}{2} \text{Tr } \boldsymbol{\Theta} & 0 \\ 0 & 0 & \boldsymbol{\Theta} \end{pmatrix} \begin{pmatrix} \mathbf{v} \\ 0 \\ 2\boldsymbol{\omega} \end{pmatrix} \\
&= \frac{1}{2|A|^2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & \nu & -\boldsymbol{\omega}^\top \\ 0 & \boldsymbol{\omega} & \nu \mathbf{I} + \tilde{\boldsymbol{\omega}} \end{pmatrix} \begin{pmatrix} m\mathbf{v} \\ 0 \\ 2\boldsymbol{\Theta}\boldsymbol{\omega} \end{pmatrix} = \frac{1}{|A|^2} \begin{pmatrix} 0 \\ -\boldsymbol{\omega}^\top \\ \nu \mathbf{I} + \tilde{\boldsymbol{\omega}} \end{pmatrix} \boldsymbol{\Theta}\boldsymbol{\omega}
\end{aligned} \tag{4.115}$$

for the product occurring in (4.112). Inserting the simplification (4.115) and the generalized force (4.100) together with the constraint force (4.105) into equation (4.112) one gets the variational equation

$$\delta \mathbf{u}^\top \mathbf{M} \dot{\mathbf{u}} + \delta \mathbf{u}^\top \frac{1}{|A|^2} \begin{pmatrix} 0 \\ -\boldsymbol{\omega}^\top \\ \nu \mathbf{I} + \tilde{\boldsymbol{\omega}} \end{pmatrix} \boldsymbol{\Theta}\boldsymbol{\omega} - \delta \mathbf{u}^\top \begin{pmatrix} \mathbf{F} \\ S_C + \lambda \\ \mathbf{M}_C \end{pmatrix} = 0 \quad \forall \delta \mathbf{u}. \tag{4.116}$$

Eliminating the variation and completing the set of equations with the kinematic relation (4.108) and the constraint equation (4.101), one gets the full DAE formulation for the dynamics of a rigid body

$$\begin{cases} m\dot{\mathbf{v}} = \mathbf{F} \\ \frac{1}{2} \text{Tr } \boldsymbol{\Theta} \dot{\nu} - \frac{1}{|A|^2} \boldsymbol{\omega}^\top \boldsymbol{\Theta}\boldsymbol{\omega} = S_C + \lambda \\ \boldsymbol{\Theta} \dot{\boldsymbol{\omega}} + \frac{1}{|A|^2} (\nu \mathbf{I} + \tilde{\boldsymbol{\omega}}) \boldsymbol{\Theta}\boldsymbol{\omega} = \mathbf{M}_C \end{cases}, \quad \begin{cases} \dot{\mathbf{r}} = \mathbf{v} \\ \dot{A} = \frac{1}{2|A|^2} A(\nu, \boldsymbol{\omega}), \quad |A|^2 = 1 \end{cases} \tag{4.117}$$

in terms of the generalized velocities  $\mathbf{v}$ ,  $\nu$ , and  $\boldsymbol{\omega}$ . This formulation is equivalent to equations (4.106). If the constraint force  $\lambda$  is set to zero and the constraint equation  $|A|^2 = 1$  is removed one recovers the full seven degree of freedom dynamics of the scalable body. The rigid body DAE (4.117) can be simplified to the classical Newton-Euler equations when the constraint is differentiated twice

$$s = |A|^2 = 1 \quad \Rightarrow \quad \dot{s} = \nu = 0, \quad \dot{\nu} = 0 \tag{4.118}$$

and all occurrences of  $|A|^2$ ,  $\nu$  and  $\dot{\nu}$  are eliminated. One then gets the ODE of a rigid body

$$\begin{cases} m\dot{\mathbf{v}} = \mathbf{F} \\ \boldsymbol{\Theta} \dot{\boldsymbol{\omega}} + \tilde{\boldsymbol{\omega}} \boldsymbol{\Theta}\boldsymbol{\omega} = \mathbf{M}_C \end{cases}, \quad \begin{cases} \dot{\mathbf{r}} = \mathbf{v} \\ \dot{A} = \frac{1}{2} A(0, \boldsymbol{\omega}) \end{cases} \tag{4.119}$$

and an equation to calculate the scaling constraint force as

$$\lambda = -\boldsymbol{\omega}^\top \boldsymbol{\Theta}\boldsymbol{\omega} - S_C. \tag{4.120}$$

The scaling constraint force  $\lambda$  has to balance the external resultant scaling force  $S_C$  and a term depending on the angular velocity. If there is no external resultant scaling force  $S_C$  and the body is not rotating, then a scalable body is identical to a rigid body.



Note that while it is very simple to reduce the rigid body DAE formulation (4.117) to the rigid body ODE formulation (4.119), the inverse way going from the ODE to the DAE based on the scalable body is not directly possible. In the Newton-Euler equations used in the ODE description of a rigid body, the scaling dynamics is no longer present. Exactly this scaling dynamics and its coupling to the Newton-Euler equations is missing when one tries to recover the scalable body based DAE of a rigid body from the Newton-Euler equations. Correspondingly the mass term  $1/2 \text{Tr } \Theta$  associated with the scaling dynamics in the second equation of (4.117) can be replaced by any other value if only the rigid body dynamics is to be described correctly.

The mechanical model of a scalable body itself might be rarely used in a technical application, due to the fact that a scalable body is rather complicated to build in reality. Nevertheless, the equations of motion of a scalable body can be valuable for the interpretation of the quaternion based rigid body DAE. The regular  $7 \times 7$  mass matrix used in the rigid body DAE formulation is exactly the mass matrix of the scalable body. The Lagrangian multiplier associated with the quaternion unit length constraint has the meaning of a constraint force, preventing the scaling body from changing its size. While already equation (4.106) completely describes the dynamics of a rigid body in DAE form, it is difficult to see the connection to the Newton-Euler equations (4.119). This is the reason why the equations of motion of a scalable body have been reformulated to (4.117) using angular and scaling velocities. The rigid body DAE formulation (4.117) can be directly recognized as a DAE generalization of the Newton-Euler equations (4.119), while at the same time being the complete description of the dynamics of a scalable body when removing the explicitly contained constraint.

Note that the  $7 \times 7$  mass matrix obtained in this section for the scalable body is decomposable as assumed in Section 3.5. In the formulation for the equations of motion derived in Section 3.5 one can directly include scalable bodies, and with additional quaternion unit length constraints also rigid bodies can be described.



## Integrators

---

The equations of motion of a mechanical system can be solved with a numerical integration scheme. The numerical integration with respect to time can be formulated as the following time step problem: For given initial time  $t^B$ , initial generalized coordinates  $\mathbf{q}^B := \mathbf{q}(t^B)$  and initial generalized velocities  $\mathbf{u}^B := \mathbf{u}(t^B)$  find approximates of the generalized coordinates  $\mathbf{q}^E := \mathbf{q}(t^E)$  and generalized velocities  $\mathbf{u}^E := \mathbf{u}(t^E)$  at the end of a chosen time interval  $[t^B, t^E]$ . The length of the time step is denoted with  $\Delta t = t^E - t^B$ . Concatenating many time steps then yields an approximation of the solution over a longer time interval. The approximation of the end time generalized coordinates  $\mathbf{q}^E$  and end time generalized velocities  $\mathbf{u}^E$  has to be made such that in the limit of vanishing time steps  $\Delta t \rightarrow 0$  the approximation converges to the exact solution. In Section 5.1 an integrator based on Moreau's midpoint rule is described. This integrator will be used as reference for the consistent integrators. Different variants of consistent integrators are developed in Section 5.2.

### 5.1 Moreau's Midpoint Rule

In this section, the numerical integration of the dynamics of a non-smooth mechanical system with an integrator based on Moreau's midpoint rule is described. Moreau's midpoint rule is an event capturing time-stepping scheme developed by Moreau in [60, 61]. In an event capturing time-stepping scheme all non-smooth events in a time step, like for example impacts or stick-slip transitions, are handled with an implicit inclusion based discretization scheme, which does not resolve the exact point in time where the events occur. In contrast to this, event driven-methods try to separate the integration into piecewise smooth parts and discrete switching points for the non-smooth events. The equations of motion of the non-smooth mechanical system are assumed to have the form

$$\begin{aligned}
 \mathbf{M}(\mathbf{q}, t)\dot{\mathbf{u}} - \mathbf{h}(\mathbf{q}, \mathbf{u}, t) - \sum_{i \in \mathcal{I}(\mathbf{q}, t)} \mathbf{W}_i(\mathbf{q}, t)\boldsymbol{\lambda}_i &= 0, & \dot{\mathbf{q}} &= \mathbf{F}(\mathbf{q}, t)\mathbf{u} + \boldsymbol{\beta}(\mathbf{q}, t) \\
 \boldsymbol{\gamma}_i &= \mathbf{W}_i^T(\mathbf{q}, t)\mathbf{u} + \boldsymbol{\chi}_i(\mathbf{q}, t), & \boldsymbol{\gamma}_i &\in \mathcal{N}_{\mathcal{C}_i}(-\boldsymbol{\lambda}_i), & i &\in \mathcal{I}(\mathbf{q}, t).
 \end{aligned} \tag{5.1}$$

This is a combination of the equations of motion (3.31) derived in Section 3.2 and Section 3.3 with set-valued force laws of normal cone type as described in Section 3.7. The force laws in (5.1) are formulated on velocity level. If the model contains unilateral contacts on displacement level, then they can be considered by keeping only the indices  $i$  of contacts that are closed on displacement level  $g_i(\mathbf{q}, t) = 0$  in the set  $\mathcal{I}$  of active force laws. The equations of motion are completed with an impact equation and Newton impact laws in inclusion form

$$\begin{aligned} \mathbf{M}(\mathbf{q}, t)(\mathbf{u}^+ - \mathbf{u}^-) - \sum_{i \in \mathcal{I}(\mathbf{q}, t)} \mathbf{W}_i(\mathbf{q}, t) \boldsymbol{\Lambda}_i &= 0 \\ \boldsymbol{\gamma}_i^\pm &= \mathbf{W}_i^\top(\mathbf{q}, t) \mathbf{u}^\pm + \boldsymbol{\chi}_i(\mathbf{q}, t), \quad \boldsymbol{\gamma}_i^+ + \varepsilon_i \boldsymbol{\gamma}_i^- \in \mathcal{N}_{\mathcal{D}_i}(-\boldsymbol{\Lambda}_i), \quad i \in \mathcal{I}(\mathbf{q}, t). \end{aligned} \quad (5.2)$$

The equations of motion in inclusion form (5.1) and the impact inclusions (5.2) constitute together the model of the non-smooth mechanical system.

Moreau's midpoint rule uses one difference scheme to approximate both the equations of motion in inclusion form (5.1) and the impact inclusions (5.2). For the generalized coordinates a midpoint rule is used, while the velocities are discretized with a partially implicit Euler step. Starting from a known state  $\mathbf{u}^B = \mathbf{u}(t^B)$  and  $\mathbf{q}^B = \mathbf{q}(t^B)$  at the time  $t^B$  of a time step  $\Delta t$ , the coordinates  $\mathbf{q}^M = \mathbf{q}(t^M)$  at the midpoint are calculated

$$\mathbf{q}^M := \mathbf{q}^B + \frac{\Delta t}{2} (\mathbf{F}(\mathbf{q}^B, t^B) \mathbf{u}^B + \boldsymbol{\beta}(\mathbf{q}^B, t^B)), \quad t^M = t^B + \frac{\Delta t}{2}. \quad (5.3)$$

Using the midpoint coordinates  $\mathbf{q}^M$  and the velocities  $\mathbf{u}^B$  from the beginning of the time step one calculates  $\mathbf{M}(\mathbf{q}^M, t^M)$ ,  $\mathbf{h}(\mathbf{q}^M, \mathbf{u}^B, t^M)$ ,  $\mathbf{W}_i(\mathbf{q}^M, t^M)$ ,  $\boldsymbol{\chi}_i(\mathbf{q}^M, t^M)$  and sets up the index set  $\mathcal{I}(\mathbf{q}^M, t^M)$ . As a next step one has to solve the inclusion

$$\begin{aligned} \mathbf{M}(\mathbf{u}^E - \mathbf{u}^B) - \mathbf{h} \Delta t - \sum_{i \in \mathcal{I}} \mathbf{W}_i \boldsymbol{\Lambda}_i &= 0, \\ \boldsymbol{\gamma}_i^B &= \mathbf{W}_i^\top \mathbf{u}^B + \boldsymbol{\chi}_i, \quad \boldsymbol{\gamma}_i^E = \mathbf{W}_i^\top \mathbf{u}^E + \boldsymbol{\chi}_i, \\ \boldsymbol{\gamma}_i^E + \varepsilon_i \boldsymbol{\gamma}_i^B &\in \mathcal{N}_{\mathcal{D}_i}(-\boldsymbol{\Lambda}_i), \end{aligned} \quad (5.4)$$

which approximates the equations of motion and the impact inclusions over a time step  $\Delta t$ , for the end time velocities  $\mathbf{u}^E$  and the discrete impulses  $\boldsymbol{\Lambda}_i$ . Note that the  $\boldsymbol{\Lambda}_i$  in the discrete problem (5.4) approximate the sum of the impact-impulses and the time integral of the forces  $\boldsymbol{\lambda}_i$ . As a result of (5.4) one obtains the end time velocities  $\mathbf{u}^E$ . Finally the end time coordinates have to be calculated

$$\mathbf{q}^E := \mathbf{q}^M + \frac{\Delta t}{2} (\mathbf{F}(\mathbf{q}^M, t^M) \mathbf{u}^E + \boldsymbol{\beta}(\mathbf{q}^M, t^M)), \quad t^E = t^M + \frac{\Delta t}{2}. \quad (5.5)$$

The core problem of an integration step is the solution of the inclusion (5.4). By introducing the abbreviations

$$\boldsymbol{\xi}_i := \boldsymbol{\gamma}_i^E + \varepsilon_i \boldsymbol{\gamma}_i^B, \quad \mathbf{G}_{ij} := \mathbf{W}_i^\top \mathbf{M}^{-1} \mathbf{W}_j, \quad \mathbf{c}_i := \mathbf{W}_i^\top \mathbf{M}^{-1} \mathbf{h} \Delta t + (1 + \varepsilon_i) \boldsymbol{\gamma}_i^B \quad (5.6)$$

it can be rewritten as an inclusion problem of the form

$$\boldsymbol{\xi}_i \in \mathcal{N}_{\mathcal{D}_i}(-\boldsymbol{\Lambda}_i), \quad \boldsymbol{\xi}_i = \sum_{j \in \mathcal{I}} \mathbf{G}_{ij} \boldsymbol{\Lambda}_j + \mathbf{c}_i, \quad i \in \mathcal{I} \quad (5.7)$$

which has to be solved for the discrete impact-impulses  $\boldsymbol{\Lambda}_i$ . The end time velocities can then be calculated in an explicit way as

$$\mathbf{u}^E = \mathbf{u}^B + \mathbf{M}^{-1} \mathbf{h} \Delta t + \mathbf{M}^{-1} \sum_{i \in \mathcal{I}} \mathbf{W}_i \boldsymbol{\Lambda}_i. \quad (5.8)$$

The inclusion problem (5.7) consists of a normal cone inclusion and a linear equation for each set-valued force law. A numerical solution for the normal cone inclusions and linear equations (5.7) can be obtained with the iterative methods described in Section 2.5. An analysis of the convergence of Moreau's midpoint rule and a comparison with other time-stepping methods can be found in [78].

## 5.2 Consistent Integrators

The numerical integration of the dynamics of a mechanical system yields an exact solution in general only in the limit of a vanishing time step. If the mechanical system conserves the total energy (see Section 3.9), then this property is usually lost in a solution obtained by numerical integration. This means at some points in the numerical integration process errors are introduced into the solution that destroy the property. One such point is the numerical solution of a difference scheme. The difference scheme is solved only up to a certain tolerance and it is solved by using floating point numbers with a finite number of digits. In general, this finite precision error does not depend on the length of the time step and it can be controlled quite well. Unfortunately, the finite precision error is not the only source of errors. Assuming the difference scheme of the numerical integrator could be solved without error, one would still not get a numerical solution that exhibits all properties of the exact solution. The reason is the discretization errors introduced when replacing the differential formulation of the equations of motion with a difference scheme. In the consistent integrators discussed in this work, the aim is to preserve properties of the exact solution in the discretization process, so that they are present in exact solutions of the difference scheme as well. The finite precision errors are not addressed specifically in this work and they are not part of any consistency consideration.

An integrator that is consistent with respect to the total energy, preserves an energy conservation property in the difference scheme if that property is present in the differential formulation. This means, for a mechanical system with the property

$$\dot{E}(\mathbf{u}, \mathbf{q}) = 0, \quad E(\mathbf{u}^+, \mathbf{q}) = E(\mathbf{u}^-, \mathbf{q}) \quad (5.9)$$

an energy consistent integration scheme has the property

$$E(\mathbf{u}^E, \mathbf{q}^E) = E(\mathbf{u}^B, \mathbf{q}^B). \quad (5.10)$$

Of course this property has to be given not only for a vanishing time step, but for all lengths of time steps. Identically, if the mechanical system is dissipative

$$\dot{E}(\mathbf{u}, \mathbf{q}) \leq 0, \quad E(\mathbf{u}^+, \mathbf{q}) \leq E(\mathbf{u}^-, \mathbf{q}) \quad (5.11)$$

then an energy consistent integration scheme has the dissipation property

$$E(\mathbf{u}^E, \mathbf{q}^E) \leq E(\mathbf{u}^B, \mathbf{q}^B) \quad (5.12)$$

as well. For a non-smooth mechanical system with unilateral contacts one can define that the consistent discretization of

$$g(\mathbf{q}) = 0 : \dot{g}(\dot{\mathbf{q}}, \mathbf{q}) \geq 0 \quad (5.13)$$

has to have the property

$$g(\mathbf{q}^B) \leq 0 : g(\mathbf{q}^E) \geq g(\mathbf{q}^B). \quad (5.14)$$

This can be seen as a consistency condition for unilateral contacts on velocity level. It does not guarantee  $g(\mathbf{q}) \geq 0$  for a numerical solution, but it makes sure that once a contact is closed, the interpenetration does not further increase. A consistency condition can be formulated as well for bilateral constraints on displacement and on velocity level, where they guarantee that also in a numerical approximation the constraints are fulfilled exactly at the end of a time step. Some consistent integrators also preserve momentum properties (see for example [13]), but this is not considered in this work.

This section is organized as follows: In Section 5.2.1, the discrete derivative introduced by Gonzalez in [31, 32] is described. The discrete derivative is an important concept that is used in the following sections for the construction of consistent integration schemes. In Section 5.2.2, energy consistent integration schemes based on a constant mass matrix formulation are developed. In Section 5.2.3, an energy consistent integrator based on the coordinate dependent mass matrix formulation from Section 3.4 is given. Using the equations of motion based on the decomposed mass matrix from Section 3.5, a non-smooth and consistent integration scheme is formulated in Section 5.2.4. A quaternion based non-smooth and consistent integrator is described in Section 5.2.5.

### 5.2.1 Discrete Derivative

The discrete derivative as introduced by Gonzalez in [31, 32] is an important concept for the design of consistent integration schemes. Given a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , the discrete derivative is a mapping  $Df/D\mathbf{x} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  which can be used to approximate the partial derivative  $\partial f/\partial \mathbf{x} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ . If the arguments of the discrete derivative are  $\mathbf{a}$  and  $\mathbf{b}$ , then it approximates the partial derivative at the midpoint  $(\mathbf{a} + \mathbf{b})/2$ . Besides being an approximation of the partial derivative at the midpoint, the discrete derivative of the points  $\mathbf{a}$  and  $\mathbf{b}$  yields exactly the difference  $f(\mathbf{b}) - f(\mathbf{a})$  of the functions values of  $f$  when multiplied with the difference vector  $\mathbf{b} - \mathbf{a}$ . This directionality property can be written as

$$\frac{Df}{D\mathbf{x}}(\mathbf{a}, \mathbf{b})(\mathbf{b} - \mathbf{a}) = f(\mathbf{b}) - f(\mathbf{a}), \quad \forall \mathbf{a}, \mathbf{b}. \quad (5.15)$$

As in [31, 32], the consistency property ensuring a good approximation is formulated as

$$\frac{Df}{D\mathbf{x}}(\mathbf{a}, \mathbf{b}) = \frac{\partial f}{\partial \mathbf{x}}\left(\frac{\mathbf{a} + \mathbf{b}}{2}\right) + \mathcal{O}(\|\mathbf{b} - \mathbf{a}\|), \quad \forall \mathbf{a}, \mathbf{b} \quad (5.16)$$

where  $\|\cdot\|$  is the standard Euclidean norm. Note that the discrete derivative has to be evaluated with two arguments and that the exact difference of the function values is obtained only when multiplying with the difference of the points used as arguments. When multiplying with the difference to other points one gets only an approximation for the difference of the function values,

$$\frac{Df}{D\mathbf{x}}(\mathbf{a}, \mathbf{b})(\mathbf{x} - \mathbf{a}) \approx f(\mathbf{x}) - f(\mathbf{a}), \quad (5.17)$$

as when multiplying the partial derivative with a difference vector

$$\frac{\partial f}{\partial \mathbf{x}}(\mathbf{a})(\mathbf{x} - \mathbf{a}) \approx f(\mathbf{x}) - f(\mathbf{a}). \quad (5.18)$$

An illustration of the discrete derivative and the partial derivative of a function is given in Figure 5.1.

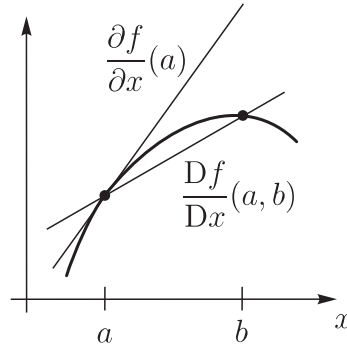


Figure 5.1: Discrete derivative and partial derivative.

For any smooth function  $f$ , the discrete derivative can be constructed with the relation

$$\frac{Df}{D\mathbf{x}}(\mathbf{a}, \mathbf{b}) = \frac{\partial f}{\partial \mathbf{x}}\left(\frac{\mathbf{a} + \mathbf{b}}{2}\right) + \frac{f(\mathbf{b}) - f(\mathbf{a}) - \frac{\partial f}{\partial \mathbf{x}}\left(\frac{\mathbf{a} + \mathbf{b}}{2}\right)(\mathbf{b} - \mathbf{a})}{\|\mathbf{b} - \mathbf{a}\|^2}(\mathbf{b} - \mathbf{a})^\top. \quad (5.19)$$

To verify that this construction always has the directionality property (5.15) we multiply (5.19) with  $(\mathbf{b} - \mathbf{a})$  from the right and obtain

$$\begin{aligned} \frac{Df}{D\mathbf{x}}(\mathbf{a}, \mathbf{b})(\mathbf{b} - \mathbf{a}) &= \frac{\partial f}{\partial \mathbf{x}}\left(\frac{\mathbf{a} + \mathbf{b}}{2}\right)(\mathbf{b} - \mathbf{a}) + \frac{f(\mathbf{b}) - f(\mathbf{a}) - \frac{\partial f}{\partial \mathbf{x}}\left(\frac{\mathbf{a} + \mathbf{b}}{2}\right)(\mathbf{b} - \mathbf{a})}{\|\mathbf{b} - \mathbf{a}\|^2}\|\mathbf{b} - \mathbf{a}\|^2 \\ &= \frac{\partial f}{\partial \mathbf{x}}\left(\frac{\mathbf{a} + \mathbf{b}}{2}\right)(\mathbf{b} - \mathbf{a}) + f(\mathbf{b}) - f(\mathbf{a}) - \frac{\partial f}{\partial \mathbf{x}}\left(\frac{\mathbf{a} + \mathbf{b}}{2}\right)(\mathbf{b} - \mathbf{a}) \\ &= f(\mathbf{b}) - f(\mathbf{a}). \end{aligned} \quad (5.20)$$

To verify that the construction (5.19) has the consistency property (5.16) we use the Taylor series of the function  $f$ , i.e.

$$f(\mathbf{x}) = f(\mathbf{y}) + \frac{\partial f}{\partial \mathbf{x}}(\mathbf{y})(\mathbf{x} - \mathbf{y}) + \frac{1}{2}(\mathbf{x} - \mathbf{y})^\top \frac{\partial^2 f}{\partial \mathbf{x}^2}(\mathbf{y})(\mathbf{x} - \mathbf{y}) + \mathcal{O}(\|\mathbf{x} - \mathbf{y}\|^3). \quad (5.21)$$

We introduce the abbreviations

$$\boldsymbol{\mu} = \frac{\mathbf{a} + \mathbf{b}}{2}, \quad \boldsymbol{\sigma} = \frac{\mathbf{b} - \mathbf{a}}{2} \quad (5.22)$$

for the midpoint  $\boldsymbol{\mu}$  and half the difference  $\boldsymbol{\sigma}$  between  $\mathbf{a}$  and  $\mathbf{b}$  to simplify the notation. With these we can now write the Taylor series of  $\mathbf{a}$  and  $\mathbf{b}$  based on the midpoint

$$\begin{aligned} f(\mathbf{a}) &= f(\boldsymbol{\mu} - \boldsymbol{\sigma}) = f(\boldsymbol{\mu}) - \frac{\partial f}{\partial \mathbf{x}}(\boldsymbol{\mu})\boldsymbol{\sigma} + \frac{1}{2}\boldsymbol{\sigma}^\top \frac{\partial^2 f}{\partial \mathbf{x}^2}(\boldsymbol{\mu})\boldsymbol{\sigma} + \mathcal{O}(\|\boldsymbol{\sigma}\|^3) \\ f(\mathbf{b}) &= f(\boldsymbol{\mu} + \boldsymbol{\sigma}) = f(\boldsymbol{\mu}) + \frac{\partial f}{\partial \mathbf{x}}(\boldsymbol{\mu})\boldsymbol{\sigma} + \frac{1}{2}\boldsymbol{\sigma}^\top \frac{\partial^2 f}{\partial \mathbf{x}^2}(\boldsymbol{\mu})\boldsymbol{\sigma} + \mathcal{O}(\|\boldsymbol{\sigma}\|^3). \end{aligned} \quad (5.23)$$

Subtracting the first equation from the second yields

$$f(\mathbf{b}) - f(\mathbf{a}) = 2\frac{\partial f}{\partial \mathbf{x}}(\boldsymbol{\mu})\boldsymbol{\sigma} + \mathcal{O}(\|\boldsymbol{\sigma}\|^3). \quad (5.24)$$

Replacing  $\boldsymbol{\mu}$  and  $\boldsymbol{\sigma}$  we get the estimate

$$f(\mathbf{b}) - f(\mathbf{a}) - \frac{\partial f}{\partial \mathbf{x}}\left(\frac{\mathbf{a} + \mathbf{b}}{2}\right)(\mathbf{b} - \mathbf{a}) = \mathcal{O}(\|\mathbf{b} - \mathbf{a}\|^3) \quad (5.25)$$

for the expression occurring above the fraction bar in (5.19). Inserting (5.25) into (5.19) one obtains

$$\begin{aligned} \frac{Df}{D\mathbf{x}}(\mathbf{a}, \mathbf{b}) &= \frac{\partial f}{\partial \mathbf{x}}\left(\frac{\mathbf{a} + \mathbf{b}}{2}\right) + \frac{\mathcal{O}(\|\mathbf{b} - \mathbf{a}\|^3)}{\|\mathbf{b} - \mathbf{a}\|^2}(\mathbf{b} - \mathbf{a})^\top \\ &= \frac{\partial f}{\partial \mathbf{x}}\left(\frac{\mathbf{a} + \mathbf{b}}{2}\right) + \mathcal{O}(\|\mathbf{b} - \mathbf{a}\|)(\mathbf{b} - \mathbf{a})^\top \\ &= \frac{\partial f}{\partial \mathbf{x}}\left(\frac{\mathbf{a} + \mathbf{b}}{2}\right) + \mathcal{O}(\|\mathbf{b} - \mathbf{a}\|^2). \end{aligned} \quad (5.26)$$

As one can see, the fraction in (5.19) is well behaved, as the numerator vanishes before the denominator. The construction (5.19) also fulfills the consistency property (5.16).

If the function is quadratic

$$g(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{G}\mathbf{x} + \mathbf{c}^\top \mathbf{x} + \mathbf{d}, \quad \mathbf{G}^\top = \mathbf{G} \quad (5.27)$$

then the discrete derivative is identical with the partial derivative at the midpoint

$$\frac{Dg}{D\mathbf{x}}(\mathbf{a}, \mathbf{b}) = \frac{\partial g}{\partial \mathbf{x}}\left(\frac{\mathbf{a} + \mathbf{b}}{2}\right) = \frac{1}{2}(\mathbf{a} + \mathbf{b})^\top \mathbf{G} + \mathbf{c}^\top \quad (5.28)$$

as the second summand in (5.19) vanishes.



For vector functions  $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  the discrete derivative can be assembled component-by-component

$$\frac{D\mathbf{f}}{D\mathbf{x}}(\mathbf{a}, \mathbf{b}) = \begin{pmatrix} \frac{Df_1}{D\mathbf{x}}(\mathbf{a}, \mathbf{b}) \\ \vdots \\ \frac{Df_m}{D\mathbf{x}}(\mathbf{a}, \mathbf{b}) \end{pmatrix} \quad (5.29)$$

yielding a discrete Jacobi matrix.

### 5.2.2 Constant Mass Matrix

In this section, a first consistent integration scheme for mechanical systems with a constant mass matrix is formulated. The equations of motion in differential-algebraic form are given by

$$\begin{aligned} \mathbf{M}\dot{\mathbf{u}} + \left(\frac{\partial V}{\partial \mathbf{q}}\right)^\top - \left(\frac{\partial \mathbf{g}}{\partial \mathbf{q}}\right)^\top \boldsymbol{\lambda} &= 0 \\ -\dot{\mathbf{q}} + \mathbf{u} &= 0 \\ \mathbf{g}(\mathbf{q}) &= 0. \end{aligned} \quad (5.30)$$

The mass matrix  $\mathbf{M} \in \mathbb{R}^{f \times f}$  is assumed to be constant, symmetric and positive definite. This yields inertial forces that consist only of the term  $\mathbf{M}\dot{\mathbf{u}}$ . Beside the inertial forces, there are potential forces described by a potential function  $V : \mathbb{R}^f \rightarrow \mathbb{R}$ . The potential  $V$  is a function of the generalized coordinates  $\mathbf{q} \in \mathbb{R}^f$ . The geometric bilateral constraints  $\mathbf{g}$  considered in the equations of motion (5.30) are assumed to be perfect and scleronomic. As follows from the discussion in Section 3.9, the total energy is preserved in this system. Note that the affine body based DAE formulation of the equations of motion of a rigid body from Section 4.2 fits exactly into the form of (5.30). Within the form of (5.30) one can still formulate a multibody system with potential forces and additional geometric, scleronomic and perfect bilateral constraints. Based on the conserving integration scheme suggested by Gonzalez [32], we create a conserving integrator by direct discretization of the DAE (5.30)

$$\begin{aligned} \mathbf{M}(\mathbf{u}^E - \mathbf{u}^B) + \Delta t \left(\frac{DV}{D\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E)\right)^\top - \left(\frac{D\mathbf{g}}{D\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E)\right)^\top \hat{\boldsymbol{\lambda}} &= 0 \\ \mathbf{q}^B - \mathbf{q}^E + \frac{\Delta t}{2}(\mathbf{u}^B + \mathbf{u}^E) &= 0 \\ \mathbf{g}(\mathbf{q}^E) &= 0 \end{aligned} \quad (5.31)$$

where the partial derivatives from the DAE have been replaced with discrete derivatives (see Section 5.2.1). As in Moreau's midpoint rule, a force time integral  $\hat{\boldsymbol{\lambda}}$  is introduced as discrete variables for the constraint forces  $\boldsymbol{\lambda}$ . This should improve the conditioning of the Jacobian matrix as it occurs in a Newton iteration used for the numerical solution of the discrete equations (5.31). Except for the scaling of the discrete constraint forces with  $\Delta t$ , the integration scheme (5.31) is identical to the one proposed by Betsch in

the beginning of [10, 11] as basis for the discrete null space method. The integrator (5.31) is an implicit integration scheme resulting in a set of equations which is nonlinear in general. The consistent discretization of the bilateral constraints on displacement level can be seen directly from the equations (5.31). The equations guarantee that the bilateral constraints are fulfilled at the end of every time step. To verify that the integrator (5.31) is conserving the total energy, we multiply the first and second equation in (5.31) from the left with  $(\mathbf{q}^E - \mathbf{q}^B)^\top$  and  $(\mathbf{u}^E - \mathbf{u}^B)^\top \mathbf{M}$  respectively and sum up the result. One obtains

$$\begin{aligned} (\mathbf{q}^E - \mathbf{q}^B)^\top \left[ \mathbf{M}(\mathbf{u}^E - \mathbf{u}^B) + \Delta t \left( \frac{DV}{D\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E) \right)^\top - \left( \frac{D\mathbf{g}}{D\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E) \right)^\top \hat{\boldsymbol{\lambda}} \right] \\ + (\mathbf{u}^E - \mathbf{u}^B)^\top \mathbf{M} \left[ \mathbf{q}^B - \mathbf{q}^E + \frac{\Delta t}{2}(\mathbf{u}^B + \mathbf{u}^E) \right] = 0. \end{aligned} \quad (5.32)$$

The sum of the two mixed products  $(\mathbf{q}^E - \mathbf{q}^B)^\top \mathbf{M}(\mathbf{u}^E - \mathbf{u}^B)$  and  $(\mathbf{u}^E - \mathbf{u}^B)^\top \mathbf{M}(\mathbf{q}^B - \mathbf{q}^E)$  disappears, because the mass matrix is symmetric. The discrete derivatives can be replaced with a difference of the corresponding functions by applying the directionality property (5.15), yielding

$$\Delta t \left( V(\mathbf{q}^E) - V(\mathbf{q}^B) \right) + \frac{\Delta t}{2} (\mathbf{u}^E - \mathbf{u}^B)^\top \mathbf{M}(\mathbf{u}^B + \mathbf{u}^E) + \left( \mathbf{g}(\mathbf{q}^B) - \mathbf{g}(\mathbf{q}^E) \right)^\top \hat{\boldsymbol{\lambda}} = 0. \quad (5.33)$$

The bilateral constraints are fulfilled at the end of every time step. If the integration is started with compatible initial conditions, then we can assume

$$\mathbf{g}(\mathbf{q}^B) = 0, \quad \mathbf{g}(\mathbf{q}^E) = 0 \quad (5.34)$$

for all time steps. Canceling also the remaining  $\Delta t$  yields

$$V(\mathbf{q}^E) - V(\mathbf{q}^B) + \frac{1}{2} (\mathbf{u}^E - \mathbf{u}^B)^\top \mathbf{M}(\mathbf{u}^B + \mathbf{u}^E) = 0. \quad (5.35)$$

The quadratic form in (5.35) is just the difference of the kinetic energy between beginning and end of the time step. This means the total energy is preserved

$$V(\mathbf{q}^E) + T(\mathbf{u}^E) = V(\mathbf{q}^B) + T(\mathbf{u}^B) \quad (5.36)$$

and the integrator (5.31) is consistent with respect to the total energy. The geometric bilateral constraints in (5.30) give rise to constraints on velocity level

$$\mathbf{g}(\mathbf{q}) = 0 \quad \Rightarrow \quad \frac{\partial \mathbf{g}}{\partial \mathbf{q}}(\mathbf{q}) \mathbf{u} = 0 \quad (5.37)$$

which are perfectly fulfilled for an exact solution of (5.30). From the constraint equations at the beginning and the end of the time step it follows, that the difference  $\mathbf{q}^E - \mathbf{q}^B$  is orthogonal to the discrete derivative of the constraint

$$\mathbf{g}(\mathbf{q}^B) = 0, \quad \mathbf{g}(\mathbf{q}^E) = 0 \quad \Rightarrow \quad \frac{D\mathbf{g}}{D\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E)(\mathbf{q}^E - \mathbf{q}^B) = 0. \quad (5.38)$$

Replacing the difference  $\mathbf{q}^E - \mathbf{q}^B$  with the help of the second equation from (5.31) yields

$$\frac{D\mathbf{g}}{D\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E)(\mathbf{u}^B + \mathbf{u}^E) = 0. \quad (5.39)$$

Comparing this with the velocity constraint (5.37) shows, that in the discrete approximation the velocity constraints are fulfilled only for the average velocity of a time step. This is illustrated in Figure 5.2. A consistent velocity at the beginning of a time step does in

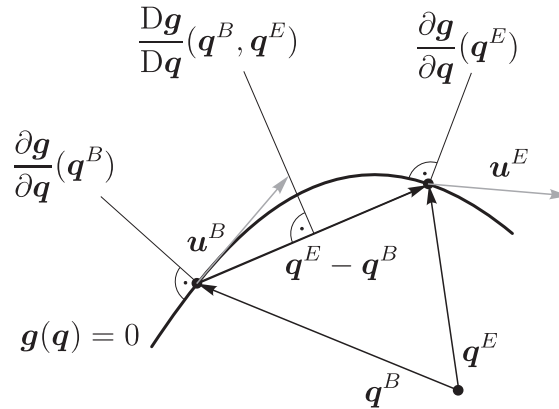


Figure 5.2: Velocities in the discrete approximation.

general not yield a consistent velocity at the end of the step,

$$\frac{\partial \mathbf{g}}{\partial \mathbf{q}}(\mathbf{q}^B)\mathbf{u}^B = 0 \quad \not\Rightarrow \quad \frac{\partial \mathbf{g}}{\partial \mathbf{q}}(\mathbf{q}^E)\mathbf{u}^E = 0, \quad (5.40)$$

as only the average velocity is forced to be consistent. A violation of the velocity constraint at the beginning results in a violation at the end, which is approximately of the same size, but into the opposite direction. In numerical simulations this usually results in velocities oscillating around the correct value with amplitudes that can increase over time. The oscillations of the velocities has been described for a similar integration scheme as well by Gonzalez [33]. One of the sources of this problem is the discretization

$$\mathbf{q}^E - \mathbf{q}^B = \frac{\Delta t}{2}(\mathbf{u}^B + \mathbf{u}^E) \quad (5.41)$$

of the kinematic equation  $\dot{\mathbf{q}} = \mathbf{u}$  used in the scheme (5.31). This discretization is too restrictive and imposes constraints on  $\mathbf{q}^E$  and  $\mathbf{u}^E$  which might be incompatible with the bilateral constraints. For example in a situation as depicted in Figure 5.3, enforcing the difference  $\mathbf{q}^E - \mathbf{q}^B$  to be exactly parallel to the velocities is not compatible with the bilateral constraints. The problem of inconsistent velocities could be solved by replacing the geometric bilateral constraints with kinematic bilateral constraints. This would yield automatically consistent velocities. Unfortunately, such an approach results in drift of the constraints on displacement level, which again can't be repaired, without changing the discretization (5.41). Note that removing the drift problem with a projection to

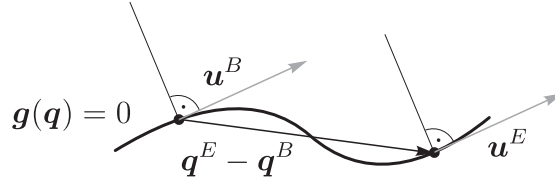


Figure 5.3: Velocities and position differences.

valid generalized coordinates is difficult when energy and other consistency conditions should not be broken. For example, given the drifted generalized coordinates  $\mathbf{q}^E$  which do not fulfill the constraint, i.e.  $\mathbf{g}(\mathbf{q}^E) \neq 0$ , one would have to find a corrected generalized coordinate  $\mathbf{q}^C$  that fulfills the constraint equation while the value of the potential functions remains unchanged, i.e.

$$\mathbf{g}(\mathbf{q}^C) = 0, \quad V(\mathbf{q}^C) = V(\mathbf{q}^E). \quad (5.42)$$

Depending on the situation there might be no solution to (5.42). Imagine for example a mass point in  $\mathbb{R}^3$  which is constrained to the surface of a sphere and which is subject to a gravity potential force. If the mass point has drifted below the lowest (with respect to the gravity potential) point of the sphere, then the problem (5.42) can't be solved, because any point on the sphere has a larger potential energy. In this case also a modification of the velocities would be required in order to preserve the total energy.

The constraint consistency problem could be solved by adding the geometric and the induced kinematic bilateral constraints to the DAE (5.30), while inserting additional multipliers into the kinematic equation as suggested by Gear et al. [22]. Inspired by the method of Gear et al., Betsch and Steinmann developed a variant to this method in [14] for a Hamiltonian based formulation. The method by Betsch and Steinmann allows not only to enforce consistent constraints on both levels, but also preserves the total energy. In the following we will apply this method to the DAE (5.30). In contrast to [14] we will still use generalized velocities instead of generalized momenta. First we introduce a function

$$\boldsymbol{\gamma}(\mathbf{u}, \mathbf{q}) = \frac{\partial \mathbf{g}}{\partial \mathbf{q}} \mathbf{u} \quad (5.43)$$

to express the induced kinematic bilateral constraint as  $\boldsymbol{\gamma}(\mathbf{u}, \mathbf{q}) = 0$ . Then the DAE (5.30) is extended with the kinematic constraint and an additional multiplier  $\boldsymbol{\kappa}$  in the kinematic equation and the equations of motion

$$\begin{aligned} \mathbf{M}\dot{\mathbf{u}} + \left(\frac{\partial V}{\partial \mathbf{q}}\right)^\top - \left(\frac{\partial \mathbf{g}}{\partial \mathbf{q}}\right)^\top \boldsymbol{\lambda} - \left(\frac{\partial \boldsymbol{\gamma}}{\partial \mathbf{q}}\right)^\top \boldsymbol{\kappa} &= 0 \\ -\dot{\mathbf{q}} + \mathbf{u} - \mathbf{M}^{-1} \left(\frac{\partial \boldsymbol{\gamma}}{\partial \mathbf{u}}\right)^\top \boldsymbol{\kappa} &= 0 \\ \mathbf{g}(\mathbf{q}) &= 0 \\ \boldsymbol{\gamma}(\mathbf{u}, \mathbf{q}) &= 0. \end{aligned} \quad (5.44)$$

Adding the induced kinematic constraint equation to the set of equations does not change the solution, as it is contained in the geometric constraint equation already. The extended

DAE (5.44) is identical to the original DAE (5.30) if the multiplier  $\boldsymbol{\kappa}$  is always zero. To verify that this is always the case, we note that a solution of (5.44) fulfills the geometric constraint equation  $\mathbf{g}(\mathbf{q}) = 0$  at all times, from which follows

$$\frac{\partial \mathbf{g}}{\partial \mathbf{q}} \dot{\mathbf{q}} = 0. \quad (5.45)$$

Inserting the second equation of (5.44) into (5.45) yields

$$\frac{\partial \mathbf{g}}{\partial \mathbf{q}} \mathbf{u} - \left( \frac{\partial \mathbf{g}}{\partial \mathbf{q}} \right) \mathbf{M}^{-1} \left( \frac{\partial \boldsymbol{\gamma}}{\partial \mathbf{u}} \right)^\top \boldsymbol{\kappa} = 0 \quad (5.46)$$

As the first term in (5.46) is always zero, one gets

$$\left( \frac{\partial \mathbf{g}}{\partial \mathbf{q}} \right) \mathbf{M}^{-1} \left( \frac{\partial \boldsymbol{\gamma}}{\partial \mathbf{u}} \right)^\top \boldsymbol{\kappa} = 0 \quad (5.47)$$

as equation for the multiplier  $\boldsymbol{\kappa}$ . If the bilateral constraints are linearly independent, then the Delassus matrix in (5.47) is regular and the only solution is

$$\boldsymbol{\kappa} \equiv 0. \quad (5.48)$$

An integration scheme based on the extended DAE (5.44) can be obtained again by direct discretization using the discrete derivative

$$\begin{aligned} \mathbf{M}(\mathbf{u}^E - \mathbf{u}^B) + \Delta t \left( \frac{D\mathbf{V}}{D\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E) \right)^\top - \left( \frac{D\mathbf{g}}{D\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E) \right)^\top \hat{\boldsymbol{\lambda}} - \left( \frac{D\boldsymbol{\gamma}}{D\mathbf{q}}(\mathbf{u}^E, \mathbf{q}^B, \mathbf{q}^E) \right)^\top \hat{\boldsymbol{\kappa}} &= 0 \\ \mathbf{M}(\mathbf{q}^B - \mathbf{q}^E) + \frac{\Delta t}{2} \mathbf{M}(\mathbf{u}^B + \mathbf{u}^E) - \left( \frac{\partial \boldsymbol{\gamma}}{\partial \mathbf{u}}(\mathbf{q}^B) \right)^\top \hat{\boldsymbol{\kappa}} &= 0 \\ \mathbf{g}(\mathbf{q}^E) &= 0 \\ \boldsymbol{\gamma}(\mathbf{u}^E, \mathbf{q}^E) &= 0. \end{aligned} \quad (5.49)$$

Note that the partial derivative  $\partial \boldsymbol{\gamma} / \partial \mathbf{u}$  is identical to the discrete derivative  $D\boldsymbol{\gamma} / D\mathbf{u}$ , because the function  $\boldsymbol{\gamma}$  is linear in  $\mathbf{u}$ . The second equation in (5.49) has been multiplied with the regular mass matrix  $\mathbf{M}$  in order to remove the  $\mathbf{M}^{-1}$ . While there was no difference between the exact solution of the original DAE (5.30) and the extended DAE (5.44), we get a different approximation from the extended integrator (5.49) when compared with the original one (5.31). The extended integrator (5.49) guarantees the consistency of the constraints on displacement level and on velocity level. This is achieved by weakening the original discretization of the kinematic equation (5.41) with additional degrees of freedom introduced by  $\hat{\boldsymbol{\kappa}}$ . In the discretization, the multipliers  $\hat{\boldsymbol{\kappa}}$  should be small, but they are no longer zero in general. At this point we can see why an additional  $\boldsymbol{\kappa}$ -term has been added not only to the kinematic equation but also to the first equation in (5.44). Without it there would be no chance of maintaining the energy consistency, as the additional  $\boldsymbol{\kappa}$ -term

in the kinematic equation alone would destroy the total energy balance. For the difference in total energy between end and beginning of a time step we introduce the abbreviation

$$\Delta E := E(\mathbf{u}^E, \mathbf{q}^E) - E(\mathbf{u}^B, \mathbf{q}^B) = V(\mathbf{q}^E) - V(\mathbf{q}^B) + T(\mathbf{u}^E) - T(\mathbf{u}^B). \quad (5.50)$$

To verify that the integrator (5.49) is consistent with respect to the total energy, we multiply the first and second equation in (5.49) from the left with  $(\mathbf{q}^E - \mathbf{q}^B)^\top$  and  $(\mathbf{u}^E - \mathbf{u}^B)^\top$  respectively and sum up the result. One gets the differences of the potential and the kinetic energy as above plus the terms in  $\hat{\boldsymbol{\kappa}}$ , i.e.

$$\begin{aligned} \Delta t \Delta E &= (\mathbf{q}^E - \mathbf{q}^B)^\top \left( \frac{D\boldsymbol{\gamma}}{D\mathbf{q}}(\mathbf{u}^E, \mathbf{q}^B, \mathbf{q}^E) \right)^\top \hat{\boldsymbol{\kappa}} + (\mathbf{u}^E - \mathbf{u}^B)^\top \left( \frac{\partial \boldsymbol{\gamma}}{\partial \mathbf{u}}(\mathbf{q}^B) \right)^\top \hat{\boldsymbol{\kappa}} \\ &= (\boldsymbol{\gamma}(\mathbf{u}^E, \mathbf{q}^E) - \boldsymbol{\gamma}(\mathbf{u}^E, \mathbf{q}^B) + \boldsymbol{\gamma}(\mathbf{u}^E, \mathbf{q}^B) - \boldsymbol{\gamma}(\mathbf{u}^B, \mathbf{q}^B))^\top \hat{\boldsymbol{\kappa}} \\ &= 0 \end{aligned} \quad (5.51)$$

In the last step of (5.51) the equations  $\boldsymbol{\gamma}(\mathbf{u}^E, \mathbf{q}^E) = 0$  and  $\boldsymbol{\gamma}(\mathbf{u}^B, \mathbf{q}^B) = 0$  have been used by assuming that the integration is started with compatible initial conditions. As a next step one could add perfect unilateral contacts into this formulation, but we will omit that as there is still another problem: The formulation is often very inefficient in terms of the number of unknowns and number of equations. The constant mass matrix assumption does not allow to use a set of minimal coordinates in most models. To simulate one rigid body one needs to use the affine body based DAE formulation from Section 4.2. This results in  $\mathbf{q} \in \mathbb{R}^{12}$ ,  $\mathbf{u} \in \mathbb{R}^{12}$ ,  $\boldsymbol{\lambda} \in \mathbb{R}^6$  and  $\boldsymbol{\kappa} \in \mathbb{R}^6$  which makes a total of 36 unknowns and equations just for one rigid body. In some situations the much simpler mass point model might be sufficient to represent inertial forces, but in general the lack of an efficient rigid body model is a problem. The number of unknowns could be reduced by using the discrete null space method developed by Betsch et al. in [10, 11, 52, 53]. Unfortunately, the bilateral constraints are enforced only on displacement level when applying the discrete null space method, which cause difficulties especially when considering unilateral constraints with impacts.

### 5.2.3 Coordinate Dependent Mass Matrix

In this section, an attempt at solving the problems from the previous section is discussed. The aim is to reduce the number of unknowns and equations in a consistent integrator while keeping bilateral constraints enforced on displacement and velocity level. The basis for this section is the equations of motion using a coordinate dependent mass matrix

$$\mathbf{M}(\mathbf{q})\dot{\mathbf{u}} + \mathbf{G}(\mathbf{u}, \mathbf{q})\mathbf{u} + \left( \frac{\partial T}{\partial \mathbf{q}} \right)^\top - \mathbf{f}_q = 0, \quad \dot{\mathbf{q}} = \mathbf{u} \quad (5.52)$$

which has been described in Section 3.4. The kinetic energy has the form

$$T(\mathbf{u}, \mathbf{q}) = \frac{1}{2} \mathbf{u}^\top \mathbf{M}(\mathbf{q}) \mathbf{u}. \quad (5.53)$$

The skew symmetric gyro matrix is given by

$$G_{ij}(\mathbf{u}, \mathbf{q}) = \left( \frac{\partial M_{ik}}{\partial q_j} - \frac{\partial M_{jk}}{\partial q_i} \right) u_k. \quad (5.54)$$

With a coordinate dependent mass matrix, many perfect bilateral constraints can be eliminated by using an appropriate set of generalized coordinates. A rigid body for example can be described with seven generalized coordinates and one additional bilateral constraint using a DAE formulation based on the scalable body (see Section 4.3). For a constant mass matrix formulation of a rigid body (as required in Section 5.2.2) one was restricted to the affine body based DAE formulation using twelve generalized coordinates and six bilateral constraints. Beside the inertial forces already present in (5.52) we consider potential forces and a set of geometric, scleronomic and perfect bilateral constraints. This yields the following DAE formulation:

$$\begin{aligned} M(\mathbf{q})\dot{\mathbf{u}} + \mathbf{G}(\mathbf{u}, \mathbf{q})\dot{\mathbf{q}} + \left( \frac{\partial T}{\partial \mathbf{q}} \right)^\top + \left( \frac{\partial V}{\partial \mathbf{q}} \right)^\top - \left( \frac{\partial \mathbf{g}}{\partial \mathbf{q}} \right)^\top \boldsymbol{\lambda} &= 0 \\ -\dot{\mathbf{q}} + \mathbf{u} &= 0 \\ \mathbf{g}(\mathbf{q}) &= 0. \end{aligned} \quad (5.55)$$

As in Section 5.2.2, the DAE is extended with additional multipliers  $\boldsymbol{\kappa}$  and the kinematic bilateral constraints

$$\boldsymbol{\gamma}(\mathbf{u}, \mathbf{q}) = \frac{\partial \mathbf{g}}{\partial \mathbf{q}} \mathbf{u} = 0 \quad (5.56)$$

induced by the geometric bilateral constraints. One gets the extended DAE

$$\begin{aligned} M(\mathbf{q})\dot{\mathbf{u}} + \mathbf{G}(\mathbf{u}, \mathbf{q})\dot{\mathbf{q}} + \left( \frac{\partial T}{\partial \mathbf{q}} \right)^\top + \left( \frac{\partial V}{\partial \mathbf{q}} \right)^\top - \left( \frac{\partial \mathbf{g}}{\partial \mathbf{q}} \right)^\top \boldsymbol{\lambda} - \left( \frac{\partial \boldsymbol{\gamma}}{\partial \mathbf{q}} \right)^\top \boldsymbol{\kappa} &= 0 \\ -\dot{\mathbf{q}} + \mathbf{u} - M^{-1}(\mathbf{q}) \left( \frac{\partial \boldsymbol{\gamma}}{\partial \mathbf{u}} \right)^\top \boldsymbol{\kappa} &= 0 \\ \mathbf{g}(\mathbf{q}) &= 0 \\ \boldsymbol{\gamma}(\mathbf{u}, \mathbf{q}) &= 0. \end{aligned} \quad (5.57)$$

which is equivalent to (5.55) by the same argument as discussed in Section 5.2.2. A consistent integrator for the equations of motion (5.57) can be constructed again by direct discretization. We replace all time derivatives with a difference and all the partial derivatives with the discrete derivative. The discretization of the  $\boldsymbol{\kappa}$  terms are chosen such that the integrator is consistent with respect to the total energy. The consistent integrator

for (5.57) is given by

$$\begin{aligned}
& \mathbf{M}(\mathbf{q}^B)(\mathbf{u}^E - \mathbf{u}^B) + \mathbf{G}(\mathbf{u}^B, \mathbf{q}^B)(\mathbf{q}^E - \mathbf{q}^B) + \Delta t \left( \frac{DT}{D\mathbf{q}}(\mathbf{u}^E, \mathbf{q}^B, \mathbf{q}^E) \right)^\top \\
& + \Delta t \left( \frac{DV}{D\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E) \right)^\top - \left( \frac{D\mathbf{g}}{D\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E) \right)^\top \hat{\boldsymbol{\lambda}} - \left( \frac{D\boldsymbol{\gamma}}{D\mathbf{q}}(\mathbf{u}^E, \mathbf{q}^B, \mathbf{q}^E) \right)^\top \hat{\boldsymbol{\kappa}} = 0 \\
& \mathbf{M}(\mathbf{q}^B)(\mathbf{q}^B - \mathbf{q}^E) + \frac{\Delta t}{2} \mathbf{M}(\mathbf{q}^B)(\mathbf{u}^B + \mathbf{u}^E) - \left( \frac{\partial \boldsymbol{\gamma}}{\partial \mathbf{u}}(\mathbf{q}^B) \right)^\top \hat{\boldsymbol{\kappa}} = 0 \\
& \mathbf{g}(\mathbf{q}^E) = 0 \\
& \boldsymbol{\gamma}(\mathbf{u}^E, \mathbf{q}^E) = 0.
\end{aligned} \tag{5.58}$$

When comparing with the integration scheme (5.49) from Section 5.2.2 one can see that the main differences are the additional gyro matrix  $\mathbf{G}$ , the discrete derivative of the kinetic energy and the mass matrix evaluated at the beginning of the time step. Obviously the integrator (5.58) is consistent with respect to the bilateral constraints on displacement level and on velocity level. The consistency with respect to the total energy can be verified by multiplying the first and second equation in (5.58) with  $(\mathbf{q}^E - \mathbf{q}^B)^\top$  and  $(\mathbf{u}^E - \mathbf{u}^B)^\top$  respectively and summing up the products. One gets the following:

$$\begin{aligned}
0 &= (\mathbf{q}^E - \mathbf{q}^B)^\top \mathbf{M}(\mathbf{q}^B)(\mathbf{u}^E - \mathbf{u}^B) + (\mathbf{q}^E - \mathbf{q}^B)^\top \mathbf{G}(\mathbf{u}^B, \mathbf{q}^B)(\mathbf{q}^E - \mathbf{q}^B) \\
& + \Delta t (\mathbf{q}^E - \mathbf{q}^B)^\top \left( \frac{DT}{D\mathbf{q}}(\mathbf{u}^E, \mathbf{q}^B, \mathbf{q}^E) \right)^\top + \Delta t (\mathbf{q}^E - \mathbf{q}^B)^\top \left( \frac{DV}{D\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E) \right)^\top \\
& - (\mathbf{q}^E - \mathbf{q}^B)^\top \left( \frac{D\mathbf{g}}{D\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E) \right)^\top \hat{\boldsymbol{\lambda}} - (\mathbf{q}^E - \mathbf{q}^B)^\top \left( \frac{D\boldsymbol{\gamma}}{D\mathbf{q}}(\mathbf{u}^E, \mathbf{q}^B, \mathbf{q}^E) \right)^\top \hat{\boldsymbol{\kappa}} \\
& + (\mathbf{u}^E - \mathbf{u}^B)^\top \mathbf{M}(\mathbf{q}^B)(\mathbf{q}^B - \mathbf{q}^E) + \frac{\Delta t}{2} (\mathbf{u}^E - \mathbf{u}^B)^\top \mathbf{M}(\mathbf{q}^B)(\mathbf{u}^B + \mathbf{u}^E) \\
& - (\mathbf{u}^E - \mathbf{u}^B)^\top \left( \frac{\partial \boldsymbol{\gamma}}{\partial \mathbf{u}}(\mathbf{q}^B) \right)^\top \hat{\boldsymbol{\kappa}}.
\end{aligned} \tag{5.59}$$

The sum of the products  $(\mathbf{q}^E - \mathbf{q}^B)^\top \mathbf{M}(\mathbf{q}^B)(\mathbf{u}^E - \mathbf{u}^B)$  and  $(\mathbf{u}^E - \mathbf{u}^B)^\top \mathbf{M}(\mathbf{q}^B)(\mathbf{q}^B - \mathbf{q}^E)$  is zero, because the mass matrix  $\mathbf{M}$  is symmetric. The product formed with the gyro matrix  $\mathbf{G}$  is equal to zero as well, because  $\mathbf{G}$  is skew symmetric. Simplifying the remaining expressions in (5.59) one gets

$$\begin{aligned}
0 &= \Delta t \left( T(\mathbf{u}^E, \mathbf{q}^E) - T(\mathbf{u}^E, \mathbf{q}^B) + V(\mathbf{q}^E) - V(\mathbf{q}^B) + T(\mathbf{u}^E, \mathbf{q}^B) - T(\mathbf{u}^B, \mathbf{q}^B) \right) \\
& - \left( \boldsymbol{\gamma}(\mathbf{u}^E, \mathbf{q}^E) - \boldsymbol{\gamma}(\mathbf{u}^E, \mathbf{q}^B) + \boldsymbol{\gamma}(\mathbf{u}^E, \mathbf{q}^B) - \boldsymbol{\gamma}(\mathbf{u}^B, \mathbf{q}^B) \right)^\top \hat{\boldsymbol{\kappa}} \\
& - \left( \mathbf{g}(\mathbf{q}^B) - \mathbf{g}(\mathbf{q}^E) \right)^\top \hat{\boldsymbol{\lambda}}
\end{aligned} \tag{5.60}$$

The bilateral constraints are fulfilled at the beginning and at the end of the time step

$$\boldsymbol{\gamma}(\mathbf{u}^E, \mathbf{q}^E) = 0, \quad \boldsymbol{\gamma}(\mathbf{u}^B, \mathbf{q}^B) = 0, \quad \mathbf{g}(\mathbf{q}^E) = 0, \quad \mathbf{g}(\mathbf{q}^B) = 0. \tag{5.61}$$



Inserting (5.61) into (5.60) finally yields the consistency

$$T(\mathbf{u}^E, \mathbf{q}^E) + V(\mathbf{q}^E) = T(\mathbf{u}^B, \mathbf{q}^B) + V(\mathbf{q}^B) \quad (5.62)$$

of the integrator (5.58) with respect to the total energy. For rigid body models, the consistent integrator described in this section reduces the number of variables and equations compared to the integrator based on the constant mass matrix. With the coordinate dependent mass matrix, one can use the rigid body DAE based on the scalable body. This results in  $\mathbf{q} \in \mathbb{R}^7$ ,  $\mathbf{u} \in \mathbb{R}^7$ ,  $\boldsymbol{\lambda} \in \mathbb{R}$  and  $\boldsymbol{\kappa} \in \mathbb{R}$  which makes a total of 16 unknowns and equations for one body (compared to 36 with a constant mass matrix approach). The disadvantage of the formulation presented in this section lies in the additional terms for the inertial forces. Besides the gyro matrix there is as well a partial derivative of the kinetic energy. The partial derivative of the kinetic energy can result in complicated expressions already. Constructing a discrete derivative from this is possible, but increases the complexity even further. Already for smaller mechanical systems, formulating the Jacobi matrix required for solving the nonlinear equations of the integrator with a Newton method becomes impracticable in closed form. For this reason we will omit the steps required for adding further force laws like unilateral contacts to this integrator formulation.

### 5.2.4 Decomposed Mass Matrix

In this section, we will formulate a consistent integrator based on the equations of motion using a decomposed mass matrix (see Section 3.5). The kinetic energy of the mechanical system is assumed to have the form

$$T(\dot{\mathbf{q}}, \mathbf{q}) = \frac{1}{2} \dot{\mathbf{q}}^\top \mathbf{Q}^\top(\mathbf{q}) \mathbf{M} \mathbf{Q}(\mathbf{q}) \dot{\mathbf{q}}. \quad (5.63)$$

The mass matrix with respect to  $\dot{\mathbf{q}} \in \mathbb{R}^f$  can be decomposed using a regular matrix  $\mathbf{Q}(\mathbf{q}) \in \mathbb{R}^{f \times f}$ . The constant, symmetric and positive definite matrix  $\mathbf{M} \in \mathbb{R}^{f \times f}$  is the mass matrix with respect to the generalized velocities

$$\mathbf{u} = \mathbf{Q}(\mathbf{q}) \dot{\mathbf{q}}. \quad (5.64)$$

obtained by transforming  $\dot{\mathbf{q}}$  with  $\mathbf{Q}$ . Beside the inertial forces, we consider potential forces described by the potential  $V(\mathbf{q})$  and geometric, scleronomic and perfect bilateral constraints described by  $\mathbf{g}_S(\mathbf{q})$ . With these additional forces, the equations of motion (3.52) then become the following DAE:

$$\begin{aligned} \mathbf{Q}^\top(\mathbf{q}) \mathbf{M} \dot{\mathbf{u}} + \mathbf{H}(\mathbf{u}, \mathbf{q}) \dot{\mathbf{q}} + \left( \frac{\partial V}{\partial \mathbf{q}} \right)^\top - \left( \frac{\partial \mathbf{g}_S}{\partial \mathbf{q}} \right)^\top \boldsymbol{\lambda}_S &= 0 \\ - \mathbf{Q}(\mathbf{q}) \dot{\mathbf{q}} + \mathbf{u} &= 0 \\ \mathbf{g}_S(\mathbf{q}) &= 0. \end{aligned} \quad (5.65)$$

The total energy function for this system is given by

$$E(\mathbf{u}, \mathbf{q}) = \frac{1}{2} \mathbf{u}^\top \mathbf{M} \mathbf{u} + V(\mathbf{q}). \quad (5.66)$$

In contrast to the previous sections, the DAE (5.65) is formulated in terms of  $\mathbf{q}$  and  $\mathbf{u}$ , where  $\dot{\mathbf{q}} \neq \mathbf{u}$  in general, but the equations of motion are still formulated based on projections associated with  $\mathbf{q}$ . Note that the constant mass matrix approach discussed in Section 5.2.2 is contained as special case for  $\mathbf{Q}(\mathbf{q}) = \mathbf{I}$  in the approach discussed in this section. The idea for the formulation (5.65) is based on the work by Lens et al. [51], but here generalized velocities are used instead of generalized momenta. Also the discretization developed in the following will be different from [51]. A direct discretization of (5.65) would also show the constraint velocity consistency problem discussed in Section 5.2.2. To avoid this, the DAE is extended with additional multipliers  $\boldsymbol{\kappa}_S$  and the kinematic bilateral constraints

$$\boldsymbol{\gamma}_S(\mathbf{u}, \mathbf{q}) = \frac{\partial \mathbf{g}_S}{\partial \mathbf{q}} \mathbf{Q}^{-1}(\mathbf{q}) \mathbf{u} = 0 \quad (5.67)$$

induced by the geometric bilateral constraints. The extension is done by the same approach as in Section 5.2.2. The extended DAE has the form

$$\begin{aligned} \mathbf{Q}^\top(\mathbf{q}) \mathbf{M} \dot{\mathbf{u}} + \mathbf{H}(\mathbf{u}, \mathbf{q}) \dot{\mathbf{q}} + \left( \frac{\partial V}{\partial \mathbf{q}} \right)^\top - \left( \frac{\partial \mathbf{g}_S}{\partial \mathbf{q}} \right)^\top \boldsymbol{\lambda}_S - \left( \frac{\partial \boldsymbol{\gamma}_S}{\partial \mathbf{q}} \right)^\top \boldsymbol{\kappa}_S &= 0 \\ - \mathbf{Q}(\mathbf{q}) \dot{\mathbf{q}} + \mathbf{u} - \mathbf{M}^{-1} \left( \frac{\partial \boldsymbol{\gamma}_S}{\partial \mathbf{u}} \right)^\top \boldsymbol{\kappa}_S &= 0 \\ \mathbf{g}_S(\mathbf{q}) &= 0 \\ \boldsymbol{\gamma}_S(\mathbf{u}, \mathbf{q}) &= 0. \end{aligned} \quad (5.68)$$

The extended DAE (5.68) is identical to the original DAE (5.65) for  $\boldsymbol{\kappa}_S = 0$ , because the additional induced kinematic constraint equation is already contained in the geometric constraint equation. From  $\mathbf{g}_S(\mathbf{q}) = 0$  at all times it follows

$$\frac{\partial \mathbf{g}_S}{\partial \mathbf{q}} \dot{\mathbf{q}} = 0. \quad (5.69)$$

Solving the second equation of (5.68) for  $\dot{\mathbf{q}}$  and inserting it into (5.69) yields

$$\begin{aligned} \frac{\partial \mathbf{g}_S}{\partial \mathbf{q}} \mathbf{Q}^{-1}(\mathbf{q}) \mathbf{u} - \frac{\partial \mathbf{g}_S}{\partial \mathbf{q}} \mathbf{Q}^{-1}(\mathbf{q}) \mathbf{M}^{-1} \left( \frac{\partial \boldsymbol{\gamma}_S}{\partial \mathbf{u}} \right)^\top \boldsymbol{\kappa}_S &= 0. \\ \underbrace{\frac{\partial \mathbf{g}_S}{\partial \mathbf{q}} \mathbf{Q}^{-1}(\mathbf{q}) \mathbf{u}}_{= \boldsymbol{\gamma}_S(\mathbf{u}, \mathbf{q})} &= 0 \end{aligned} \quad (5.70)$$

This can be simplified to the equation

$$\left( \frac{\partial \boldsymbol{\gamma}_S}{\partial \mathbf{u}} \right) \mathbf{M}^{-1} \left( \frac{\partial \boldsymbol{\gamma}_S}{\partial \mathbf{u}} \right)^\top \boldsymbol{\kappa}_S = 0. \quad (5.71)$$

If the constraints are linearly independent, then the Delassus matrix in (5.71) is regular and  $\boldsymbol{\kappa}_S \equiv 0$  is the only solution. From this it follows that the extended DAE (5.68) has the same solution as the original DAE (5.65). A consistent integrator for the extended

DAE (5.68) can be created by direct discretization

$$\begin{aligned}
& \mathbf{Q}^\top(\mathbf{q}^B)\mathbf{M}(\mathbf{u}^E - \mathbf{u}^B) + \mathbf{H}(\mathbf{u}^B, \mathbf{q}^B)(\mathbf{q}^E - \mathbf{q}^B) + \Delta t \left( \frac{DV}{D\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E) \right)^\top \\
& - \left( \frac{D\mathbf{g}_S}{D\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E) \right)^\top \hat{\boldsymbol{\lambda}}_S - \left( \frac{D\gamma_S}{D\mathbf{q}}(\mathbf{u}^E, \mathbf{q}^B, \mathbf{q}^E) \right)^\top \hat{\boldsymbol{\kappa}}_S = 0 \\
& \mathbf{M}\mathbf{Q}(\mathbf{q}^B)(\mathbf{q}^B - \mathbf{q}^E) + \frac{\Delta t}{2}\mathbf{M}(\mathbf{u}^B + \mathbf{u}^E) - \left( \frac{\partial\gamma_S}{\partial\mathbf{u}}(\mathbf{q}^B) \right)^\top \hat{\boldsymbol{\kappa}}_S = 0 \\
& \mathbf{g}_S(\mathbf{q}^E) = 0 \\
& \gamma_S(\mathbf{u}^E, \mathbf{q}^E) = 0
\end{aligned} \tag{5.72}$$

where the partial derivatives have been replaced with discrete derivatives and the time derivatives with differences. The integrator (5.72) is consistent with respect to the bilateral constraints on displacement and velocity level. This follows directly from the last two equations in (5.72). The total energy balance of the integrator is obtained by multiplying the first and second equation in (5.72) from the left with  $(\mathbf{q}^E - \mathbf{q}^B)^\top$  and  $(\mathbf{u}^E - \mathbf{u}^B)^\top$  respectively and summing up the result. The product formed with the gyro matrix

$$(\mathbf{q}^E - \mathbf{q}^B)^\top \mathbf{H}(\mathbf{u}^B, \mathbf{q}^B)(\mathbf{q}^E - \mathbf{q}^B) = 0 \tag{5.73}$$

is equal to zero, because  $\mathbf{H}$  is skew symmetric. The resulting sum

$$(\mathbf{q}^E - \mathbf{q}^B)^\top \mathbf{Q}^\top(\mathbf{q}^B)\mathbf{M}(\mathbf{u}^E - \mathbf{u}^B) + (\mathbf{u}^E - \mathbf{u}^B)^\top \mathbf{M}\mathbf{Q}(\mathbf{q}^B)(\mathbf{q}^B - \mathbf{q}^E) = 0 \tag{5.74}$$

vanishes, because the mass matrix  $\mathbf{M}$  is symmetric. The remaining terms are

$$\begin{aligned}
0 = & \Delta t \left( V(\mathbf{q}^E) - V(\mathbf{q}^B) + \frac{1}{2}(\mathbf{u}^E - \mathbf{u}^B)^\top \mathbf{M}(\mathbf{u}^E - \mathbf{u}^B) \right) \\
& - \left( \gamma_S(\mathbf{u}^E, \mathbf{q}^E) - \gamma_S(\mathbf{u}^E, \mathbf{q}^B) + \gamma_S(\mathbf{u}^E, \mathbf{q}^B) - \gamma_S(\mathbf{u}^B, \mathbf{q}^B) \right)^\top \hat{\boldsymbol{\kappa}}_S \\
& - \left( \mathbf{g}_S(\mathbf{q}^E) - \mathbf{g}_S(\mathbf{q}^B) \right)^\top \hat{\boldsymbol{\lambda}}_S
\end{aligned} \tag{5.75}$$

With consistent initial conditions, the following is true for every time step:

$$\gamma_S(\mathbf{u}^E, \mathbf{q}^E) = \gamma_S(\mathbf{u}^B, \mathbf{q}^B) = 0, \quad \mathbf{g}_S(\mathbf{q}^E) = \mathbf{g}_S(\mathbf{q}^B) = 0. \tag{5.76}$$

Inserting this into (5.75), canceling the remaining  $\Delta t$  and using the total energy function (5.66) yields

$$E(\mathbf{u}^E, \mathbf{q}^E) = E(\mathbf{u}^B, \mathbf{q}^B) \tag{5.77}$$

which is exactly the energy consistency condition for the integrator (5.72). The constant mass matrix based integrator from Section 5.2.2 required in general the affine body based rigid body formulation for three-dimensional multibody system. The integrator (5.72) allows the usage of the scalable body formulation from Section 4.3 as well, which helps reducing the number of variables and equations. At the same time the complexity of the integrator from Section 5.2.3 is avoided.

### Rigid Body

In the following, the gyro matrix and the constraint functions used in the integrator (5.72) will be evaluated for the rigid body formulation from Section 4.3, which is based on the scalable body. As described in Section 4.3.3, the generalized coordinates and generalized velocities of the body are given by

$$\mathbf{q}_i = \begin{pmatrix} \mathbf{r}_i \\ a_{0i} \\ \mathbf{a}_i \end{pmatrix}, \quad \mathbf{u}_i = \begin{pmatrix} \mathbf{v}_i \\ \nu_i \\ \boldsymbol{\omega}_i \end{pmatrix}. \quad (5.78)$$

The regular matrix

$$\mathbf{Q}_{ii}(\mathbf{q}) = \begin{pmatrix} \mathbf{I} & 0 \\ 0 & 2\boldsymbol{\varphi}(A_i^*) \end{pmatrix} = \begin{pmatrix} \mathbf{I} & 0 & 0 \\ 0 & 2a_{0i} & 2\mathbf{a}_i^\top \\ 0 & -2\mathbf{a}_i & 2a_{0i}\mathbf{I} - 2\tilde{\mathbf{a}}_i \end{pmatrix} \quad (5.79)$$

describes the transformation between  $\dot{\mathbf{q}}_i$  and  $\mathbf{u}_i$  as defined in (4.72). There is no direct connection to any other generalized velocities, i.e.

$$\mathbf{Q}_{ij}(\mathbf{q}) = 0, \quad \mathbf{Q}_{ji}(\mathbf{q}) = 0, \quad j \neq i. \quad (5.80)$$

The regular mass matrix of the scalable body

$$\mathbf{M}_{ii} = \begin{pmatrix} m_i\mathbf{I} & 0 & 0 \\ 0 & \frac{1}{2}\text{Tr}\boldsymbol{\Theta}_i & 0 \\ 0 & 0 & \boldsymbol{\Theta}_i \end{pmatrix}, \quad \mathbf{M}_{ij} = 0, \quad j \neq i \quad (5.81)$$

has been given in (4.88). We will use the formula (3.49) to evaluate the gyro matrix  $\mathbf{H}_{ii}(\mathbf{u}, \mathbf{q})$  of the body. This requires the function  $\mathbf{p}_i(\mathbf{u}, \mathbf{q})$  defined in (3.48), which can be simplified as follows

$$\begin{aligned} \mathbf{p}_i(\mathbf{u}, \mathbf{q}) &= \mathbf{Q}_{ii}(\mathbf{q})^\top \mathbf{M}_{ii} \mathbf{u}_i \\ &= \begin{pmatrix} \mathbf{I} & 0 \\ 0 & 2\boldsymbol{\varphi}(A_i) \end{pmatrix} \begin{pmatrix} m_i \mathbf{v}_i \\ \psi\left(\left(\frac{1}{2}\text{Tr}\boldsymbol{\Theta}_i \nu_i, \boldsymbol{\Theta}_i \boldsymbol{\omega}_i\right)\right) \end{pmatrix} \\ &= \begin{pmatrix} m_i \mathbf{v}_i \\ \psi\left(A_i\left(\text{Tr}\boldsymbol{\Theta}_i \nu_i, 2\boldsymbol{\Theta}_i \boldsymbol{\omega}_i\right)\right) \end{pmatrix} \\ &= \begin{pmatrix} m_i \mathbf{v}_i \\ \mathbf{T} \boldsymbol{\varphi}\left(\left(\text{Tr}\boldsymbol{\Theta}_i \nu_i, -2\boldsymbol{\Theta}_i \boldsymbol{\omega}_i\right)\right) \mathbf{T} \psi(A_i) \end{pmatrix}. \end{aligned} \quad (5.82)$$

Inserting the partial derivative

$$\frac{\partial \mathbf{p}_i}{\partial \mathbf{q}_i} = \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{T} \boldsymbol{\varphi}\left(\left(\text{Tr}\boldsymbol{\Theta}_i \nu_i, -2\boldsymbol{\Theta}_i \boldsymbol{\omega}_i\right)\right) \mathbf{T} \end{pmatrix}, \quad \frac{\partial \mathbf{p}_i}{\partial \mathbf{q}_j} = 0, \quad j \neq i \quad (5.83)$$

into the formula (3.49) yields

$$\mathbf{H}_{ii}(\mathbf{u}, \mathbf{q}) = \left(\frac{\partial \mathbf{p}_i}{\partial \mathbf{q}_i}\right) - \left(\frac{\partial \mathbf{p}_i}{\partial \mathbf{q}_i}\right)^\top = \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{T} \boldsymbol{\varphi}\left(\left(0, -4\boldsymbol{\Theta}_i \boldsymbol{\omega}_i\right)\right) \mathbf{T} \end{pmatrix}. \quad (5.84)$$

After evaluating the quaternion matrix function we get the skew-symmetric gyro matrix

$$\mathbf{H}_{ii}(\mathbf{u}, \mathbf{q}) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -4\boldsymbol{\omega}_i^\top \boldsymbol{\Theta}_i \\ 0 & 4\boldsymbol{\Theta}_i \boldsymbol{\omega}_i & -4(\boldsymbol{\Theta}_i \boldsymbol{\omega}_i)^\sim \end{pmatrix}, \quad \mathbf{H}_{ij}(\mathbf{u}, \mathbf{q}) = 0, \quad j \neq i \quad (5.85)$$

for the scalable body. The constraint

$$g_{Si}(\mathbf{q}) = a_{0i}^2 + \mathbf{a}_i^\top \mathbf{a}_i - 1, \quad (5.86)$$

reduces the scalable body to a rigid body. For the induced kinematic constraint function as defined by (5.67) one gets

$$\gamma_{Si}(\mathbf{u}, \mathbf{q}) = \nu_i. \quad (5.87)$$

The partial derivatives of the geometric constraint function

$$\frac{\partial g_{Si}}{\partial \mathbf{q}_i} = (0 \quad 2a_{0i} \quad 2\mathbf{a}_i), \quad \frac{\partial g_{Si}}{\partial \mathbf{q}_j} = 0, \quad j \neq i, \quad \frac{Dg_{Si}}{D\mathbf{q}}(\mathbf{a}, \mathbf{b}) = \frac{\partial g_{Si}}{\partial \mathbf{q}} \left( \frac{\mathbf{a} + \mathbf{b}}{2} \right) \quad (5.88)$$

and the kinematic constraint function

$$\frac{D\gamma_{Si}}{D\mathbf{q}} = 0, \quad \frac{\partial \gamma_{Si}}{\partial \mathbf{u}_i} = (0 \quad 1 \quad 0), \quad \frac{\partial \gamma_{Si}}{\partial \mathbf{u}_j} = 0, \quad j \neq i \quad (5.89)$$

complete the collection of expressions required for using rigid body models with the integrator (5.72). Beside three-dimensional rigid bodies, also point masses and two-dimensional rigid bodies can be considered as primitives in the decomposed mass matrix based formulation. Also for certain minimal coordinates formulations an analytical decomposition of the mass matrix can be found.

### Unilateral Constraints

As a next step we add a collection of geometric, scleronomic and perfect unilateral constraints to the equations of motion (5.68). The force law for one unilateral contact is given by

$$g_{Ni}(\mathbf{q}) \in \mathcal{N}_{\mathbb{R}_0^-}(-\lambda_{Ni}), \quad \mathbf{f}_q = \left( \frac{\partial g_{Ni}}{\partial \mathbf{q}} \right)^\top \lambda_{Ni}. \quad (5.90)$$

The vector

$$\mathbf{w}_{Ni}(\mathbf{q}) := \left( \frac{\partial g_{Ni}}{\partial \mathbf{q}} \right)^\top \quad (5.91)$$

is introduced for the generalized force direction of a contact. All generalized force direction of the unilateral contacts are collected in the matrix

$$\mathbf{W}_N(\mathbf{q}) := \left( \frac{\partial \mathbf{g}_N}{\partial \mathbf{q}} \right)^\top. \quad (5.92)$$

Here the generalized force directions based on variations of  $\mathbf{q}$  are used. Extending the equations of motion (5.68) with the force laws of the unilateral contacts on velocity level yields the following set of equations and inclusions:

$$\begin{aligned}
\mathbf{Q}^\top(\mathbf{q})\mathbf{M}\dot{\mathbf{u}} + \mathbf{H}(\mathbf{u}, \mathbf{q})\dot{\mathbf{q}} + \left(\frac{\partial V}{\partial \mathbf{q}}\right)^\top - \left(\frac{\partial \mathbf{g}_S}{\partial \mathbf{q}}\right)^\top \boldsymbol{\lambda}_S - \left(\frac{\partial \gamma_S}{\partial \mathbf{q}}\right)^\top \boldsymbol{\kappa}_S - \mathbf{W}_N(\mathbf{q})\boldsymbol{\lambda}_N &= 0 \\
-\mathbf{Q}(\mathbf{q})\dot{\mathbf{q}} + \mathbf{u} - \mathbf{M}^{-1} \left(\frac{\partial \gamma_S}{\partial \mathbf{u}}\right)^\top \boldsymbol{\kappa}_S &= 0 \\
\mathbf{g}_S(\mathbf{q}) &= 0 \\
\gamma_S(\mathbf{u}, \mathbf{q}) &= 0 \\
g_{Ni}(\mathbf{q}) > 0 : \lambda_{Ni} &= 0 \\
g_{Ni}(\mathbf{q}) = 0 : \mathbf{w}_{Ni}^\top(\mathbf{q})\dot{\mathbf{q}} \in \mathcal{N}_{\mathbb{R}_0^-}(-\lambda_{Ni}). &
\end{aligned} \tag{5.93}$$

Impacts can occur at unilateral contacts, for example when there is a collision. To describe the discontinuity of the velocities at an impact, the equations of motion (5.93) have to be equipped with impact equations and for the unilateral contacts an impact law has to be formulated (see Section 3.8). For the contacts a perfectly elastic Newton impact law in inclusion form is used. The impact equations are obtained by integrating (5.93) over a singleton in time. Together with the impact laws we obtain the following inclusion:

$$\begin{aligned}
\mathbf{Q}^\top(\mathbf{q})\mathbf{M}(\mathbf{u}^+ - \mathbf{u}^-) - \left(\frac{\partial \mathbf{g}_S}{\partial \mathbf{q}}\right)^\top \boldsymbol{\Lambda}_S - \mathbf{W}_N(\mathbf{q})\boldsymbol{\Lambda}_N &= 0 \\
\gamma_S(\mathbf{u}^+, \mathbf{q}) &= 0 \\
g_{Ni}(\mathbf{q}) > 0 : \Lambda_{Ni} &= 0 \\
g_{Ni}(\mathbf{q}) = 0 : \mathbf{w}_{Ni}^\top(\mathbf{q})\mathbf{Q}^{-1}(\mathbf{q})(\mathbf{u}^+ + \mathbf{u}^-) \in \mathcal{N}_{\mathbb{R}_0^-}(-\Lambda_{Ni}) &
\end{aligned} \tag{5.94}$$

Note that also the bilateral constraints have to be considered, as they can propagate impact-impulses as well. Since the generalized force directions  $\mathbf{w}_{Ni}(\mathbf{q})$  are formulated based on variations of  $\mathbf{q}$ , the additional transformation  $\mathbf{Q}^{-1}$  from  $\dot{\mathbf{q}}$  to  $\mathbf{u}$  occurs in the impact laws of the unilateral contacts. As integration scheme we use

$$\begin{aligned}
\mathbf{Q}^\top(\mathbf{q}^B)\mathbf{M}(\mathbf{u}^E - \mathbf{u}^B) + \mathbf{H}(\mathbf{u}^B, \mathbf{q}^B)(\mathbf{q}^E - \mathbf{q}^B) + \Delta t \left(\frac{DV}{D\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E)\right)^\top \\
- \left(\frac{D\mathbf{g}_S}{D\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E)\right)^\top \hat{\boldsymbol{\lambda}}_S - \left(\frac{D\gamma_S}{D\mathbf{q}}(\mathbf{u}^E, \mathbf{q}^B, \mathbf{q}^E)\right)^\top \hat{\boldsymbol{\kappa}}_S - \mathbf{W}_N \hat{\boldsymbol{\lambda}}_N &= 0 \\
\mathbf{M}\mathbf{Q}(\mathbf{q}^B)(\mathbf{q}^B - \mathbf{q}^E) + \frac{\Delta t}{2}\mathbf{M}(\mathbf{u}^B + \mathbf{u}^E) - \left(\frac{\partial \gamma_S}{\partial \mathbf{u}}(\mathbf{q}^B)\right)^\top \hat{\boldsymbol{\kappa}}_S &= 0 \\
\mathbf{g}_S(\mathbf{q}^E) &= 0 \\
\gamma_S(\mathbf{u}^E, \mathbf{q}^E) &= 0 \\
g_{Ni}(\mathbf{q}^B) > 0 : \hat{\lambda}_{Ni} &= 0 \\
g_{Ni}(\mathbf{q}^B) \leq 0 : \mathbf{w}_{Ni}^\top(\mathbf{q}^E - \mathbf{q}^B) \in \mathcal{N}_{\mathbb{R}_0^-}(-\hat{\lambda}_{Ni}) &
\end{aligned} \tag{5.95}$$

which approximates both (5.93) and (5.94). For the generalized force directions in the integrator (5.95) one can use either a variant based on the partial derivative

$$\mathbf{w}_N = \left( \frac{\partial \mathbf{g}_N}{\partial \mathbf{q}}(\mathbf{q}^B) \right)^\top \quad (5.96)$$

or one based on the discrete derivative

$$\mathbf{w}_N = \left( \frac{D\mathbf{g}_N}{D\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E) \right)^\top. \quad (5.97)$$

But of course the same approximation should be used in the first and last line of (5.95). The discrete variable  $\hat{\boldsymbol{\lambda}}_N$  approximates the sum of the impact-impulses  $\boldsymbol{\Lambda}_N$  and the integral of the contact forces  $\boldsymbol{\lambda}_N$  over a time step. The total energy balance can be obtained with the same approach as for the integrator (5.72). Already using the simplifications discussed above one gets

$$\Delta t \left( E(\mathbf{u}^E, \mathbf{q}^E) - E(\mathbf{u}^B, \mathbf{q}^B) \right) - \sum_i (\mathbf{q}^E - \mathbf{q}^B)^\top \mathbf{w}_{Ni} \hat{\lambda}_{Ni} = 0. \quad (5.98)$$

The contact forces  $\hat{\lambda}_{Ni}$  are zero for all open contacts ( $g_{Ni}(\mathbf{q}^B) > 0$ ), resulting in a zero contribution in (5.98). For the contacts that are closed at the beginning of a time step, the inclusion in (5.95) or the equivalent complementarity

$$(\mathbf{q}^E - \mathbf{q}^B)^\top \mathbf{w}_{Ni} \hat{\lambda}_{Ni} = 0, \quad \mathbf{w}_{Ni}^\top (\mathbf{q}^E - \mathbf{q}^B) \geq 0, \quad \hat{\lambda}_{Ni} \geq 0 \quad (5.99)$$

holds. It follows that the contribution of the unilateral constraints in (5.98) is zero for all closed contacts. Canceling the  $\Delta t$  in (5.98) yields

$$E(\mathbf{u}^E, \mathbf{q}^E) = E(\mathbf{u}^B, \mathbf{q}^B), \quad (5.100)$$

which shows that the total energy is preserved by the integrator (5.95). This means the integrator (5.95) is consistent with respect to the total energy, because the mechanical system described by (5.93) and (5.94) preserves the total energy as well. The matrix of generalized force directions given in (5.96) is usually simple to implement and is good enough for an energy consistent integrator. The more complicated variant given by (5.97) can be used to achieve a weak form of unilateral constraint consistency. From the last line of the integrator (5.95) we have the relation

$$g_{Ni}(\mathbf{q}^B) \leq 0 : \mathbf{w}_{Ni}^\top (\mathbf{q}^E - \mathbf{q}^B) \geq 0. \quad (5.101)$$

Inserting  $\mathbf{w}_{Ni}$  as given by (5.97) into (5.101) yields

$$g_{Ni}(\mathbf{q}^B) \leq 0 : \frac{Dg_{Ni}}{D\mathbf{q}}(\mathbf{q}^E, \mathbf{q}^B) (\mathbf{q}^E - \mathbf{q}^B) \geq 0 \quad (5.102)$$

This can be simplified to

$$g_{Ni}(\mathbf{q}^B) \leq 0 : g_{Ni}(\mathbf{q}^E) \geq g_{Ni}(\mathbf{q}^B) \quad (5.103)$$

which is exactly the consistency condition for unilateral contacts on velocity level discussed at the beginning of Section 5.2.

The integrator (5.95) has been formulated for perfectly elastic impacts only. While adding coefficients of restitution to the impact inclusions (5.94) is trivial, finding a one-step discretization similar to (5.95) with consistent dissipation seems to be difficult, even if the consistency condition for unilateral contacts on velocity level would be dropped. The difficulties arise mainly from the additional multipliers  $\boldsymbol{\kappa}_S$  and the combined approximation of impacts and impact-free motion in one set of inclusions.

### Normal Cone Inclusions

The restriction of set-valued force laws to perfect unilateral constraints as in (5.95) can be very limiting for applications. To remove this limitation, we extend (5.68) with set-valued force laws of normal cone type (for examples see Section 3.7). We get the following equations of motion:

$$\begin{aligned} \mathbf{Q}^\top(\mathbf{q})\mathbf{M}\dot{\mathbf{u}} + \mathbf{H}(\mathbf{u}, \mathbf{q})\dot{\mathbf{q}} + \left(\frac{\partial V}{\partial \mathbf{q}}\right)^\top - \left(\frac{\partial \mathbf{g}_S}{\partial \mathbf{q}}\right)^\top \boldsymbol{\lambda}_S - \left(\frac{\partial \boldsymbol{\gamma}_S}{\partial \mathbf{q}}\right)^\top \boldsymbol{\kappa}_S - \sum_{i \in \mathcal{I}(\mathbf{q})} \mathbf{W}_i(\mathbf{q}, t)\boldsymbol{\lambda}_i &= 0 \\ -\mathbf{Q}(\mathbf{q})\dot{\mathbf{q}} + \mathbf{u} - \mathbf{M}^{-1} \left(\frac{\partial \boldsymbol{\gamma}_S}{\partial \mathbf{u}}\right)^\top \boldsymbol{\kappa}_S &= 0 \end{aligned} \quad (5.104)$$

$$\mathbf{g}_S(\mathbf{q}) = 0$$

$$\boldsymbol{\gamma}_S(\mathbf{u}, \mathbf{q}) = 0$$

$$i \in \mathcal{I}(\mathbf{q}) : \mathbf{W}_i^\top(\mathbf{q}, t)\dot{\mathbf{q}} \in \mathcal{N}_{\mathcal{C}_i}(-\boldsymbol{\lambda}_i)$$

The convex sets  $\mathcal{C}_i$  are assumed to contain zero, which leads to a (not strictly) dissipative system. The set-valued force laws of normal cone type in (5.104) are formulated on velocity level. The set  $\mathcal{I}(\mathbf{q})$  can be used to consider scleronomic geometric unilateral constraints by keeping only the indices  $i$  in the set  $\mathcal{I}(\mathbf{q})$  where the gap function  $\mathbf{g}_i(\mathbf{q})$  is equal to zero. The equations of motion (5.104) are equipped with impact equations and impact laws

$$\begin{aligned} \mathbf{Q}^\top(\mathbf{q})\mathbf{M}(\mathbf{u}^+ - \mathbf{u}^-) - \left(\frac{\partial \mathbf{g}_S}{\partial \mathbf{q}}\right)^\top \boldsymbol{\Lambda}_S - \sum_{i \in \mathcal{I}(\mathbf{q})} \mathbf{W}_i(\mathbf{q}, t)\boldsymbol{\Lambda}_i &= 0 \\ \boldsymbol{\gamma}_S(\mathbf{u}^+, \mathbf{q}) &= 0 \\ i \in \mathcal{I}(\mathbf{q}) : \mathbf{W}_i^\top(\mathbf{q}, t)\mathbf{Q}^{-1}(\mathbf{q})(\mathbf{u}^+ + \varepsilon_i \mathbf{u}^-) &\in \mathcal{N}_{\mathcal{D}_i}(-\boldsymbol{\Lambda}_i). \end{aligned} \quad (5.105)$$

The impact laws are of Newton type in inclusion form. The coefficients of restitution are denoted by  $\varepsilon_i$ . To get an energy consistent integrator, a two-step scheme is employed. At the beginning of a time step an impact inclusion of the form

$$\begin{aligned} \mathbf{Q}^\top(\mathbf{q}^B)\mathbf{M}(\mathbf{u}^C - \mathbf{u}^B) - \left(\frac{\partial \mathbf{g}_S}{\partial \mathbf{q}}(\mathbf{q}^B)\right)^\top \boldsymbol{\Lambda}_S - \sum_{i \in \mathcal{I}(\mathbf{q}^B)} \mathbf{W}_i(\mathbf{q}^B, t^B)\boldsymbol{\Lambda}_i &= 0 \\ \boldsymbol{\gamma}_S(\mathbf{u}^C, \mathbf{q}^B) &= 0 \\ i \in \mathcal{I}(\mathbf{q}^B) : \mathbf{W}_i^\top(\mathbf{q}^B, t^B)\mathbf{Q}^{-1}(\mathbf{q}^B)(\mathbf{u}^C + \varepsilon_i \mathbf{u}^B) &\in \mathcal{N}_{\mathcal{D}_i}(-\boldsymbol{\Lambda}_i) \end{aligned} \quad (5.106)$$



is solved, where the generalized coordinates are kept constant. Starting with the beginning state  $(\mathbf{q}^B, \mathbf{u}^B)$  this yields the post impact state  $(\mathbf{q}^B, \mathbf{u}^C)$ . From the intermediate point  $(\mathbf{q}^B, \mathbf{u}^C)$ , a time step approximating the impact-free dynamics is made. This time step is described by

$$\begin{aligned}
& \mathbf{Q}^\top(\mathbf{q}^B) \mathbf{M}(\mathbf{u}^E - \mathbf{u}^C) + \mathbf{H}(\mathbf{u}^C, \mathbf{q}^B)(\mathbf{q}^E - \mathbf{q}^B) + \Delta t \left( \frac{\mathrm{D}V}{\mathrm{D}\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E) \right)^\top \\
& - \left( \frac{\mathrm{D}\mathbf{g}_S}{\mathrm{D}\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E) \right)^\top \hat{\boldsymbol{\lambda}}_S - \left( \frac{\mathrm{D}\boldsymbol{\gamma}_S}{\mathrm{D}\mathbf{q}}(\mathbf{u}^E, \mathbf{q}^B, \mathbf{q}^E) \right)^\top \hat{\boldsymbol{\kappa}}_S - \sum_{i \in \mathcal{I}(\mathbf{q}^B)} \mathbf{W}_i(\mathbf{q}^B, t^B) \hat{\boldsymbol{\lambda}}_i = 0 \\
& \mathbf{M}\mathbf{Q}(\mathbf{q}^B)(\mathbf{q}^B - \mathbf{q}^E) + \frac{\Delta t}{2} \mathbf{M}(\mathbf{u}^C + \mathbf{u}^E) - \left( \frac{\partial \boldsymbol{\gamma}_S}{\partial \mathbf{u}}(\mathbf{q}^B) \right)^\top \hat{\boldsymbol{\kappa}}_S = 0 \\
& \mathbf{g}_S(\mathbf{q}^E) = 0 \\
& \boldsymbol{\gamma}_S(\mathbf{u}^E, \mathbf{q}^E) = 0 \\
& i \in \mathcal{I}(\mathbf{q}^B) : \mathbf{W}_i^\top(\mathbf{q}^B, t^B)(\mathbf{q}^E - \mathbf{q}^B) \in \mathcal{N}_{\hat{c}_i}(-\hat{\boldsymbol{\lambda}}_i)
\end{aligned} \tag{5.107}$$

and yields the end time state  $(\mathbf{q}^E, \mathbf{u}^E)$ . A complete time step of the integrator consists of (5.106) and (5.107). The two-step formulation allows to overcome the difficulties of a combined approximation of impacts and impact-free motion in one set of inclusions. The downside of using a two-step scheme is the increased computational effort required to solve two sets of inclusions per time step.

In the following, we will discuss the consistency of the integrator (5.106) and (5.107) with respect to the total energy. As follows from the discussion in Section 3.9, the total energy of the mechanical system described by (5.104) is remaining constant or decreasing for impact-free motion, i.e.

$$\dot{E}(\mathbf{u}, \mathbf{q}) \leq 0. \tag{5.108}$$

For the impacts described by (5.105), we assume that the coefficients of restitution  $\varepsilon_i$  are chosen such that the inequality

$$E(\mathbf{u}^+, \mathbf{q}) \leq E(\mathbf{u}^-, \mathbf{q}) \tag{5.109}$$

holds. This means the total energy is assumed to remain constant or decrease for all impacts. The first step (5.106) of the integrator has exactly the same structure as the impact inclusions (5.105). Since (5.109) is assumed to hold for the impact inclusions (5.105), we get automatically the condition

$$E(\mathbf{u}^C, \mathbf{q}^B) \leq E(\mathbf{u}^B, \mathbf{q}^B) \tag{5.110}$$

for the first integrator step (5.106). To get an energy balance for the second integrator step (5.107), we multiply the first and second equation of (5.107) with  $(\mathbf{q}^E - \mathbf{q}^B)^\top$  and  $(\mathbf{u}^E - \mathbf{u}^C)^\top$  respectively. Using the simplifications discussed above, one gets

$$\Delta t \left( E(\mathbf{u}^E, \mathbf{q}^E) - E(\mathbf{u}^C, \mathbf{q}^B) \right) = \sum_{i \in \mathcal{I}(\mathbf{q}^B)} (\mathbf{q}^E - \mathbf{q}^B)^\top \mathbf{W}_i(\mathbf{q}^B, t^B) \hat{\boldsymbol{\lambda}}_i. \tag{5.111}$$

In the simplifications we have used the guaranteed fulfillment of the bilateral constraints on velocity and displacement level

$$\gamma_S(\mathbf{u}^E, \mathbf{q}^E) = \gamma_S(\mathbf{u}^C, \mathbf{q}^B) = 0, \quad \mathbf{g}_S(\mathbf{q}^E) = \mathbf{g}_S(\mathbf{q}^B) = 0 \quad (5.112)$$

for the states  $(\mathbf{u}^E, \mathbf{q}^E)$  and  $(\mathbf{u}^C, \mathbf{q}^B)$  as given by (5.106) and (5.107). From the normal cone inclusions in (5.107) it follows that all contributions from the  $\hat{\boldsymbol{\lambda}}_i$  in (5.111) are negative

$$\mathbf{W}_i^\top(\mathbf{q}^B, t^B)(\mathbf{q}^E - \mathbf{q}^B) \in \mathcal{N}_{\hat{C}_i}(-\hat{\boldsymbol{\lambda}}_i), \quad 0 \in \hat{C}_i \quad \Rightarrow \quad (\mathbf{q}^E - \mathbf{q}^B)^\top \mathbf{W}_i(\mathbf{q}^B, t^B) \hat{\boldsymbol{\lambda}}_i \leq 0. \quad (5.113)$$

Inserting (5.113) into (5.111) and canceling the remaining  $\Delta t$  yields

$$E(\mathbf{u}^E, \mathbf{q}^E) \leq E(\mathbf{u}^C, \mathbf{q}^B). \quad (5.114)$$

By combining (5.110) and (5.114) one gets the inequality

$$E(\mathbf{u}^E, \mathbf{q}^E) \leq E(\mathbf{u}^B, \mathbf{q}^B) \quad (5.115)$$

for the total energy at the beginning and the end of a time step. The inequality (5.115) is consistent with the properties (5.108) and (5.109) of the exact solution. This means the integrator (5.106) and (5.107) is consistent with respect to the total energy. Considering only unilateral constraints with perfect elastic impact laws yields an equality in (5.108) and (5.109). Of course in this case an equality holds in (5.110), (5.113) and (5.114), from which follows that the total energy is preserved in the integrator. So in this special case, the integrator (5.106) and (5.107) is consistent with respect to the total energy even for the stricter consistency criterion.

The matrices of generalized force directions are approximated with  $\mathbf{W}_i(\mathbf{q}^B, t^B)$  in the second step (5.107) of the integrator. For unilateral constraints, also an approximation

$$\mathbf{W}_N = \left( \frac{D\mathbf{g}_N}{D\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E) \right)^\top \quad (5.116)$$

based on the discrete derivative of the gap function  $\mathbf{g}_N$  can be used. With the discrete derivative based approximation, one gets again the consistency condition for unilateral constraints on velocity level. The consistency is obtained for the second step (5.107) of the integrator with the same arguments as discussed above for the integrator (5.95). In the first step (5.106) of the integrator the generalized coordinates  $\mathbf{q}^B$  remain unchanged, accordingly the consistency condition for geometric unilateral constraints remains unaffected. Note that using the approximation (5.116) instead of  $\mathbf{W}_N(\mathbf{q}^B)$  does not affect the consistency with respect to the total energy.

The problems (5.106) and (5.107) formulated for the two steps of the integrator have to be solved numerically in each time step. The problem (5.106) of the first integrator step can be simplified analytically by eliminating  $\mathbf{u}_C$  and  $\boldsymbol{\Lambda}_S$  first. Solving the first equation

in (5.106) for  $\mathbf{u}_C$  and introducing the variables  $\boldsymbol{\xi}_i$  for the normal cone inclusion yields

$$\begin{aligned} \mathbf{u}^C &= \mathbf{u}^B + M^{-1} \left( \frac{\partial \gamma_S}{\partial \mathbf{u}}(\mathbf{q}^B) \right)^\top \boldsymbol{\Lambda}_S + M^{-1} \mathbf{Q}^{-\top}(\mathbf{q}^B) \sum_{i \in \mathcal{I}(\mathbf{q}^B)} \mathbf{W}_i(\mathbf{q}^B, t^B) \boldsymbol{\Lambda}_i = 0 \\ \frac{\partial \gamma_S}{\partial \mathbf{u}}(\mathbf{q}^B) \mathbf{u}^C &= 0 \\ i \in \mathcal{I}(\mathbf{q}^B) : \boldsymbol{\xi}_i &= \mathbf{W}_i^\top(\mathbf{q}^B, t^B) \mathbf{Q}^{-1}(\mathbf{q}^B) (\mathbf{u}^C + \varepsilon_i \mathbf{u}^B), \quad \boldsymbol{\xi}_i \in \mathcal{N}_{\mathcal{D}_i}(-\boldsymbol{\Lambda}_i). \end{aligned} \quad (5.117)$$

In (5.117) the bilateral constraint has been rewritten to make the linear dependency in  $\mathbf{u}_C$  explicit. After inserting  $\mathbf{u}_C$  from the first equation into the second equation in (5.117) one gets

$$\begin{aligned} \underbrace{\frac{\partial \gamma_S}{\partial \mathbf{u}}(\mathbf{q}^B) M^{-1} \left( \frac{\partial \gamma_S}{\partial \mathbf{u}}(\mathbf{q}^B) \right)^\top \boldsymbol{\Lambda}_S}_{=: \mathbf{G}_{SS}} + \underbrace{\gamma_S(\mathbf{q}^B, \mathbf{u}^B)}_{=0} \\ + \frac{\partial \gamma_S}{\partial \mathbf{u}}(\mathbf{q}^B) M^{-1} \mathbf{Q}^{-\top}(\mathbf{q}^B) \sum_{i \in \mathcal{I}(\mathbf{q}^B)} \mathbf{W}_i(\mathbf{q}^B, t^B) \boldsymbol{\Lambda}_i = 0. \end{aligned} \quad (5.118)$$

This equation can be solved for the bilateral constraint impulses, yielding

$$\boldsymbol{\Lambda}_S = -\mathbf{G}_{SS}^{-1} \frac{\partial \gamma_S}{\partial \mathbf{u}}(\mathbf{q}^B) M^{-1} \mathbf{Q}^{-\top}(\mathbf{q}^B) \sum_{i \in \mathcal{I}(\mathbf{q}^B)} \mathbf{W}_i(\mathbf{q}^B, t^B) \boldsymbol{\Lambda}_i. \quad (5.119)$$

Next, the equation (5.119) is inserted into the first equation of (5.117), which itself is used to replace  $\mathbf{u}_C$  in the definition of  $\boldsymbol{\xi}_i$ . This reduces the problem (5.117) of the first integrator step to the normal cone inclusion with linear equations of the form

$$i \in \mathcal{I}(\mathbf{q}^B) : \boldsymbol{\xi}_i = \sum_{j \in \mathcal{I}(\mathbf{q}^B)} \mathbf{G}_{ij}(\mathbf{q}^B) \boldsymbol{\Lambda}_j + \mathbf{c}_i(\mathbf{q}^B), \quad \boldsymbol{\xi}_i \in \mathcal{N}_{\mathcal{D}_i}(-\boldsymbol{\Lambda}_i) \quad (5.120)$$

where the matrices  $\mathbf{G}_{ij}$  are given by

$$\begin{aligned} \mathbf{G}_{ij} &:= \mathbf{W}_i^\top(\mathbf{q}^B, t^B) \mathbf{Q}^{-1}(\mathbf{q}^B) M^{-1} \mathbf{Q}^{-\top}(\mathbf{q}^B) \mathbf{W}_j(\mathbf{q}^B, t^B) \\ &\quad - \mathbf{W}_i^\top(\mathbf{q}^B, t^B) \mathbf{Q}^{-1}(\mathbf{q}^B) M^{-1} \left( \frac{\partial \gamma_S}{\partial \mathbf{u}}(\mathbf{q}^B) \right)^\top \mathbf{G}_{SS}^{-1} \frac{\partial \gamma_S}{\partial \mathbf{u}}(\mathbf{q}^B) M^{-1} \mathbf{Q}^{-\top}(\mathbf{q}^B) \mathbf{W}_j(\mathbf{q}^B, t^B) \end{aligned} \quad (5.121)$$

and the vectors  $\mathbf{c}_i$  are defined as

$$\mathbf{c}_i := (1 + \varepsilon_i) \mathbf{W}_i^\top(\mathbf{q}^B, t^B) \mathbf{Q}^{-1}(\mathbf{q}^B) \mathbf{u}^B. \quad (5.122)$$

The problem (5.120) can be solved for  $\boldsymbol{\Lambda}_i$  with the projected Jacobi iteration or the projected Gauss-Seidel iteration discussed in Section 2.5. The solution can be substituted back into (5.119) and (5.117) to get the solution for  $\mathbf{u}^C$  required for the second step (5.107) of the integrator. The second step contains in general also nonlinear equations beside the normal cone inclusions. It can be solved with the projected Newton iteration discussed in Section 2.5.5.

### Classical Dissipation

Beside the dissipation based on force laws of normal cone type, one can also consider damping forces of the generalized form

$$\mathbf{f}_q = -\mathbf{D}(\mathbf{u}, \mathbf{q}, t)\dot{\mathbf{q}} \quad (5.123)$$

where  $\mathbf{D}(\mathbf{u}, \mathbf{q}, t) \in \mathbb{R}^{f \times f}$  is a positive semidefinite damping matrix. The contribution to the time derivative of the total energy (see Section 3.9) is then given by

$$\mathbf{f}_q^\top \dot{\mathbf{q}} = -\dot{\mathbf{q}}^\top \mathbf{D}(\mathbf{u}, \mathbf{q}, t)\dot{\mathbf{q}} \leq 0. \quad (5.124)$$

As expected the forces of the type (5.123) are (not strictly) dissipative. To consider these damping forces in the equations of motion, the first equation in (5.104) has to be replaced with

$$\begin{aligned} \mathbf{Q}^\top(\mathbf{q})\mathbf{M}\dot{\mathbf{u}} + \mathbf{H}(\mathbf{u}, \mathbf{q})\dot{\mathbf{q}} + \mathbf{D}(\mathbf{u}, \mathbf{q}, t)\dot{\mathbf{q}} + \left(\frac{\partial V}{\partial \mathbf{q}}\right)^\top \\ - \left(\frac{\partial \mathbf{g}_S}{\partial \mathbf{q}}\right)^\top \boldsymbol{\lambda}_S - \left(\frac{\partial \gamma_S}{\partial \mathbf{q}}\right)^\top \boldsymbol{\kappa}_S - \sum_{i \in \mathcal{I}(\mathbf{q})} \mathbf{W}_i(\mathbf{q})\boldsymbol{\lambda}_i = 0. \end{aligned} \quad (5.125)$$

Since the damping forces (5.123) are never impulsive they have to be considered only in the second step of the integrator by replacing the first equation of (5.107) with

$$\begin{aligned} \mathbf{Q}^\top(\mathbf{q}^B)\mathbf{M}(\mathbf{u}^E - \mathbf{u}^C) + \mathbf{H}(\mathbf{u}^C, \mathbf{q}^B)(\mathbf{q}^E - \mathbf{q}^B) + \mathbf{D}(\mathbf{u}^C, \mathbf{q}^B, t^B)(\mathbf{q}^E - \mathbf{q}^B) \\ + \Delta t \left(\frac{\mathrm{D}V}{\mathrm{D}\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E)\right)^\top - \left(\frac{\mathrm{D}\mathbf{g}_S}{\mathrm{D}\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E)\right)^\top \hat{\boldsymbol{\lambda}}_S - \left(\frac{\mathrm{D}\gamma_S}{\mathrm{D}\mathbf{q}}(\mathbf{u}^E, \mathbf{q}^B, \mathbf{q}^E)\right)^\top \hat{\boldsymbol{\kappa}}_S \\ - \sum_{i \in \mathcal{I}(\mathbf{q}^B)} \mathbf{W}_i \hat{\boldsymbol{\lambda}}_i = 0. \end{aligned} \quad (5.126)$$

Considering the additional force term  $\mathbf{D}(\mathbf{u}^C, \mathbf{q}^B, t^B)(\mathbf{q}^E - \mathbf{q}^B)$  of the second integrator in the energy balance (5.111) yields the following relation:

$$\begin{aligned} \Delta t \left( E(\mathbf{u}^E, \mathbf{q}^E) - E(\mathbf{u}^C, \mathbf{q}^B) \right) = \underbrace{-(\mathbf{q}^E - \mathbf{q}^B)^\top \mathbf{D}(\mathbf{u}^C, \mathbf{q}^B, t^B)(\mathbf{q}^E - \mathbf{q}^B)}_{\leq 0} \\ + \underbrace{\sum_{i \in \mathcal{I}(\mathbf{q}^B)} (\mathbf{q}^E - \mathbf{q}^B)^\top \mathbf{W}_i(\mathbf{q}^B, t^B) \hat{\boldsymbol{\lambda}}_i}_{\leq 0} \end{aligned} \quad (5.127)$$

From this the inequality  $E(\mathbf{u}^E, \mathbf{q}^E) \leq E(\mathbf{u}^C, \mathbf{q}^B)$  follows, which makes the integrator with damping force energetically consistent as well.

### 5.2.5 Quaternion Based Formulation

In this section, a quaternion based consistent integrator for non-smooth rigid body systems will be formulated. In this formulation, the quaternion based rigid body ODE (4.119) will be used as basis. Instead of replacing this ODE with the scalable body based rigid body DAE as in the previous sections, we will use a discretization of the kinematic equation, which preserves the unit length of the quaternion exactly. The discretization is specialized for the quaternion unit length constraint, and can not be used for other constraints in general.

Some of the complexity of the integrators from Section 5.2.2-5.2.4 is caused by the enforcement of the perfect bilateral constraints on displacement and velocity level. But removing the perfect bilateral constraints from these integrators would also remove the support for rigid body formulations, as affine and scalable body based rigid body DAE formulations require bilateral constraints. With the ODE based integrator described in this section, the general support for enforcing additional bilateral constraints on displacement level will be dropped, as it is not required for rigid body systems. This will allow to formulate a non-smooth consistent integrator with less complexity. Beside the rigid body model, we will consider potential forces and set-valued force laws of normal cone type. For each rigid body there is one vector of generalized coordinates

$$\mathbf{q}_i = \begin{pmatrix} \mathbf{r}_i \\ \boldsymbol{\psi}(A_i) \end{pmatrix} \quad (5.128)$$

where  $\mathbf{r}_i \in \mathbb{R}^3$  is the location of the center of mass and  $\boldsymbol{\psi}(A_i) \in \mathbb{R}^4$  the vector of the quaternion describing the orientation of the body. The complete generalized coordinates  $\mathbf{q}$  are obtained by assembling all  $\mathbf{q}_i \in \mathbb{R}^7$  into one vector. The generalized velocities for a body are given by

$$\mathbf{u}_i = \begin{pmatrix} \mathbf{v}_i \\ \boldsymbol{\omega}_i \end{pmatrix}. \quad (5.129)$$

The vector  $\mathbf{v}_i \in \mathbb{R}^3$  is the time derivative of  $\mathbf{r}_i$ . The angular velocity of the body is described by  $\boldsymbol{\omega}_i \in \mathbb{R}^3$ . Assembling all  $\mathbf{u}_i \in \mathbb{R}^6$  into one vector yields the complete vector of generalized velocities  $\mathbf{u}$ . The equations of motion of the mechanical system have the form

$$\begin{aligned} \mathbf{M}\dot{\mathbf{u}} + \mathbf{H}(\mathbf{u})\mathbf{u} + \mathbf{F}^\top(\mathbf{q}) \left( \frac{\partial V}{\partial \mathbf{q}} \right)^\top - \sum_{i \in \mathcal{I}(\mathbf{q})} \mathbf{W}_i(\mathbf{q}, t) \boldsymbol{\lambda}_i &= 0 \\ \dot{\mathbf{q}} &= \mathbf{F}(\mathbf{q})\mathbf{u} \\ i \in \mathcal{I}(\mathbf{q}) : \mathbf{W}_i(\mathbf{q}, t)^\top \mathbf{u} &\in \mathcal{N}_{\mathcal{C}_i}(-\boldsymbol{\lambda}_i). \end{aligned} \quad (5.130)$$

This is an assembled version of the equations of motion of each rigid body given by the ODE (4.119). The mass matrix  $\mathbf{M}$  is obtained by assembling the blocks

$$\mathbf{M}_{ii} = \begin{pmatrix} m_i \mathbf{I} & 0 \\ 0 & \boldsymbol{\Theta}_i \end{pmatrix}, \quad \mathbf{M}_{ij} = 0 \quad \forall i \neq j \quad (5.131)$$

into one matrix. The mass of the  $i$ -th body is abbreviated with  $m_i$ . The classical inertia tensor of the  $i$ -th body is denoted by  $\boldsymbol{\Theta}_i$ . The complete mass matrix  $\mathbf{M}$  is constant,

symmetric and positive definite. To express the inertial forces  $\tilde{\omega}_i \Theta_i \omega_i$  occurring in the Newton-Euler equations, the skew symmetric gyro matrix  $\mathbf{H}$  is introduced. The blocks to assemble  $\mathbf{H}$  are given by

$$\mathbf{H}_{ii}(\mathbf{u}) = \begin{pmatrix} 0 & 0 \\ 0 & \tilde{\omega}_i \Theta_i + \Theta_i \tilde{\omega}_i \end{pmatrix}, \quad \mathbf{H}_{ij}(\mathbf{u}) = 0 \quad \forall i \neq j, \quad \mathbf{H}^\top(\mathbf{u}) = -\mathbf{H}(\mathbf{u}). \quad (5.132)$$

The product

$$\mathbf{H}_{ii}(\mathbf{u})\mathbf{u}_i = \begin{pmatrix} 0 \\ \tilde{\omega}_i \Theta_i \omega_i \end{pmatrix} \quad (5.133)$$

then yields exactly the inertial forces occurring in the Newton-Euler equations. For the kinematic equation of each body (see right hand side of (4.119)) the matrix  $\mathbf{F}$  is introduced

$$\mathbf{F}_{ii}(\mathbf{q}) = \begin{pmatrix} \mathbf{I} & 0 \\ 0 & \frac{1}{2}\varphi(A_i) \end{pmatrix} \begin{pmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{I} \end{pmatrix}, \quad \mathbf{F}_{ij}(\mathbf{q}) = 0 \quad \forall i \neq j. \quad (5.134)$$

This allows to write all kinematic equations as  $\dot{\mathbf{q}} = \mathbf{F}(\mathbf{q})\mathbf{u}$ . The matrices of generalized force directions of the set-valued force laws are formulated with respect to the generalized velocities  $\mathbf{u}$ . One gets

$$\mathbf{W}_i(\mathbf{q}) = \mathbf{F}^\top \left( \frac{\partial \mathbf{g}_i}{\partial \mathbf{q}} \right)^\top. \quad (5.135)$$

for an originally geometric force law and

$$\mathbf{W}_i(\mathbf{q}, t) = \left( \frac{\partial \gamma_i}{\partial \mathbf{u}} \right)^\top \quad (5.136)$$

for a kinematic force law. For geometric unilateral constraints, the set  $\mathcal{I}(\mathbf{q})$  of active force laws can be used to disable the constraints on velocity level when the constraints are open on displacement level. The impact equations and impact laws associated with (5.130) are given by

$$\begin{aligned} \mathbf{M}(\mathbf{u}^+ - \mathbf{u}^-) - \sum_{i \in \mathcal{I}(\mathbf{q})} \mathbf{W}_i(\mathbf{q}, t) \Lambda_i &= 0 \\ i \in \mathcal{I}(\mathbf{q}) : \mathbf{W}_i^\top(\mathbf{q}, t)(\mathbf{u}^+ + \varepsilon_i \mathbf{u}^-) &\in \mathcal{N}_{\mathcal{D}_i}(-\Lambda_i). \end{aligned} \quad (5.137)$$

A Newton impact law in inclusion form with a coefficient of restitution  $\varepsilon_i$  is used for each force law. Before discretization, we will reformulate the impact laws

$$\begin{aligned} \mathbf{W}_i^\top(\mathbf{q}, t)(\mathbf{u}^+ + \varepsilon_i \mathbf{u}^-) &\in \mathcal{N}_{\mathcal{D}_i}(-\Lambda_i) \\ \frac{1}{2} \mathbf{W}_i^\top(\mathbf{q}, t)((1 + \varepsilon_i)(\mathbf{u}^+ + \mathbf{u}^-) + (1 - \varepsilon_i)(\mathbf{u}^+ - \mathbf{u}^-)) &\in \mathcal{N}_{\mathcal{D}_i}(-\Lambda_i) \\ \mathbf{W}_i^\top(\mathbf{q}, t)(\mathbf{u}^+ + \mathbf{u}^- + \frac{1 - \varepsilon_i}{1 + \varepsilon_i}(\mathbf{u}^+ - \mathbf{u}^-)) &\in \mathcal{N}_{\mathcal{D}_i}(-\Lambda_i) \end{aligned} \quad (5.138)$$

in terms of the sum  $\mathbf{u}^+ + \mathbf{u}^-$  and the difference  $\mathbf{u}^+ - \mathbf{u}^-$  as in Section 3.9. Solving the impact equation for the pre- and post-impact difference

$$\mathbf{u}^+ - \mathbf{u}^- = \mathbf{M}^{-1} \sum_{i \in \mathcal{I}(\mathbf{q})} \mathbf{W}_i(\mathbf{q}, t) \boldsymbol{\Lambda}_i \quad (5.139)$$

and inserting it into the impact law (5.138) yields the formulation

$$\mathbf{W}_i^\top(\mathbf{q}, t) (\mathbf{u}^+ + \mathbf{u}^- + \frac{1 - \varepsilon_i}{1 + \varepsilon_i} \mathbf{M}^{-1} \sum_{j \in \mathcal{I}(\mathbf{q})} \mathbf{W}_j(\mathbf{q}, t) \boldsymbol{\Lambda}_j) \in \mathcal{N}_{\mathcal{D}_i}(-\boldsymbol{\Lambda}_i). \quad (5.140)$$

which is still equivalent to the one used in (5.137). The reformulation of the impact laws will allow a one-step discretization that is consistent with respect to the total energy. As direct discretization of the equations of motion (5.130), the impact equations (5.137) and the impact laws (5.140) we propose the following difference scheme:

$$\begin{aligned} & \mathbf{M}(\mathbf{u}^E - \mathbf{u}^B) + \frac{\Delta t}{2} \mathbf{H} \left( \frac{\mathbf{u}^E + \mathbf{u}^B}{2} \right) (\mathbf{u}^E + \mathbf{u}^B) + \Delta t \mathbf{F}^\top \left( \frac{\mathbf{q}^E + \mathbf{q}^B}{2} \right) \left( \frac{\text{D}V}{\text{D}\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E) \right)^\top \\ & - \sum_{i \in \mathcal{I}(\mathbf{q}^B)} \mathbf{W}_i \hat{\boldsymbol{\lambda}}_i = 0 \\ & \mathbf{q}^E - \mathbf{q}^B = \frac{\Delta t}{2} \mathbf{F} \left( \frac{\mathbf{q}^E + \mathbf{q}^B}{2} \right) (\mathbf{u}^E + \mathbf{u}^B) \\ & i \in \mathcal{I}(\mathbf{q}^B) : \mathbf{W}_i^\top (\mathbf{u}^E + \mathbf{u}^B + \frac{1 - \varepsilon_i}{1 + \varepsilon_i} \mathbf{M}^{-1} \sum_{i \in \mathcal{I}(\mathbf{q}^B)} \mathbf{W}_i \hat{\boldsymbol{\lambda}}_i) \in \mathcal{N}_{\mathcal{D}_i}(-\hat{\boldsymbol{\lambda}}_i). \end{aligned} \quad (5.141)$$

In this integrator again the partial derivative of the potential energy is approximated using the discrete derivative. The vectors  $\hat{\boldsymbol{\lambda}}_i$  approximate the time integral of  $\boldsymbol{\lambda}_i$  over the time step plus the impact-impulses  $\boldsymbol{\Lambda}_i$ . For the discrete approximation of the matrices of generalized force directions, a value

$$\mathbf{W}_i = \mathbf{W}_i(\mathbf{q}^B, t^B) \quad (5.142)$$

based on the state at the beginning of a time step can be used. An alternative for the generalized force directions of unilateral contacts is the approximation

$$\mathbf{W}_i = \mathbf{F}^\top \left( \frac{\mathbf{q}^E + \mathbf{q}^B}{2} \right) \left( \frac{\text{D}g_i}{\text{D}\mathbf{q}}(\mathbf{q}^E, \mathbf{q}^B) \right)^\top. \quad (5.143)$$

based on the discrete derivative of the gap function. With a coefficient of restitution  $\varepsilon_i$  equal to one, the inclusion in (5.141) and the generalized force direction approximation (5.143)

can be simplified as follows:

$$\begin{aligned}
\mathbf{W}_i^\top(\mathbf{u}^E + \mathbf{u}^B) &\in \mathcal{N}_{\mathcal{D}_i}(-\hat{\boldsymbol{\lambda}}_i) \\
\Leftrightarrow \frac{D\mathbf{g}_i}{D\mathbf{q}}(\mathbf{q}^E, \mathbf{q}^B)\mathbf{F}\left(\frac{\mathbf{q}^E + \mathbf{q}^B}{2}\right)(\mathbf{u}^E + \mathbf{u}^B) &\in \mathcal{N}_{\mathcal{D}_i}(-\hat{\boldsymbol{\lambda}}_i) \\
\Leftrightarrow \frac{D\mathbf{g}_i}{D\mathbf{q}}(\mathbf{q}^E, \mathbf{q}^B)(\mathbf{q}^E - \mathbf{q}^B) &\in \mathcal{N}_{\mathcal{D}_i}(-\hat{\boldsymbol{\lambda}}_i) \\
\Leftrightarrow \mathbf{g}_i(\mathbf{q}^E) - \mathbf{g}_i(\mathbf{q}^B) &\in \mathcal{N}_{\mathcal{D}_i}(-\hat{\boldsymbol{\lambda}}_i)
\end{aligned} \tag{5.144}$$

From (5.144) we get directly the consistency condition on velocity level for perfect elastic unilateral contacts discussed at the beginning of Section 5.2.

In the following, we will verify that the discretization of the kinematic equation  $\dot{\mathbf{q}} = \mathbf{F}(\mathbf{q})\mathbf{u}$  used in the integrator (5.141) preserves the unit length of a quaternion exactly. The discrete kinematic equation from (5.141) concerning the quaternion of the  $i$ -th rigid body can be written as

$$\boldsymbol{\psi}(A_i^E) - \boldsymbol{\psi}(A_i^B) = \frac{\Delta t}{4}\boldsymbol{\varphi}\left(\frac{A_i^E + A_i^B}{2}\right)\begin{pmatrix} 0 \\ \boldsymbol{\omega}_i^E + \boldsymbol{\omega}_i^B \end{pmatrix} \tag{5.145}$$

by inserting (5.134) into the second equation of (5.141). Premultiplying (5.145) from the left with  $\boldsymbol{\psi}^\top(A_i^E + A_i^B)$  and simplifying the resulting equation yields

$$\begin{aligned}
\boldsymbol{\psi}^\top(A_i^E + A_i^B)(\boldsymbol{\psi}(A_i^E) - \boldsymbol{\psi}(A_i^B)) &= \frac{\Delta t}{8}\boldsymbol{\psi}^\top(A_i^E + A_i^B)\boldsymbol{\varphi}(A_i^E + A_i^B)\begin{pmatrix} 0 \\ \boldsymbol{\omega}_i^E + \boldsymbol{\omega}_i^B \end{pmatrix} \\
(\boldsymbol{\psi}^\top(A_i^E) + \boldsymbol{\psi}^\top(A_i^B))(\boldsymbol{\psi}(A_i^E) - \boldsymbol{\psi}(A_i^B)) &= \frac{\Delta t}{8}\boldsymbol{\psi}^\top(A_i^E + A_i^B)\boldsymbol{\varphi}(A_i^E + A_i^B)\begin{pmatrix} 0 \\ \boldsymbol{\omega}_i^E + \boldsymbol{\omega}_i^B \end{pmatrix} \\
|A_i^E|^2 - |A_i^B|^2 &= \frac{\Delta t}{8}\boldsymbol{\psi}^\top((A_i^E + A_i^B)^*(A_i^E + A_i^B))\begin{pmatrix} 0 \\ \boldsymbol{\omega}_i^E + \boldsymbol{\omega}_i^B \end{pmatrix} \\
|A_i^E|^2 - |A_i^B|^2 &= \frac{\Delta t}{8}(|A_i^E + A_i^B|^2 \quad 0)\begin{pmatrix} 0 \\ \boldsymbol{\omega}_i^E + \boldsymbol{\omega}_i^B \end{pmatrix} \\
|A_i^E|^2 - |A_i^B|^2 &= 0.
\end{aligned} \tag{5.146}$$

From (5.146) then follows the equation

$$|A_i^E| = |A_i^B|. \tag{5.147}$$

A unit quaternion  $|A_i^B| = 1$  at the beginning of a time step yields a quaternion  $A_i^E$  at the end of the time step which has exactly unit length as well.

The total energy  $E(\mathbf{u}, \mathbf{q})$  of the rigid body system (5.130) is the sum of the kinetic energy

$$T(\mathbf{u}) = \frac{1}{2}\mathbf{u}^\top \mathbf{M} \mathbf{u} \tag{5.148}$$



and the potential energy  $V(\mathbf{q})$ . The consistency of the integrator (5.141) with respect to the total energy can be verified by multiplying the first equation of the integrator (5.141) with  $(\mathbf{u}^E + \mathbf{u}^B)^\top/2$  from the left. This yields the following:

$$\begin{aligned} & \underbrace{\frac{1}{2}(\mathbf{u}^E + \mathbf{u}^B)^\top \mathbf{M}(\mathbf{u}^E - \mathbf{u}^B)}_{= T(\mathbf{u}^E) - T(\mathbf{u}^B)} + \underbrace{\frac{\Delta t}{4}(\mathbf{u}^E + \mathbf{u}^B)^\top \mathbf{H} \left( \frac{\mathbf{u}^E + \mathbf{u}^B}{2} \right) (\mathbf{u}^E + \mathbf{u}^B)}_{= 0} \\ & + \underbrace{\frac{\Delta t}{2}(\mathbf{u}^E + \mathbf{u}^B)^\top \mathbf{F}^\top \left( \frac{\mathbf{q}^E + \mathbf{q}^B}{2} \right)}_{= \mathbf{q}^E - \mathbf{q}^B} \left( \frac{DV}{D\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E) \right)^\top - \frac{1}{2}(\mathbf{u}^E + \mathbf{u}^B)^\top \sum_{i \in \mathcal{I}(\mathbf{q}^B)} \mathbf{W}_i \hat{\boldsymbol{\lambda}}_i = 0 \end{aligned} \quad (5.149)$$

The first term is obviously the difference in kinetic energy between the end and the beginning of the time step. The second term is zero, because the gyro matrix  $\mathbf{H}$  is skew symmetric. The expression in front of the discrete derivative of the potential energy is just the difference  $\mathbf{q}^E - \mathbf{q}^B$ . Multiplying this difference with the discrete derivative of the potential energy yields the difference in potential energy between the end and the beginning of the time step. Replacing the resulting kinetic and potential energy differences in (5.149) with the difference in total energy yields

$$E(\mathbf{u}^E, \mathbf{q}^E) - E(\mathbf{u}^B, \mathbf{q}^B) = \frac{1}{2} \sum_{i \in \mathcal{I}(\mathbf{q}^B)} \hat{\boldsymbol{\lambda}}_i^\top \mathbf{W}_i^\top (\mathbf{u}^E + \mathbf{u}^B). \quad (5.150)$$

If no set-valued force laws are present or all  $\hat{\boldsymbol{\lambda}}_i$  are zero, then the total energy is preserved. This is consistent with the equations of motion (5.130). From the normal cone inclusion in (5.141) and convex sets  $\mathcal{D}_i$  that contain the zero element, we obtain the inequality

$$\hat{\boldsymbol{\lambda}}_i^\top \mathbf{W}_i^\top (\mathbf{u}^E + \mathbf{u}^B + \frac{1 - \varepsilon_i}{1 + \varepsilon_i} \mathbf{M}^{-1} \sum_{j \in \mathcal{I}(\mathbf{q}^B)} \mathbf{W}_j \hat{\boldsymbol{\lambda}}_j) \leq 0. \quad (5.151)$$

for each index in the set  $\mathcal{I}(\mathbf{q}^B)$ . Summing up these inequalities over all indices yields

$$\sum_{i \in \mathcal{I}(\mathbf{q}^B)} \hat{\boldsymbol{\lambda}}_i^\top \mathbf{W}_i^\top (\mathbf{u}^E + \mathbf{u}^B) + \sum_{i \in \mathcal{I}(\mathbf{q}^B)} \hat{\boldsymbol{\lambda}}_i^\top \frac{1 - \varepsilon_i}{1 + \varepsilon_i} \mathbf{W}_i^\top \mathbf{M}^{-1} \sum_{j \in \mathcal{I}(\mathbf{q}^B)} \mathbf{W}_j \hat{\boldsymbol{\lambda}}_j \leq 0. \quad (5.152)$$

One obtains the inequality

$$E(\mathbf{u}^E, \mathbf{q}^E) - E(\mathbf{u}^B, \mathbf{q}^B) \leq -\frac{1}{2} \sum_{i \in \mathcal{I}(\mathbf{q})} \sum_{j \in \mathcal{I}(\mathbf{q}^B)} \hat{\boldsymbol{\lambda}}_i^\top \frac{1 - \varepsilon_i}{1 + \varepsilon_i} \mathbf{W}_i^\top \mathbf{M}^{-1} \mathbf{W}_j \hat{\boldsymbol{\lambda}}_j \quad (5.153)$$

after inserting (5.150) into (5.152). Assembling the vectors  $\hat{\boldsymbol{\lambda}}_i$  and the matrices  $\mathbf{W}_i$  into a global vector  $\hat{\boldsymbol{\lambda}}$  and a global matrix  $\mathbf{W}$  and using the matrix of dissipation indices  $\boldsymbol{\Delta}$  from Section 3.9 simplifies (5.153) to

$$E(\mathbf{u}^E, \mathbf{q}^E) - E(\mathbf{u}^B, \mathbf{q}^B) \leq -\frac{1}{2} \hat{\boldsymbol{\lambda}}^\top \boldsymbol{\Delta} \mathbf{W}^\top \mathbf{M}^{-1} \mathbf{W} \hat{\boldsymbol{\lambda}}. \quad (5.154)$$

If the matrix  $\Delta \mathbf{W}^\top \mathbf{M}^{-1} \mathbf{W}$  is positive semidefinite, then the total energy is guaranteed not to increase, which is consistent with the properties of (5.130) and (5.137) shown in Section 3.9. For example if we use one global coefficient of restitution

$$\Delta = \frac{1 - \varepsilon}{1 + \varepsilon} \mathbf{I} \quad (5.155)$$

then the matrix  $\Delta \mathbf{W}^\top \mathbf{M}^{-1} \mathbf{W}$  is positive semidefinite because  $\mathbf{W}^\top \mathbf{M}^{-1} \mathbf{W}$  is positive semidefinite. From (5.154) then follows that the total energy in the integrator (5.141) is never increasing

$$E(\mathbf{u}^E, \mathbf{q}^E) \leq E(\mathbf{u}^B, \mathbf{q}^B) \quad (5.156)$$

for this case. The inclusion problem (5.141) for a time step of the integrator contains beside the normal cone inclusions also equations which are nonlinear in the unknowns. For the numerical solution the projected Newton iteration described in Section 2.5.5 can be used.

The integrator (5.141) discussed above formulates one inclusion problem to approximate both, the impact-free motion and the impacts. Alternatively, also a two-step integrator as discussed in Section 5.2.4 can be formulated. The first step approximating the impacts is then given by

$$\begin{aligned} \mathbf{M}(\mathbf{u}^C - \mathbf{u}^B) - \sum_{i \in \mathcal{I}(\mathbf{q}^B)} \mathbf{W}_i(\mathbf{q}^B, t^B) \Lambda_i &= 0 \\ i \in \mathcal{I}(\mathbf{q}^B) : \mathbf{W}_i^\top(\mathbf{q}^B, t^B)(\mathbf{u}^C + \varepsilon_i \mathbf{u}^B) &\in \mathcal{N}_{\mathcal{D}_i}(-\Lambda_i). \end{aligned} \quad (5.157)$$

This impact inclusion yields the transition from the beginning state  $(\mathbf{q}^B, \mathbf{u}^B)$  to the post impact state  $(\mathbf{q}^B, \mathbf{u}^C)$ . The second step approximating the impact free motion is formulated as

$$\begin{aligned} \mathbf{M}(\mathbf{u}^E - \mathbf{u}^C) + \frac{\Delta t}{2} \mathbf{H} \left( \frac{\mathbf{u}^E + \mathbf{u}^C}{2} \right) (\mathbf{u}^E + \mathbf{u}^C) + \Delta t \mathbf{F}^\top \left( \frac{\mathbf{q}^E + \mathbf{q}^B}{2} \right) \left( \frac{D\mathbf{V}}{D\mathbf{q}}(\mathbf{q}^B, \mathbf{q}^E) \right)^\top \\ - \sum_{i \in \mathcal{I}(\mathbf{q}^B)} \mathbf{W}_i \hat{\lambda}_i &= 0 \\ \mathbf{q}^E - \mathbf{q}^B &= \frac{\Delta t}{2} \mathbf{F} \left( \frac{\mathbf{q}^E + \mathbf{q}^B}{2} \right) (\mathbf{u}^E + \mathbf{u}^C) \\ i \in \mathcal{I}(\mathbf{q}^B) : \mathbf{W}_i^\top(\mathbf{u}^E + \mathbf{u}^C) &\in \mathcal{N}_{\hat{\mathcal{C}}_i}(-\hat{\lambda}_i) \end{aligned} \quad (5.158)$$

and yields the end time state  $(\mathbf{q}^E, \mathbf{u}^E)$  starting from the intermediate state  $(\mathbf{q}^B, \mathbf{u}^C)$ . The first step (5.157) of the integrator does not change the generalized coordinates. The consistency with respect to the total energy can be obtained by the same argument as in Section 5.2.4, since the first step (5.157) has exactly the same structure as the impact inclusion (5.137). The second step (5.158) approximates the equations of motion (5.130) and has the same properties as (5.141) for the special case  $\varepsilon_i = 1$ . This means the total energy is consistent with respect to the equations of motion, and also the consistency for unilateral contacts on velocity level as shown in (5.144) holds. While the two-step

---

version of the quaternion based consistent integrator is usually more expensive to evaluate numerically, it allows to get the consistency for unilateral contacts on velocity level for any coefficients of restitution, and not only for the perfectly elastic contacts like the one-step integrator (5.141). The inclusion problem of the first step (5.157) can be solved with the projected Jacobi iteration or the projected Gauss-Seidel iteration discussed in Section 2.5. The second step (5.158) contains also nonlinear equations beside the normal cone inclusions. It can be solved with the projected Newton iteration discussed in Section 2.5.5.



## Examples

---

In this chapter, a few examples of non-smooth mechanical systems are described and their numerical solution with the integrators from Chapter 5 is discussed.

### 6.1 Oscillator

In this section, a mass-spring system with a unilateral contact as shown in Figure 6.1 is considered. The mass  $m_2$  is chosen much smaller than the mass  $m_1$ . The vertical position

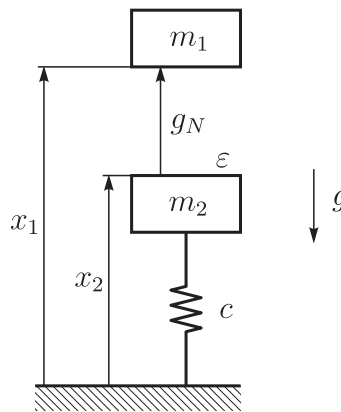


Figure 6.1: Oscillator with unilateral contact.

and vertical velocity of the two masses are used as generalized coordinates and generalized velocities, respectively

$$\mathbf{q} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \mathbf{u} = \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix}, \quad \mathbf{u} = \dot{\mathbf{q}}. \quad (6.1)$$

The mass matrix and kinetic energy of the system is given by

$$\mathbf{M} = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix}, \quad T(\dot{\mathbf{q}}) = \frac{1}{2} \dot{\mathbf{q}}^\top \mathbf{M} \dot{\mathbf{q}}. \quad (6.2)$$

The potential function

$$V(\mathbf{q}) = \frac{c}{2}(x_2 - l)^2 + g(m_1 x_1 + m_2 x_2) \quad (6.3)$$

describes the gravitational forces and the force of the spring. The spring has a force free length equal to  $l$ . The gap function for the unilateral contact is just the difference of the two coordinates

$$g_N(\mathbf{q}) = x_1 - x_2. \quad (6.4)$$

The generalized force direction of the unilateral contact is obtained as

$$\mathbf{w}_N = \left( \frac{Dg_N}{D\mathbf{q}} \right)^\top = \left( \frac{\partial g_N}{\partial \mathbf{q}} \right)^\top = \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (6.5)$$

The equations of motion of the oscillator with unilateral contact are given by

$$\begin{aligned} \mathbf{M}\dot{\mathbf{u}} + \left( \frac{\partial V}{\partial \mathbf{q}} \right)^\top - \mathbf{w}_N \lambda_N &= 0 \\ -\dot{\mathbf{q}} + \mathbf{u} &= 0 \\ g_N(\mathbf{q}) &\in \mathcal{N}_{\mathbb{R}_0^-}(-\lambda_N) \end{aligned} \quad (6.6)$$

For the unilateral contact an impact law of Newton type is used. The impact equations and the impact law have the form

$$\begin{aligned} \mathbf{M}(\mathbf{u}^+ - \mathbf{u}^-) - \mathbf{w}_N \Lambda_N &= 0 \\ g_N(\mathbf{q}) > 0 : \Lambda_N &= 0 \\ g_N(\mathbf{q}) = 0 : \mathbf{w}_N^\top (\mathbf{u}^+ + \varepsilon \mathbf{u}^-) &\in \mathcal{N}_{\mathbb{R}_0^-}(-\Lambda_N). \end{aligned} \quad (6.7)$$

The parameters and initial conditions listed in Table 6.1 are used for the numerical simulation with the Moreau's midpoint rule and the consistent two-step integrator given by (5.106) and (5.107).

Mass	$m_1 = 10 \text{ kg}$ $m_2 = 10^{-3} \text{ kg}$
Stiffness	$c = 4 \cdot 10^5 \text{ N m}^{-1}$
Force free length	$l = 0.1 \text{ m}$
Gravitational acceleration	$g = 9.81 \text{ m s}^{-2}$
Coefficient of restitution	$\varepsilon = 0$
Initial coordinates	$x_1(0) = 0.11 \text{ m}$ $x_2(0) = 0.1 \text{ m}$
Initial velocities	$\dot{x}_{1,2}(0) = 0 \text{ m s}^{-1}$

Table 6.1: Parameters and initial conditions.

In Figure 6.2 the results obtained with the consistent two-step integrator and a time step size equal to  $10^{-3} \text{ s}$  are shown. Using larger time steps decreases the accuracy of the

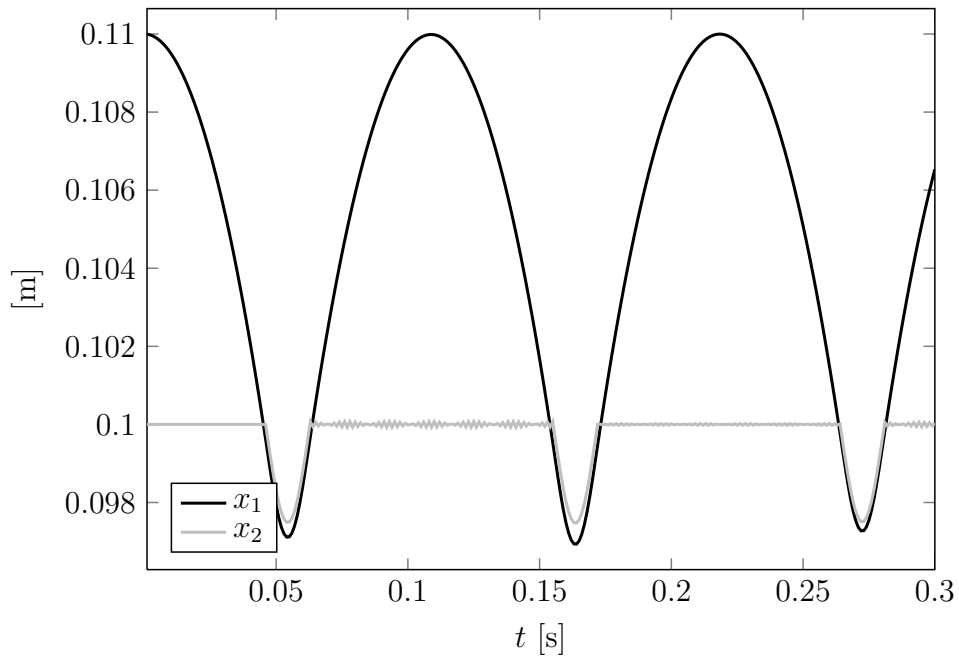


Figure 6.2: Consistent integrator solution for the oscillator ( $\Delta t = 10^{-3}$  s).

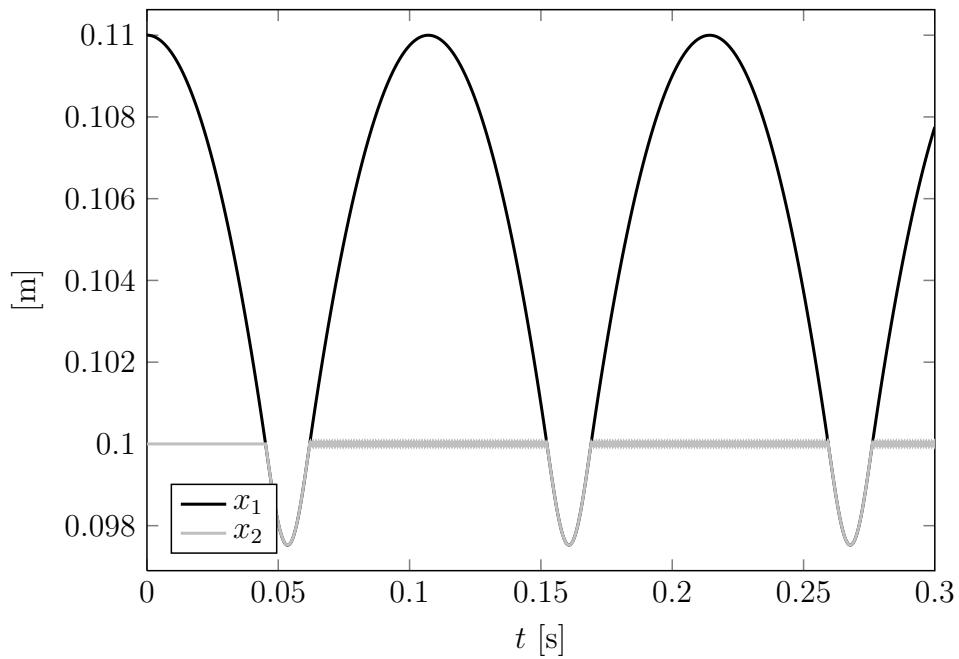


Figure 6.3: Solution for the oscillator ( $\Delta t = 10^{-5}$  s).

solution, but the consistent integrator remains stable. The total energy as approximated by the numerical solution is decreasing stepwise, which is consistent with the mechanical model. An attempt to approximate a solution with the Moreau's midpoint rule and the same step size of  $10^{-3}$  s fails. The integrator becomes unstable and the solution diverges towards infinity. To get a non-divergent solution with Moreau's midpoint rule, a time step size smaller than  $10^{-4}$  s is required. The reason for this, is the relatively high eigenfrequency of the mass spring system formed by the mass  $m_2$  and the spring. Moreau's midpoint rule uses an explicit discretization of the classical forces, which has a limited stability region. In Figure 6.3 the solution obtained for both integrators at a time step equal to  $10^{-5}$  s is shown. In order to remove the small time step requirement imposed on the integrator one can either remove the high frequency dynamics from the model or use an integrator which discretizes the relevant parts with an implicit approach, as for example the consistent integrator. Removing the high frequency dynamics from the model for all parts of the dynamics could be difficult in general. Note that for the system discussed in this section, also a variable time step scheme like the one developed in [78] would not reduce the total number of time steps required.

## 6.2 Unilateral Double Pendulum

The planar unilateral double pendulum considered in this section consists of two mass points with mass  $m$ , two massless rigid rods of length  $l$  and a rigid wall. The rigid rods connect the mass points to each other and to the wall as illustrated in Figure 6.4. The pendulum is under the influence of gravity  $g$ . The unilateral contacts between mass points

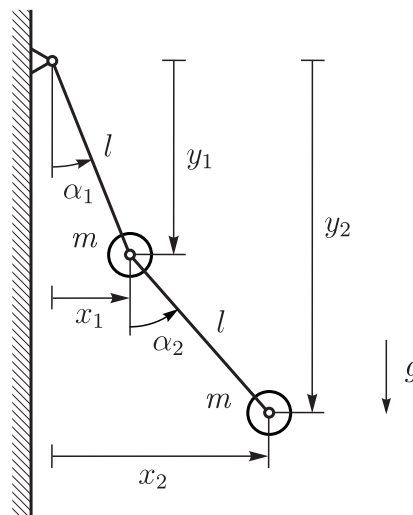


Figure 6.4: Double pendulum with wall.

and wall are assumed to be perfectly elastic. The horizontal and vertical displacements



of the mass points form the generalized coordinates

$$\mathbf{q} = \begin{pmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \end{pmatrix}, \quad \mathbf{u} = \begin{pmatrix} \dot{x}_1 \\ \dot{y}_1 \\ \dot{x}_2 \\ \dot{y}_2 \end{pmatrix}, \quad \mathbf{u} = \dot{\mathbf{q}}. \quad (6.8)$$

The time derivatives of the generalized coordinates are used as generalized velocities. With the constant mass matrix

$$\mathbf{M} = \begin{pmatrix} m & 0 & 0 & 0 \\ 0 & m & 0 & 0 \\ 0 & 0 & m & 0 \\ 0 & 0 & 0 & m \end{pmatrix} \quad (6.9)$$

one can write the kinetic energy of the double pendulum as

$$T(\dot{\mathbf{q}}) = \frac{1}{2} \dot{\mathbf{q}}^\top \mathbf{M} \dot{\mathbf{q}}. \quad (6.10)$$

The gravitational force are described with the potential function

$$V(\mathbf{q}) = 3mgl - mg(y_1 + y_2). \quad (6.11)$$

The potential is chosen such that it is zero in the lowest equilibrium position of the double pendulum. The perfect bilateral constraints imposed by the rigid rods are described with the constraint function

$$\mathbf{g}_S(\mathbf{q}) = \frac{1}{2} \begin{pmatrix} x_1^2 + y_1^2 - l^2 \\ (x_2 - x_1)^2 + (y_2 - y_1)^2 - l^2 \end{pmatrix}. \quad (6.12)$$

With the partial derivative

$$\left( \frac{\partial \mathbf{g}_S}{\partial \mathbf{q}} \right)^\top = \begin{pmatrix} x_1 & x_1 - x_2 \\ y_1 & y_1 - y_2 \\ 0 & x_2 - x_1 \\ 0 & y_2 - y_1 \end{pmatrix} \quad (6.13)$$

the constraint function

$$\boldsymbol{\gamma}_S(\mathbf{u}, \mathbf{q}) = \frac{\partial \mathbf{g}_S}{\partial \mathbf{q}} \mathbf{u} = \begin{pmatrix} x_1 \dot{x}_1 + y_1 \dot{y}_1 \\ (x_1 - x_2) \dot{x}_1 + (y_1 - y_2) \dot{y}_1 + (x_2 - x_1) \dot{x}_2 + (y_2 - y_1) \dot{y}_2 \end{pmatrix} \quad (6.14)$$

of the kinematic constraints associated with  $\mathbf{g}_S$  can be formulated. The normal contacts between the wall and the pendulum restrict the horizontal displacements to positive values. To formulate the normal contact force law the gap function

$$\mathbf{g}_N(\mathbf{q}) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (6.15)$$

is used. With the help of Chapter 3 one gets the equations of motion

$$\begin{aligned} \mathbf{M}\dot{\mathbf{u}} + \left(\frac{\partial V}{\partial \mathbf{q}}\right)^\top - \left(\frac{\partial \mathbf{g}_S}{\partial \mathbf{q}}\right)^\top \boldsymbol{\lambda}_S - \left(\frac{\partial \mathbf{g}_N}{\partial \mathbf{q}}\right)^\top \boldsymbol{\lambda}_N &= 0 \\ -\dot{\mathbf{q}} + \mathbf{u} &= 0 \\ \mathbf{g}_S(\mathbf{q}) &= 0 \\ g_{Ni}(\mathbf{q}) &\in \mathcal{N}_{\mathbb{R}_0^-}(-\lambda_{Ni}) \end{aligned} \quad (6.16)$$

for the unilateral double pendulum. The impact equations and the impact laws are then given by

$$\begin{aligned} \mathbf{M}(\mathbf{u}^+ - \mathbf{u}^-) - \left(\frac{\partial \mathbf{g}_S}{\partial \mathbf{q}}\right)^\top \boldsymbol{\Lambda}_S - \sum_{i \in \mathcal{I}(\mathbf{q})} \left(\frac{\partial g_{Ni}}{\partial \mathbf{q}}\right)^\top \boldsymbol{\Lambda}_{Ni} &= 0 \\ \boldsymbol{\gamma}_S(\mathbf{u}^+, \mathbf{q}) &= 0 \\ i \in \mathcal{I}(\mathbf{q}) : \frac{\partial g_{Ni}}{\partial \mathbf{q}}(\mathbf{u}^+ + \varepsilon \mathbf{u}^-) &\in \mathcal{N}_{\mathbb{R}_0^-}(-\boldsymbol{\Lambda}_{Ni}) \end{aligned} \quad (6.17)$$

where the set of active normal contacts is defined as

$$\mathcal{I}(\mathbf{q}) = \{i \mid g_{Ni}(\mathbf{q}) = 0\}. \quad (6.18)$$

After reformulating the unilateral constraints on velocity level, the description (6.16) and (6.17) has the form used in Section 5.2.4 for the consistent two-step integrator. For the numerical integration, the partial and discrete derivatives of the potential and constraint functions are required. One obtains

$$\left(\frac{DV}{D\mathbf{q}}\right)^\top = \left(\frac{\partial V}{\partial \mathbf{q}}\right)^\top = \begin{pmatrix} 0 \\ -mg \\ 0 \\ -mg \end{pmatrix}, \quad \left(\frac{D\mathbf{g}_N}{D\mathbf{q}}\right)^\top = \left(\frac{\partial \mathbf{g}_N}{\partial \mathbf{q}}\right)^\top = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \quad (6.19)$$

for the potential and gap function. The function  $\mathbf{g}_S$  is quadratic in the generalized coordinates which yields the discrete derivative

$$\frac{D\mathbf{g}_S}{D\mathbf{q}}(\mathbf{a}, \mathbf{b}) = \frac{\partial \mathbf{g}_S}{\partial \mathbf{q}}\left(\frac{\mathbf{a} + \mathbf{b}}{2}\right) \quad (6.20)$$

as midpoint partial derivative (see Section 5.2.1). For the velocity level constraint function one gets

$$\left(\frac{\partial \boldsymbol{\gamma}_S}{\partial \mathbf{q}}\right)^\top = \begin{pmatrix} \dot{x}_1 & \dot{x}_1 - \dot{x}_2 \\ \dot{y}_1 & \dot{y}_1 - \dot{y}_2 \\ 0 & \dot{x}_2 - \dot{x}_1 \\ 0 & \dot{y}_2 - \dot{y}_1 \end{pmatrix}, \quad \frac{D\boldsymbol{\gamma}_S}{D\mathbf{q}}(\mathbf{u}, \mathbf{a}, \mathbf{b}) = \frac{\partial \boldsymbol{\gamma}_S}{\partial \mathbf{q}}(\mathbf{u}). \quad (6.21)$$

The parameters and initial conditions listed in Table 6.2 are used for the numerical simulation. In Figure 6.5, a numerical solution for the rod angles  $\alpha_1$  and  $\alpha_2$  obtained with the consistent two-step integrator given by (5.106) and (5.107) is shown as  $\alpha_{1C}$  and  $\alpha_{2C}$ .

Point mass	$m = 0.2 \text{ kg}$
Rod length	$l = 0.5 \text{ m}$
Gravitational acceleration	$g = 9.81 \text{ m s}^{-2}$
Coefficient of restitution	$\varepsilon = 1$
Initial coordinates	$x_1(0) = 0 \text{ m}$ $y_1(0) = 0.5 \text{ m}$ $x_2(0) = 0 \text{ m}$ $y_2(0) = 1 \text{ m}$
Initial velocities	$\dot{x}_1(0) = 2.5 \text{ m s}^{-1}$ $\dot{y}_1(0) = 0 \text{ m s}^{-1}$ $\dot{x}_2(0) = 5 \text{ m s}^{-1}$ $\dot{y}_2(0) = 0 \text{ m s}^{-1}$

Table 6.2: Parameters and initial conditions.

As numerical reference, a solution obtained with Moreau's midpoint rule (see Section 5.1) is shown as  $\alpha_{1M}$  and  $\alpha_{2M}$ . For the integration with Moreau's midpoint rule, a minimal model based on the angles  $\alpha_1$  and  $\alpha_2$  has been used. Both numerical solutions are calculated with a relatively large time step  $\Delta t = 0.01 \text{ s}$ . In Figure 6.6 the total energy  $E_C$

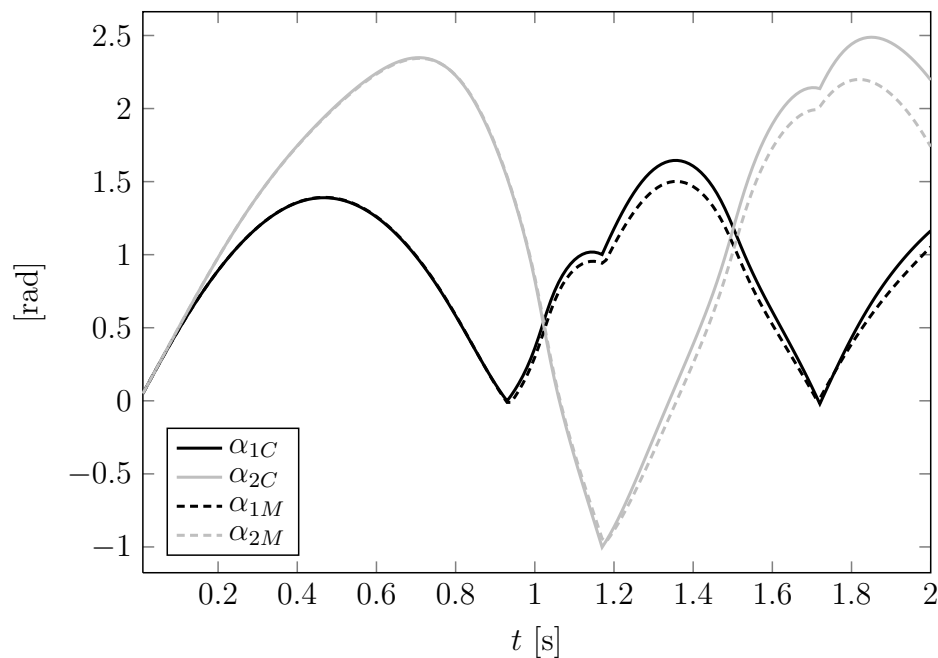


Figure 6.5: Angles of the double pendulum.

and  $E_M$  relative to the initial energy  $E_0$  is shown for the numerical solutions obtained with Moreau's midpoint rule and the consistent two-step integrator. For Moreau's midpoint rule we get a significant deviation of the energy from its initial value, although the total

energy is preserved for an exact solution. With the consistent integrator, the total energy is preserved up to the numerical precision. Reducing the size of the time step lets both solutions converge towards the exact solution. Instead of the consistent two-step integrator, also the consistent integrator (5.95) for systems with perfect unilateral constraints could have been used. Results obtained with (5.95) are very similar (but not identical) to the results from the consistent two-step integrator.

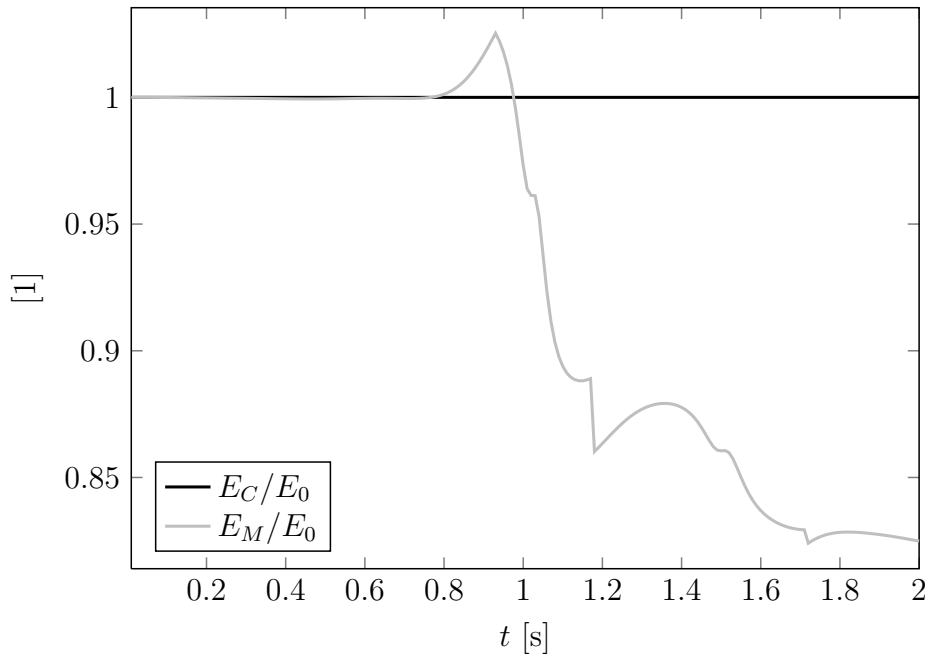


Figure 6.6: Total energy of the double pendulum.

### 6.3 Overbalanced Wheel

The overbalanced wheel was an early attempt at constructing a perpetual motion machine, dating back at least to the Middle Ages. A version of an overbalanced wheel is shown in Figure 6.7. It was argued that the lever arms on right side of the wheel axle are longer than those on the left side. So it was hoped to rotate in clockwise direction ad infinitum producing free energy. But while the lever arms are really longer on the right side in general, there are also more levers on the left side. So of course it does not produce any energy. In this section, we will use the overbalanced wheel just as an example of a mechanical system which is a bit more complicated than the unilateral double pendulum from Section 6.2.

The basis of the mechanical system depicted in Figure 6.7 is a wheel with radius  $a$  and moment of inertia  $\theta$  which is supported at its center by a frictionless bearing. Rigid and massless rods of length  $l$  connect  $n$  mass points with mass  $m$  to the wheel. The system is under the influence of gravity  $g$ . The rods are assumed to be distributed evenly around

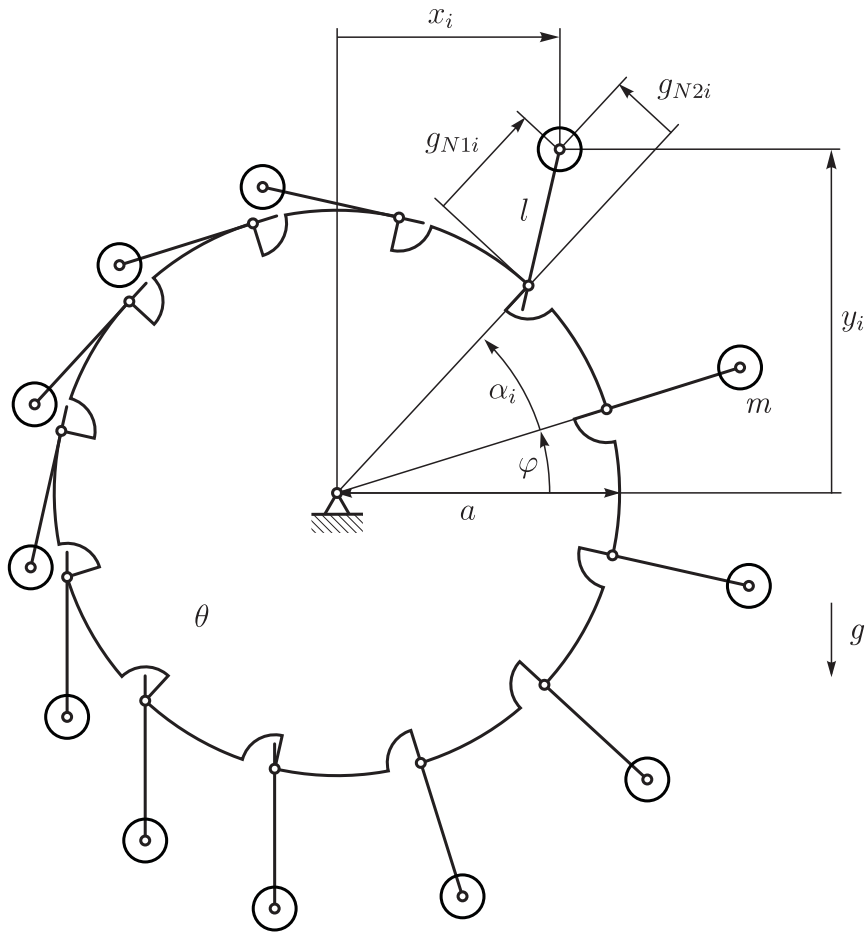


Figure 6.7: Overbalanced wheel.

the wheel. Starting from the first one, the angle to the frictionless bearing of the  $i$ -th rod is given by

$$\alpha_i = \frac{2\pi(i-1)}{n}. \quad (6.22)$$

Each lever is restricted by unilateral constraints to an at most tangential position on one side, and an at most perpendicular position on the other side (see Figure 6.7). As generalized coordinates we will use the angle  $\varphi$  of the wheel and the horizontal and vertical displacements of the mass points

$$\mathbf{q} = (\varphi \ x_1 \ x_2 \ \dots \ x_n \ y_1 \ y_2 \ \dots \ y_n)^\top. \quad (6.23)$$

The time derivatives of the generalized coordinates are used as generalized velocities

$$\mathbf{u} = \dot{\mathbf{q}}, \quad \mathbf{u} = (\dot{\varphi} \ \dot{x}_1 \ \dot{x}_2 \ \dots \ \dot{x}_n \ \dot{y}_1 \ \dot{y}_2 \ \dots \ \dot{y}_n)^\top. \quad (6.24)$$

With the constant mass matrix

$$\mathbf{M} = \begin{pmatrix} \theta & 0 \\ 0 & m\mathbf{I} \end{pmatrix} \quad (6.25)$$

we get the kinetic energy of the overbalanced wheel as

$$T(\dot{\mathbf{q}}) = \frac{1}{2} \dot{\mathbf{q}}^\top \mathbf{M} \dot{\mathbf{q}}. \quad (6.26)$$

The gravitational forces can be described with the potential energy function

$$V(\mathbf{q}) = mg \sum_{i=1}^n y_i. \quad (6.27)$$

The bilateral constraint function for the  $i$ -th rod is given by

$$g_{Si}(\mathbf{q}) = \frac{1}{2}(x_i - a \cos(\varphi + \alpha_i))^2 + \frac{1}{2}(y_i - a \sin(\varphi + \alpha_i))^2 - \frac{1}{2}l^2. \quad (6.28)$$

For the unilateral constraints we use the gap functions

$$\begin{aligned} g_{N1i}(\mathbf{q}) &= x_i \cos(\varphi + \alpha_i) + y_i \sin(\varphi + \alpha_i) - a \\ g_{N2i}(\mathbf{q}) &= -x_i \sin(\varphi + \alpha_i) + y_i \cos(\varphi + \alpha_i) \end{aligned} \quad (6.29)$$

and the unilateral contact law described in Section 3.7.2. A Newton impact law with a coefficient of restitution  $\varepsilon$  is used to describe the impacts. With the partial derivative

$$\begin{aligned} \frac{\partial g_{Si}}{\partial \varphi} &= ax_i \sin(\varphi + \alpha_i) - ay_i \cos(\varphi + \alpha_i) \\ \frac{\partial g_{Si}}{\partial x_i} &= x_i - a \cos(\varphi + \alpha_i), \quad \frac{\partial g_{Si}}{\partial x_j} = 0, \quad i \neq j \\ \frac{\partial g_{Si}}{\partial y_i} &= y_i - a \sin(\varphi + \alpha_i), \quad \frac{\partial g_{Si}}{\partial y_j} = 0, \quad i \neq j \end{aligned} \quad (6.30)$$

of the constraint function  $\mathbf{g}_S$  we can formulate the constraint function

$$\begin{aligned} \gamma_{Si}(\mathbf{u}, \mathbf{q}) &= \frac{\partial g_{Si}}{\partial \mathbf{q}} \mathbf{u} = a\dot{\varphi}(x_i \sin(\varphi + \alpha_i) - y_i \cos(\varphi + \alpha_i)) \\ &\quad + \dot{x}_i(x_i - a \cos(\varphi + \alpha_i)) \\ &\quad + \dot{y}_i(y_i - a \sin(\varphi + \alpha_i)) \end{aligned} \quad (6.31)$$

of the kinematic bilateral constraints associated with  $\mathbf{g}_S$ . The complete equations of motion have the same structure as (6.16) and (6.17) from the double pendulum example. For the numerical solution of the equations of motion, we will use again the consistent two-step integrator from Section 5.2.4. The partial derivative of the velocity level constraint function

$$\begin{aligned} \frac{\partial \gamma_{Si}}{\partial \varphi} &= a(\dot{\varphi}x_i - \dot{y}_i) \cos(\varphi + \alpha_i) + a(\dot{\varphi}y_i - \dot{x}_i) \sin(\varphi + \alpha_i) \\ \frac{\partial \gamma_{Si}}{\partial x_i} &= a\dot{\varphi} \sin(\varphi + \alpha_i) + \dot{x}_i, \quad \frac{\partial \gamma_{Si}}{\partial x_j} = 0, \quad i \neq j \\ \frac{\partial \gamma_{Si}}{\partial y_i} &= -a\dot{\varphi} \cos(\varphi + \alpha_i) + \dot{y}_i, \quad \frac{\partial \gamma_{Si}}{\partial y_j} = 0, \quad i \neq j \end{aligned} \quad (6.32)$$

and the generalized force directions

$$\begin{aligned}
\frac{\partial g_{N1i}}{\partial \varphi} &= -x_i \sin(\varphi + \alpha_i) + y_i \cos(\varphi + \alpha_i) \\
\frac{\partial g_{N1i}}{\partial x_i} &= \cos(\varphi + \alpha_i), \quad \frac{\partial g_{N1i}}{\partial x_j} = 0, \quad i \neq j \\
\frac{\partial g_{N1i}}{\partial y_i} &= \sin(\varphi + \alpha_i), \quad \frac{\partial g_{N1i}}{\partial y_j} = 0, \quad i \neq j \\
\frac{\partial g_{N2i}}{\partial \varphi} &= -x_i \cos(\varphi + \alpha_i) - y_i \sin(\varphi + \alpha_i) \\
\frac{\partial g_{N2i}}{\partial x_i} &= -\sin(\varphi + \alpha_i), \quad \frac{\partial g_{N2i}}{\partial x_j} = 0, \quad i \neq j \\
\frac{\partial g_{N2i}}{\partial y_i} &= \cos(\varphi + \alpha_i), \quad \frac{\partial g_{N2i}}{\partial y_j} = 0, \quad i \neq j
\end{aligned} \tag{6.33}$$

of the unilateral constraints will be needed for the consistent integrator. The partial derivate of the potential function  $V$  is constant and serves as well as discrete derivative

$$\frac{DV}{D\varphi} = 0, \quad \frac{DV}{Dx_i} = 0, \quad \frac{DV}{Dy_i} = mg. \tag{6.34}$$

The other discrete derivatives are calculated numerically by direct evaluation of (5.19). The parameters listed in Table 6.3 are used for the numerical simulation. The simulation is started at the initial position shown in Figure 6.8 and with an initial velocity given by

$$\mathbf{u} = -2 \text{ rad s}^{-1} (1 \quad -y_1(0) \quad -y_2(0) \quad \dots \quad -y_n(0) \quad x_1(0) \quad x_2(0) \quad \dots \quad x_n(0))^T. \tag{6.35}$$

The initial velocity corresponds to a clockwise rigid body rotation of the system with an angular velocity of  $2 \text{ rad s}^{-1}$ . The numerical integration is performed with a time step size of  $\Delta t = 10^{-4} \text{ s}$  once for absolute inelastic impacts ( $\varepsilon = 0$ ) and once for perfect elastic impacts ( $\varepsilon = 1$ ).

Number of levers	$n = 12$
Lever length	$l = 0.2 \text{ m}$
Lever mass	$m = 0.1 \text{ kg}$
Wheel radius	$a = 0.4 \text{ m}$
Wheel moment of inertia	$\theta = 0.16 \text{ kg m}^2$
Gravitational acceleration	$g = 9.81 \text{ m s}^{-2}$
Coefficient of restitution	$\varepsilon \in \{0, 1\}$

Table 6.3: Parameters of the overbalanced wheel.

The total energy in the overbalanced wheel is constant for perfect elastic impacts. For the inelastic impact law the total energy is decreasing stepwise. Both properties are preserved in the numerical solutions obtained with the consistent integrator as shown in Figure 6.9.

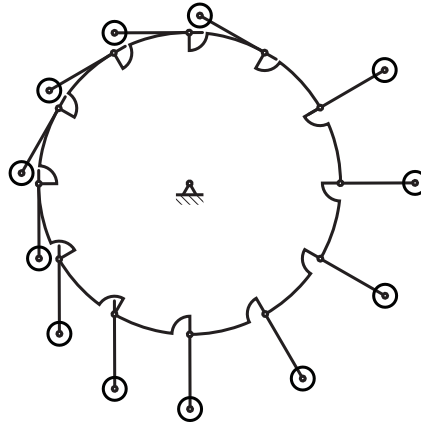


Figure 6.8: Initial position of the overbalanced wheel.

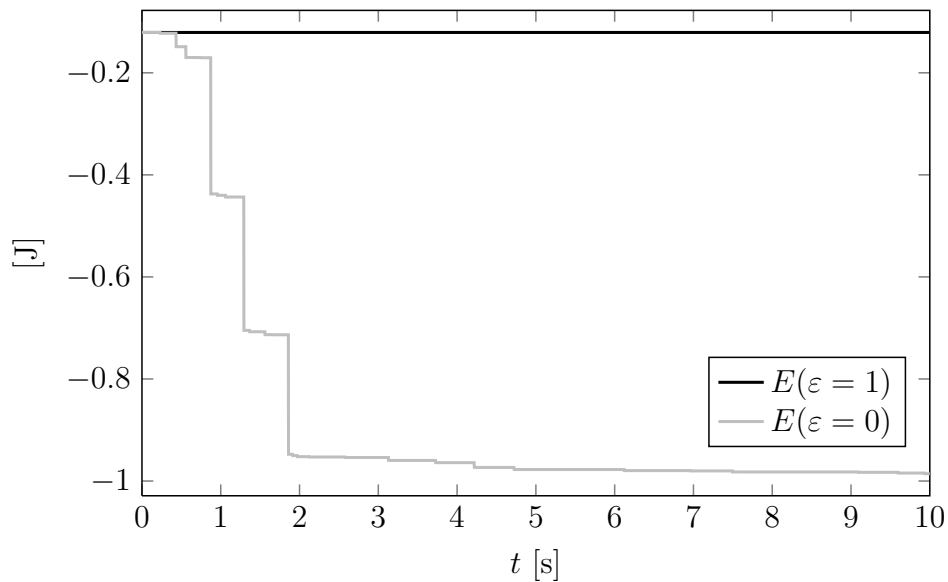
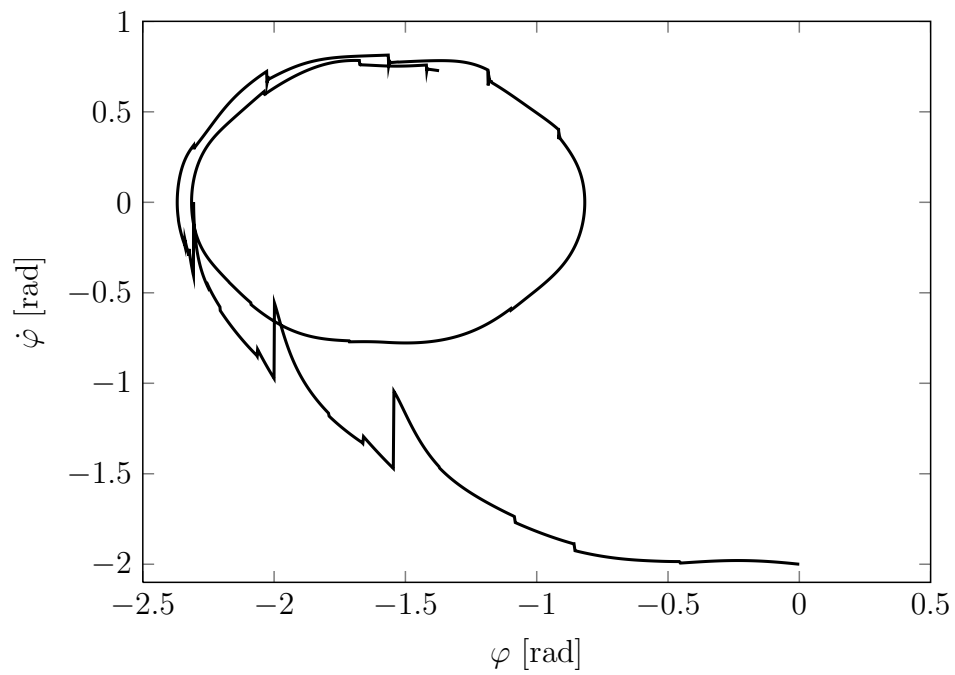
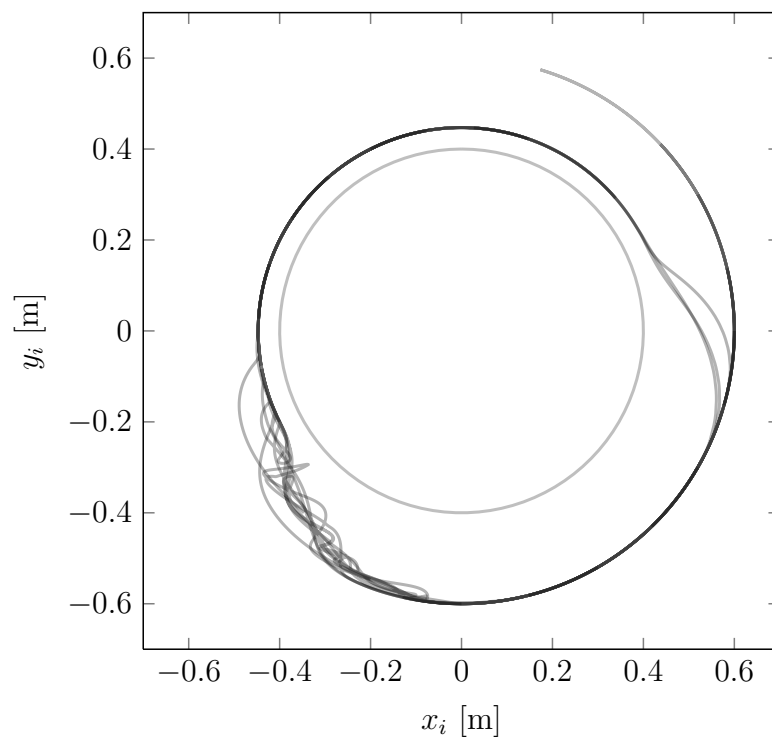
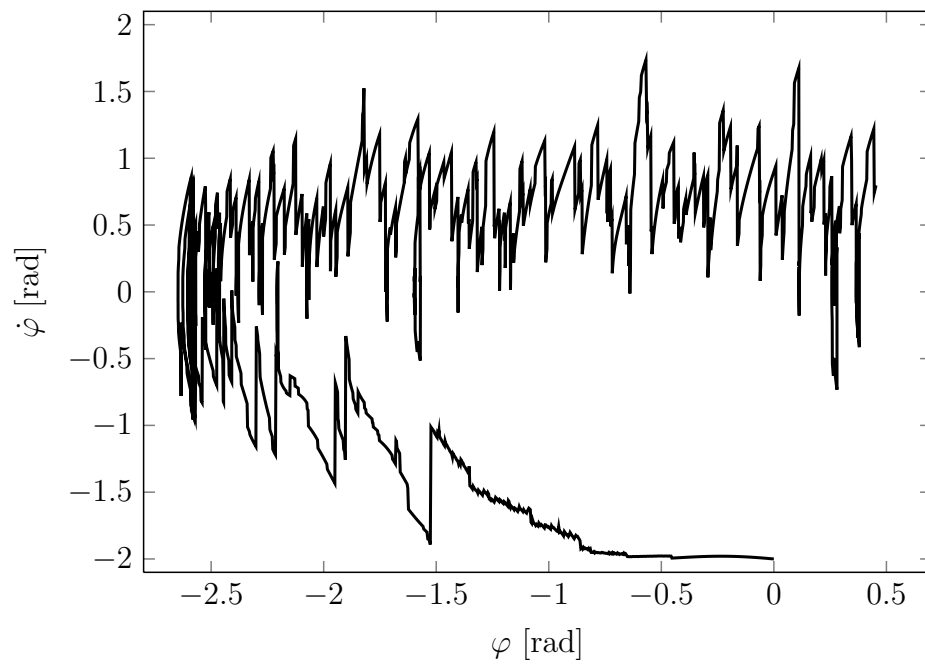
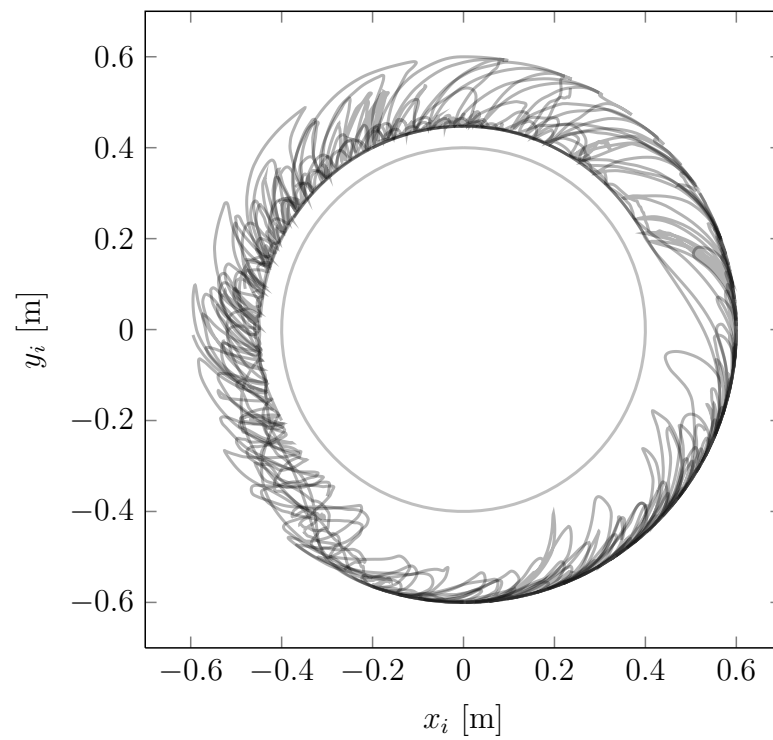


Figure 6.9: Total energy.

In Figure 6.10, the first 10s of the phase space plot for the wheel angle is shown for an absolute inelastic impact law. The large vertical jumps correspond to the impacts when a lever drops from its tangential to its perpendicular position. Energy is dissipated in every impact, which quickly reduces the rotation into a pendular motion. Once the system is in the pendular motion and no further levers drop on the right side, all the dissipation comes from small impacts on the left side of the wheel. In Figure 6.11, the traces of all mass points and an outline of the wheel are shown. Simulating the overbalanced wheel with the same parameters but a perfect elastic impact law results in the motion shown in Figure 6.12 and Figure 6.13. Short after starting the wheel, the total energy is distributed by elastic impacts into a confusing motion of all levers. The phase space plot is dominated by many impacts.



Figure 6.10: Wheel angle phase space ( $\varepsilon = 0$ ).Figure 6.11: Mass point traces ( $\varepsilon = 0$ ).

Figure 6.12: Wheel angle phase space ( $\varepsilon = 1$ ).Figure 6.13: Mass point traces ( $\varepsilon = 1$ ).

## 6.4 Woodpecker

The woodpecker toy is a well known non-smooth mechanical system that shows both impact and frictional phenomena. In this section we use the woodpecker model based on linear kinematics as formulated by Glocker [23, 30]. An illustration of the woodpecker toy is shown in Figure 6.14. It consists of a vertical pole, a sleeve with a hole that is slightly larger than the diameter of the pole, a spring and the woodpecker. When correctly started, the woodpecker moves down the pole, performing a pitching motion interrupted by impacts.

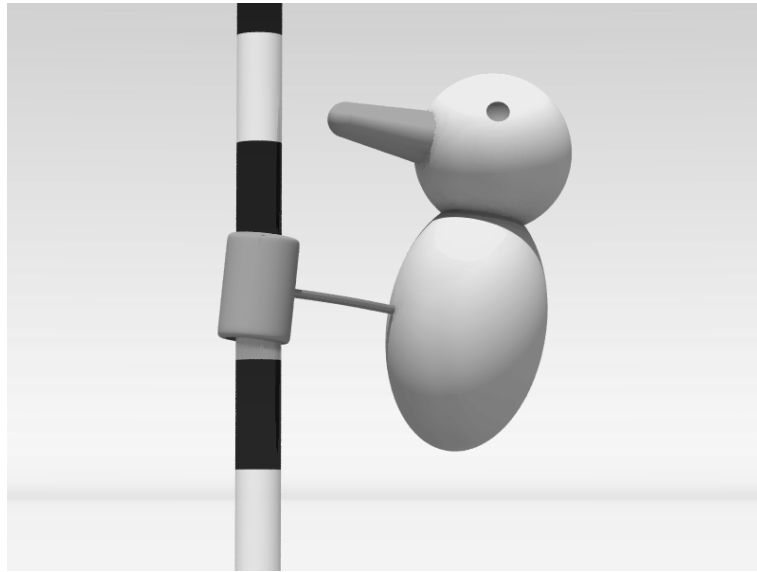


Figure 6.14: Woodpecker toy.

The mechanical model of the woodpecker toy is shown in Figure 6.15. The model has three degrees of freedom: The rotation of the sleeve, the rotation of the woodpecker and the vertical displacement of sleeve and woodpecker. As generalized coordinates the sleeve angle  $\alpha$ , the vertical position  $y$  and the angle  $\varphi$  of the woodpecker are used, i.e.

$$\mathbf{q} = \begin{pmatrix} y \\ \alpha \\ \varphi \end{pmatrix}, \quad \mathbf{u} = \begin{pmatrix} \dot{y} \\ \dot{\alpha} \\ \dot{\varphi} \end{pmatrix}, \quad \mathbf{u} = \dot{\mathbf{q}}. \quad (6.36)$$

The generalized velocities are just the time derivative of the generalized coordinates. The center of mass of the sleeve and the woodpecker are at  $C_M$  and  $C_S$ , respectively. Based on linear kinematics, the kinetic energy of the woodpecker toy can be written as

$$T(\dot{\mathbf{q}}) = \frac{1}{2} [m_M \dot{y}^2 + J_M \dot{\alpha}^2 + m_S (\dot{y} + l_M \dot{\alpha} + l_G \dot{\varphi})^2 + J_S \dot{\varphi}^2] \quad (6.37)$$

where  $m_M$ ,  $J_M$  and  $m_S$ ,  $J_S$  are the mass and moment of inertia of the sleeve and woodpecker, respectively. All geometrical parameters are shown in Figure 6.15. With the

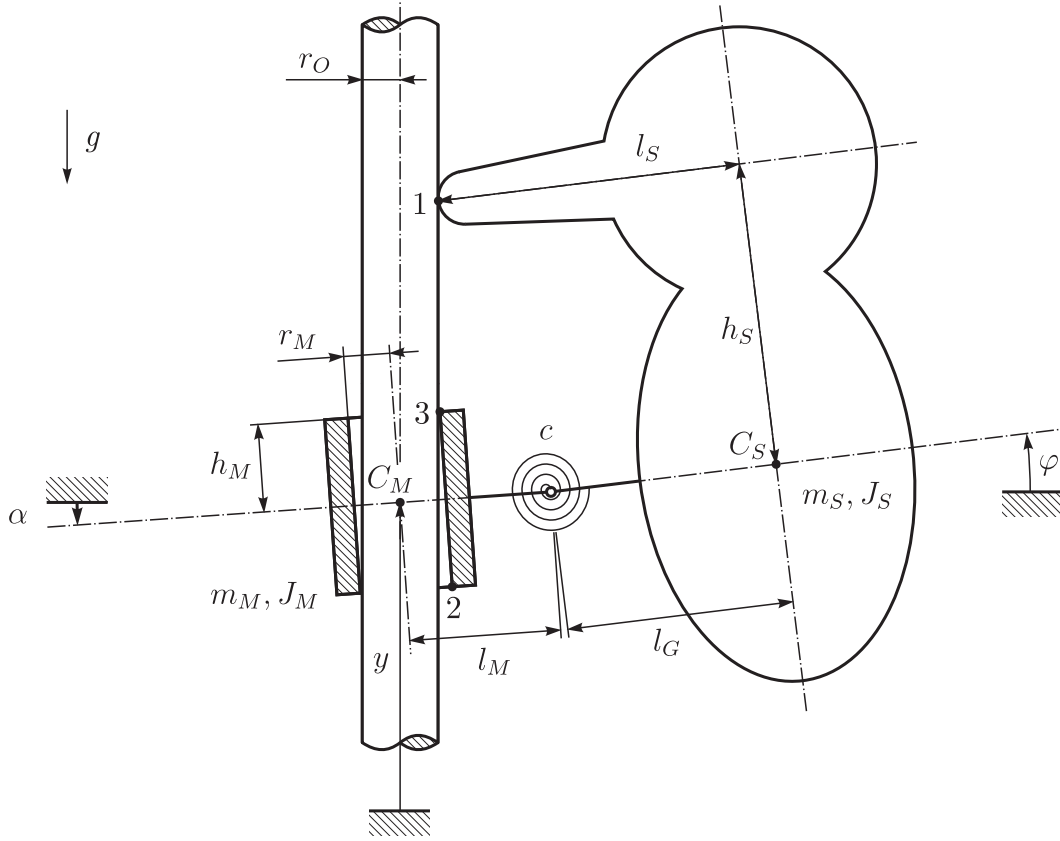


Figure 6.15: Model of the woodpecker toy as described in [23, 30].

constant mass matrix

$$\mathbf{M} = \begin{pmatrix} m_S + m_M & m_S l_M & m_S l_G \\ m_S l_M & J_M + m_S l_M^2 & m_S l_M l_G \\ m_S l_G & m_S l_M l_G & J_S + m_S l_G^2 \end{pmatrix} \quad (6.38)$$

the kinetic energy (6.37) of the woodpecker can be written as

$$T(\dot{\mathbf{q}}) = \frac{1}{2} \dot{\mathbf{q}}^\top \mathbf{M} \dot{\mathbf{q}}. \quad (6.39)$$

The stiffness matrix

$$\mathbf{K} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & c & -c \\ 0 & -c & c \end{pmatrix} \quad (6.40)$$

is introduced for the spring. This allows to formulate the potential function

$$V(\mathbf{q}) = \frac{1}{2} \mathbf{q}^\top \mathbf{K} \mathbf{q} + (g \ 0 \ 0) \mathbf{M} \mathbf{q} \quad (6.41)$$

to describe the gravitational and spring forces. The woodpecker model uses three contacts with friction. The first contact is between the beak and the pole. The second and third

contacts are between the pole and the upper and lower end of the sleeve. The gap functions based on linear kinematics are given by

$$\begin{aligned} g_{N1}(\mathbf{q}) &= (l_M + l_G - l_S - r_O) - h_S\varphi \\ g_{N2}(\mathbf{q}) &= (r_M - r_O) + h_M\alpha \\ g_{N3}(\mathbf{q}) &= (r_M - r_O) - h_M\alpha. \end{aligned} \quad (6.42)$$

The partial derivatives of the gap functions with respect to the generalized coordinates  $\mathbf{q}$  yield the generalized force directions

$$\mathbf{w}_{N1} = \begin{pmatrix} 0 \\ 0 \\ -h_S \end{pmatrix}, \quad \mathbf{w}_{N2} = \begin{pmatrix} 0 \\ h_M \\ 0 \end{pmatrix}, \quad \mathbf{w}_{N3} = \begin{pmatrix} 0 \\ -h_M \\ 0 \end{pmatrix} \quad (6.43)$$

for the normal direction of the contacts. The generalized force directions for the tangential friction forces are given by

$$\mathbf{w}_{T1} = \begin{pmatrix} 1 \\ l_M \\ l_G - l_S \end{pmatrix}, \quad \mathbf{w}_{T2} = \begin{pmatrix} 1 \\ r_M \\ 0 \end{pmatrix}, \quad \mathbf{w}_{T3} = \begin{pmatrix} 1 \\ r_M \\ 0 \end{pmatrix}. \quad (6.44)$$

The equations of motion have the form

$$\begin{aligned} \mathbf{M}\dot{\mathbf{u}} + \left(\frac{\partial V}{\partial \mathbf{q}}\right)^\top - \sum_{i \in \mathcal{I}(\mathbf{q})} (\mathbf{w}_{Ni}\lambda_{Ni} + \mathbf{w}_{Ti}\lambda_{Ti}) &= 0 \\ -\dot{\mathbf{q}} + \mathbf{u} &= 0 \\ i \in \mathcal{I}(\mathbf{q}) : \mathbf{w}_{Ni}^\top \dot{\mathbf{q}} &\in \mathcal{N}_{\mathbb{R}_0^-}(-\lambda_{Ni}) \\ i \in \mathcal{I}(\mathbf{q}) : \mathbf{w}_{Ti}^\top \dot{\mathbf{q}} &\in \mathcal{N}_{[-\mu_i\lambda_{Ni}, \mu_i\lambda_{Ni}]}(-\lambda_{Ti}) \end{aligned} \quad (6.45)$$

where the set of active contacts is defined as

$$\mathcal{I}(\mathbf{q}) = \{i \mid g_{Ni}(\mathbf{q}) = 0\}. \quad (6.46)$$

For the contact friction, a Coulomb friction law formulated as normal cone inclusion is used (see 3.7.3). The impact equations and the impact laws of Newton type in inclusion form are given by

$$\begin{aligned} \mathbf{M}(\mathbf{u}^+ - \mathbf{u}^-) - \sum_{i \in \mathcal{I}(\mathbf{q})} (\mathbf{w}_{Ni}\Lambda_{Ni} + \mathbf{w}_{Ti}\Lambda_{Ti}) &= 0 \\ i \in \mathcal{I}(\mathbf{q}) : \mathbf{w}_{Ni}^\top (\mathbf{u}^+ + \varepsilon_{Ni}\mathbf{u}^-) &\in \mathcal{N}_{\mathbb{R}_0^-}(-\Lambda_{Ni}) \\ i \in \mathcal{I}(\mathbf{q}) : \mathbf{w}_{Ti}^\top (\mathbf{u}^+ + \varepsilon_{Ti}\mathbf{u}^-) &\in \mathcal{N}_{[-\mu_i\Lambda_{Ni}, \mu_i\Lambda_{Ni}]}(-\Lambda_{Ti}). \end{aligned} \quad (6.47)$$

The description (6.45) and (6.47) of the dynamics of the woodpecker toy has the form of (5.1) and (5.2) used for Moreau's midpoint rule as well as the form of (5.104) and (5.105) used for the consistent two-step integrator from Section 5.2.4. Since all the generalized

force directions are constant and there are no bilateral constraints, only the discrete derivative

$$\left(\frac{DV}{D\mathbf{q}}(\mathbf{a}, \mathbf{b})\right)^{\top} = \frac{1}{2}\mathbf{K}(\mathbf{a} + \mathbf{b}) + \mathbf{M}(g \ 0 \ 0)^{\top} \quad (6.48)$$

of the potential function  $V$  is needed additionally for the numerical integration. Note that this system has no additional bilateral constraints and all the potential forces are linear in the generalized coordinates  $\mathbf{q}$ . In this case both inclusion problems obtained for the consistent two-step integrator consist only of normal cone inclusions coupled with linear equations and can be solved with the projected Jacobi iteration or the projected Gauss-Seidel iteration discussed in Section 2.5. Using a projected Jacobi or Gauss-Seidel iteration instead of a projected Newton iteration for the second step of the integrator can greatly reduce the numerical work required for a time step.

For the numerical simulation the parameters and initial conditions in Table 6.4 taken from [30] are used.

Pole radius	$r_0 = 0.0025 \text{ m}$
Inner sleeve radius	$r_M = 0.0031 \text{ m}$
1/2 sleeve height	$h_M = 0.0058 \text{ m}$
Sleeve-Spring distance	$l_M = 0.010 \text{ m}$
Spring-Woodpecker distance	$l_G = 0.015 \text{ m}$
Woodpecker height	$h_S = 0.02 \text{ m}$
Beak length	$l_S = 0.0201 \text{ m}$
Mass, sleeve	$m_M = 0.0003 \text{ kg}$
Mass, woodpecker	$m_S = 0.0045 \text{ kg}$
Moment of inertia, sleeve	$J_M = 5 \cdot 10^{-9} \text{ kg m}^2$
Moment of inertia, woodpecker	$J_S = 7 \cdot 10^{-7} \text{ kg m}^2$
Spring stiffness	$c = 0.0056 \text{ N m rad}^{-1}$
Gravitational acceleration	$g = 9.81 \text{ m s}^{-2}$
Restitution coefficients, normal	$\varepsilon_{N1} = 0.5$ $\varepsilon_{N2} = 0$ $\varepsilon_{N3} = 0$
Restitution coefficients, tangential	$\varepsilon_{T1,2,3} = 0$
Friction coefficients	$\mu_{1,2,3} = 0.3$
Initial coordinates	$y(0) = 0 \text{ m}$ $\alpha(0) = -0.1036 \text{ rad}$ $\varphi(0) = -0.2788 \text{ rad}$
Initial velocities	$\dot{y}(0) = -0.3411 \text{ m s}^{-1}$ $\dot{\alpha}(0) = 0 \text{ rad s}^{-1}$ $\dot{\varphi}(0) = -7.4583 \text{ rad s}^{-1}$

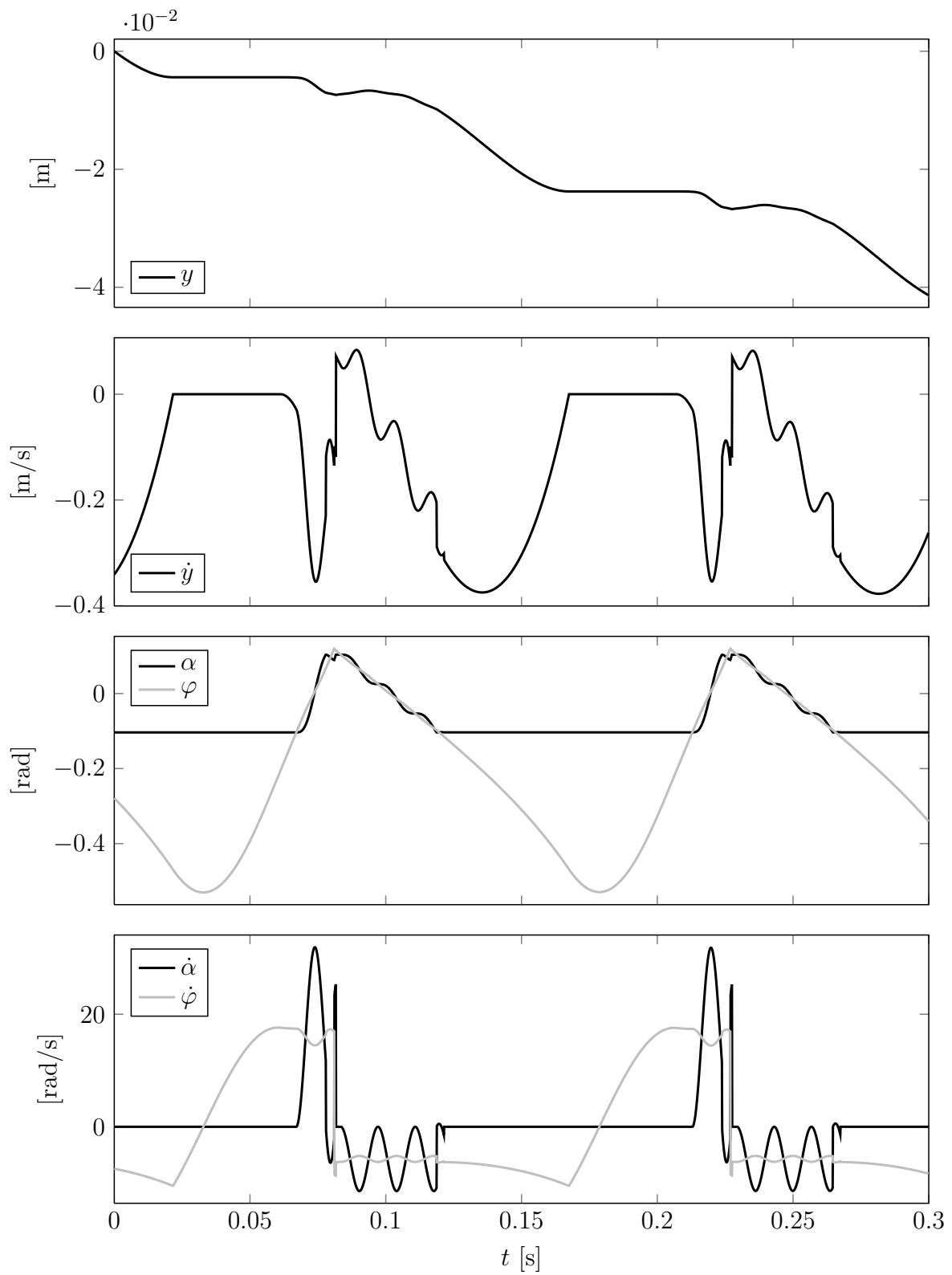
Table 6.4: Parameters and initial conditions.

In Figure 6.16, the solution obtained with the consistent two-step integrator from Section 5.2.4 and a time step  $\Delta t = 10^{-5}$  s is shown. The solution obtained with Moreau's midpoint rule and the same time step size is identical up to the precision of the shown figure. A phase space plot of the two angles  $\alpha$  and  $\varphi$  is shown in Figure 6.17 and Figure 6.18.

A solution obtained with the consistent two-step integrator for a time step  $\Delta t = 10^{-3}$  s is shown in Figure 6.19. Besides the fact that the solution starts to deviate from the exact solution visibly at  $t = 0.12$  s, there is also a high frequency oscillation of the vertical velocity  $\dot{y}$  in the time window from 0.025 s to 0.06 s. This time window is the part of the woodpecker dynamics where the sleeve is sticking. The reason for the oscillation is the discretization of the friction law in the second step (5.107) of the consistent integrator. In (5.107) the discretization is based on the difference  $\mathbf{w}_{T_i}^T(\mathbf{q}^E - \mathbf{q}^B)$  instead of using  $\mathbf{w}_{T_i}^T(\mathbf{u}^E + \varepsilon_i \mathbf{u}^B)$  like in Moreau's midpoint rule. When the friction law is in sticking mode, then the  $\mathbf{w}_{T_i}^T(\mathbf{q}^E - \mathbf{q}^B)$  discretization is dissipation free, leading to an oscillation on velocity level. The oscillation occurs, because the perfect sticking constraint is enforced only on displacement level. This is the same behavior one can observe when enforcing bilateral constraints on displacement level only. Moreau's midpoint rule profits in this case from the combined discretization of the force and impact law. The tangential coefficient of restitution which is equal to zero in this case, dissipates the remaining energy when friction is entering the sticking mode, although this happens without impact in the exact solution. The  $\mathbf{w}_{T_i}^T(\mathbf{q}^E - \mathbf{q}^B)$  discretization approach is used in the consistent two-step integrator to maintain the consistency of the total energy. The oscillation problem could be solved (in this case) by weakening the consistency requirement for the consistent integrator. One could enforce the total energy consistency for example only for the inertial and potential forces, and use a discretization of the normal cone inclusion force laws similar to Moreau's midpoint rule. This would still yield exact energetic consistency at least whenever all the forces of the set-valued force laws of normal cone type are zero (see equation (5.111)).

In Figure 6.20, a part of the total energy curves are shown as comparison for the consistent two-step integrator and Moreau's midpoint rule. For both integrators a time step of  $\Delta t = 10^{-3}$  s has been used. The considered time window shows the sticking of the sleeve and the start of the sliding. As expected, the total energy consistency of the consistent two-step integrator is fulfilled, while the integrator based on Moreau's midpoint rule shows also a slightly increasing total energy at times.

Overall there is no real advantage in using a consistent integrator for the numerical integration of the dynamics of the woodpecker toy. The main source of inaccuracy are badly approximated impact times. When small enough time step sizes are used to localize the impacts good enough in time, then also the small energy inaccuracy of Moreau's midpoint rule is no longer a problem.

Figure 6.16: Woodpecker solution with  $\Delta t = 10^{-5}$  s.



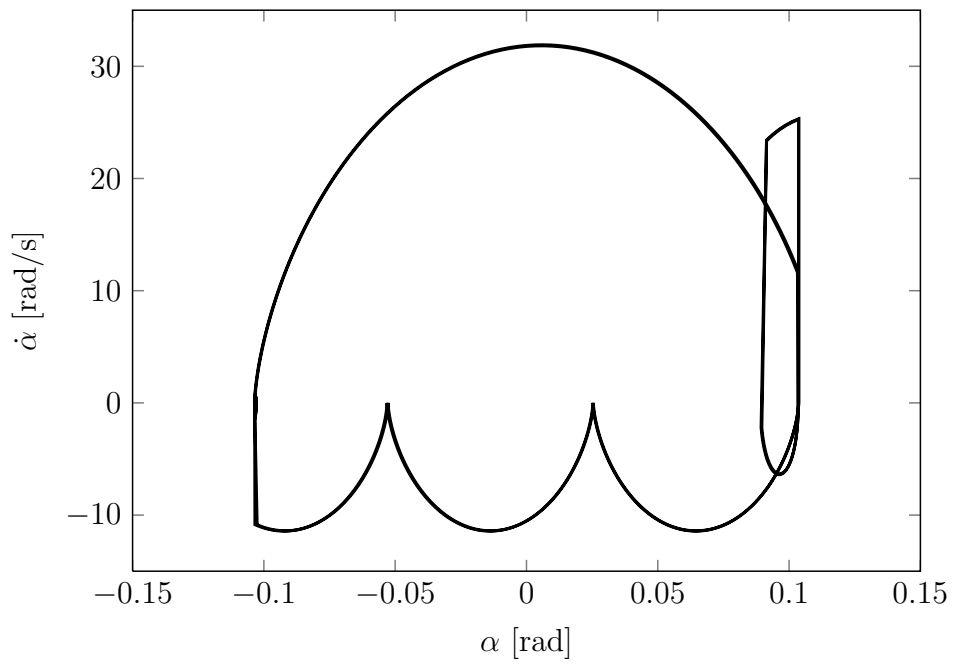


Figure 6.17: Sleeve phase space.

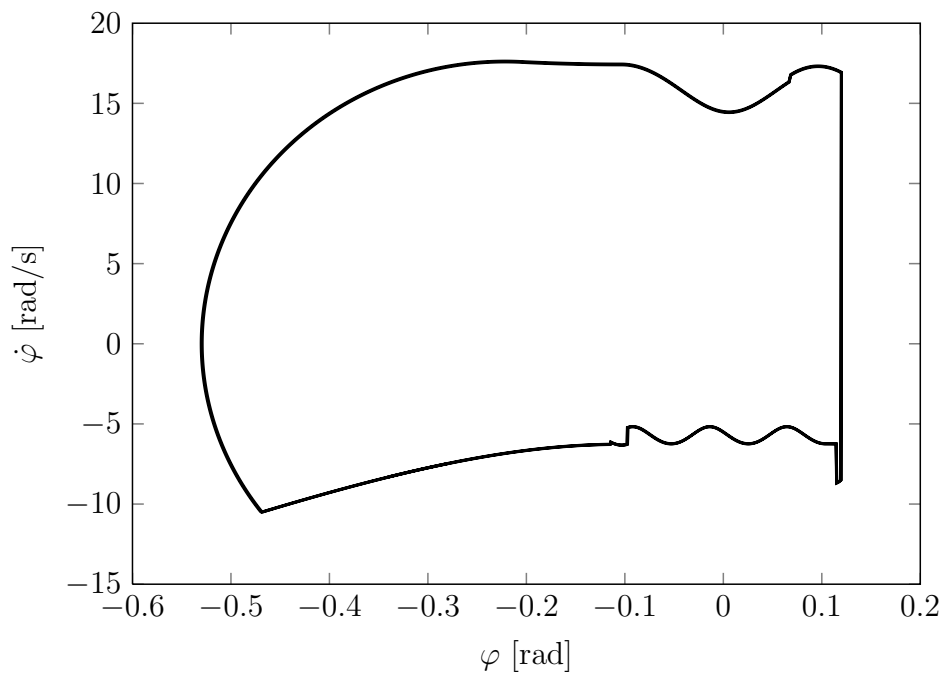


Figure 6.18: Woodpecker phase space.

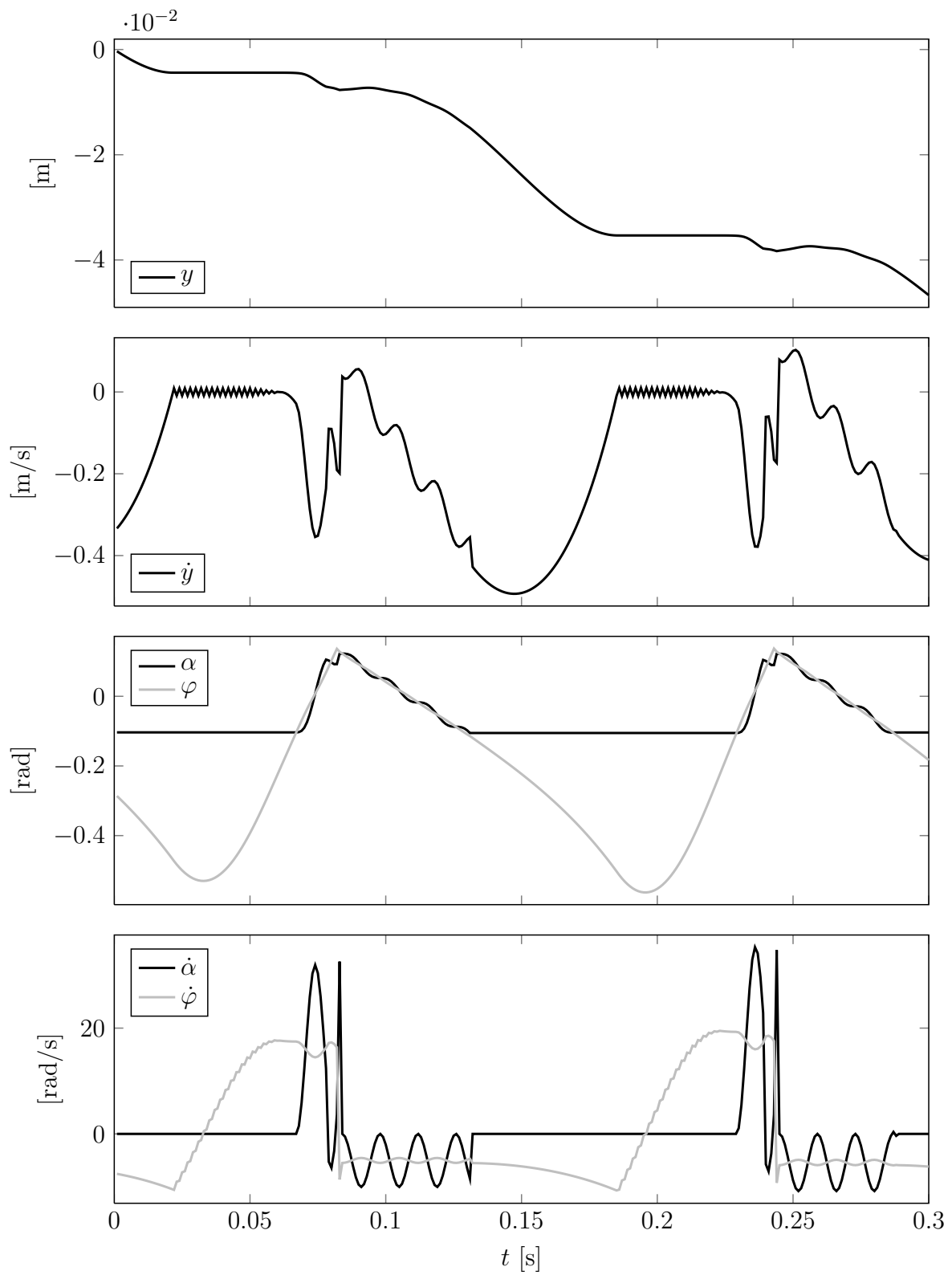


Figure 6.19: Woodpecker solution of the consistent two-step integrator with  $\Delta t = 10^{-3}$ .

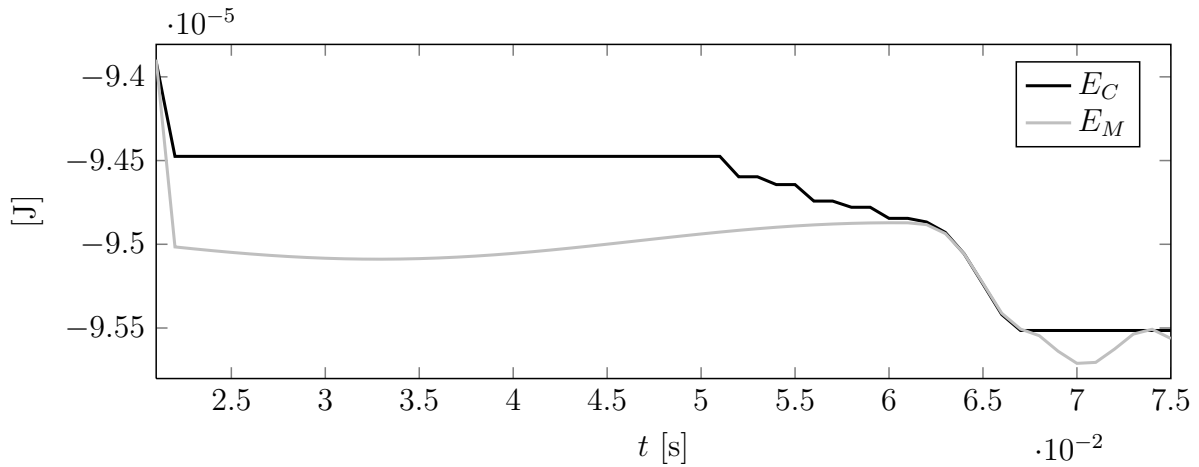


Figure 6.20: Woodpecker energy.

## 6.5 Tippe Top

The tippe top is a special kind of spinning top, which turns from the trivial position to an inverted position if the spinning velocity is high enough. Once the spinning velocity has decreased due to dissipation it tumbles over back to the trivial position. An illustration of a tippe top is shown in Figure 6.21. An analysis of the dynamics and stability of the tippe top can be found in [47]. In this section, the mechanical model of a tippe top will be used as an example application for the Coulomb-Contensou friction approximation discussed in Section 3.7.5. The normal cone inclusion problems occurring in the numerical integration are solved with the methods discussed in Section 2.5.2 and Section 2.5.3. Beside this, the quaternion based consistent integrator formulation from Section 5.2.5 is applied to the example of the tippe top.

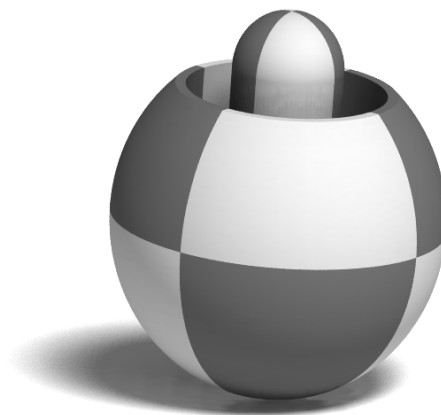


Figure 6.21: Tippe top.

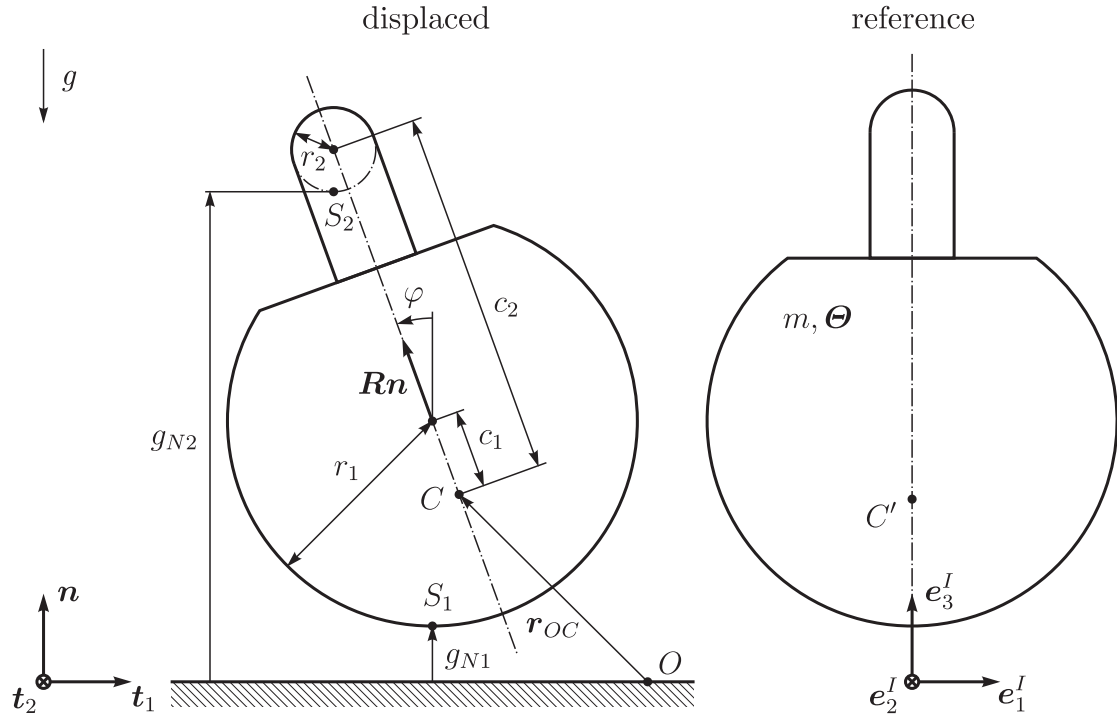


Figure 6.22: Tippe top.

The mechanical model of the tippe top is shown in Figure 6.22. The tippe top consists of one rigid body with center of mass at point  $C$ . The rigid body dynamics is described with the quaternion based ODE formulation from Section 4.3. The position  $\mathbf{r}_{OC}$  of the center of mass and a unit quaternion  $A = (a_0, \mathbf{a})$  form the generalized coordinates

$$\mathbf{q} = \begin{pmatrix} \mathbf{r}_{OC} \\ \psi(A) \end{pmatrix}. \quad (6.49)$$

The unit quaternion describes the rotation of the body from the reference configuration into the displaced configuration. The corresponding rotation matrix is given by

$$\mathbf{R} = \mathbf{a}\mathbf{a}^\top + 2a_0\tilde{\mathbf{a}} + a_0^2\mathbf{I} + \tilde{\mathbf{a}}\tilde{\mathbf{a}} \quad (6.50)$$

as follows directly from (4.57) for a unit quaternion. The generalized velocities are formed by the velocity  $\mathbf{v}_C$  of the center of mass and the vector  $\boldsymbol{\omega}$ , i.e.

$$\mathbf{u} = \begin{pmatrix} \mathbf{v}_C \\ \boldsymbol{\omega} \end{pmatrix}. \quad (6.51)$$

The vector  $\boldsymbol{\omega}$  is the angular velocity  $\boldsymbol{\Omega}$  of the body rotated from the displaced configuration back to the reference configuration

$$\boldsymbol{\omega} = \mathbf{R}^\top \boldsymbol{\Omega}. \quad (6.52)$$

The kinematic relation from the quaternion based rigid body ODE (4.119) can be expressed as

$$\dot{\mathbf{q}} = \mathbf{F}(\mathbf{q})\mathbf{u} \quad (6.53)$$

where the matrix  $\mathbf{F}(\mathbf{q})$  is given by

$$\mathbf{F}(\mathbf{q}) = \begin{pmatrix} \mathbf{I} & 0 \\ 0 & -\frac{1}{2}\mathbf{a}^\top \\ 0 & \frac{a_0}{2}\mathbf{I} + \frac{1}{2}\tilde{\mathbf{a}} \end{pmatrix}. \quad (6.54)$$

The mass and gyro matrix

$$\mathbf{M} = \begin{pmatrix} m\mathbf{I} & 0 \\ 0 & \Theta \end{pmatrix}, \quad \mathbf{H}(\mathbf{u}) = \begin{pmatrix} 0 & 0 \\ 0 & \tilde{\omega}\Theta + \Theta\tilde{\omega} \end{pmatrix}. \quad (6.55)$$

are defined as in Section 5.2.5. In the coordinate system  $\{\mathbf{e}_1^I, \mathbf{e}_2^I, \mathbf{e}_3^I\}$ , the inertia tensor is given by

$${}_I\Theta = \begin{pmatrix} J_1 & 0 & 0 \\ 0 & J_1 & 0 \\ 0 & 0 & J_3 \end{pmatrix}. \quad (6.56)$$

Gravitation is the only potential force in this system. With the normal vector  $\mathbf{n}$  we get the potential function

$$V(\mathbf{q}) = \mathbf{n}^\top \mathbf{r}_{OC} m g. \quad (6.57)$$

The contacts are modeled as unilateral contacts with Coulomb-Contensou friction. As contact geometry, the rigid floor and two spheres with radii  $r_1$  and  $r_2$  are used. To formulate the gap functions for the contacts, one needs the vector

$$\mathbf{r}_{CS_i} = c_i \mathbf{R}\mathbf{n} - r_i \mathbf{n} \quad (6.58)$$

connecting the center of mass with the potential contact points  $S_i$ . The length  $c_i$  is the distance of the center of the sphere from the center of mass (see Figure 6.22). The gap functions are then given by

$$\begin{aligned} g_{Ni}(\mathbf{q}) &= \mathbf{n}^\top \mathbf{r}_{OS_i} \\ &= \mathbf{n}^\top \mathbf{r}_{OC} + \mathbf{n}^\top \mathbf{r}_{CS_i} \\ &= \mathbf{n}^\top \mathbf{r}_{OC} + c_i \mathbf{n}^\top \mathbf{R}\mathbf{n} - r_i \underbrace{\mathbf{n}^\top \mathbf{n}}_{=1} \\ &= \mathbf{n}^\top \mathbf{r}_{OC} + c_i (\mathbf{n}^\top \mathbf{a} \mathbf{a}^\top \mathbf{n} + 2a_0 \underbrace{\mathbf{n}^\top \tilde{\mathbf{a}} \mathbf{n}}_{=0} + a_0^2 \underbrace{\mathbf{n}^\top \mathbf{n}}_{=1} + \mathbf{n}^\top \tilde{\mathbf{a}} \tilde{\mathbf{a}} \mathbf{n}) - r_i \\ &= \mathbf{n}^\top \mathbf{r}_{OC} + c_i (a_0^2 + \mathbf{a}^\top (\mathbf{n}\mathbf{n}^\top + \tilde{\mathbf{n}}\tilde{\mathbf{n}}) \mathbf{a}) - r_i. \end{aligned} \quad (6.59)$$

With the partial derivative of the gap function

$$\frac{\partial g_{Ni}}{\partial \mathbf{q}} = \begin{pmatrix} \mathbf{n}^\top & 2c_i a_0 & 2c_i \mathbf{a}^\top (\mathbf{n}\mathbf{n}^\top + \tilde{\mathbf{n}}\tilde{\mathbf{n}}) \end{pmatrix} \quad (6.60)$$

we can formulate the generalized force direction for the normal contact

$$\mathbf{w}_{Ni}(\mathbf{q}) = \mathbf{F}^\top(\mathbf{q}) \left( \frac{\partial g_{Ni}}{\partial \mathbf{q}} \right)^\top \quad (6.61)$$

based on the direction of the generalized velocities  $\mathbf{u}$ . The Coulomb-Contensou friction law requires three generalized force directions, two for the translational friction forces and one for the rotational friction moment (see Section 3.7.5). The two translational friction forces  $\lambda_{T1}$  and  $\lambda_{T2}$  are introduced in the directions  $\mathbf{t}_1$  and  $\mathbf{t}_2$ . The friction moment  $\lambda_\tau$  into the direction of  $\mathbf{n}$ . The complete Coulomb-Contensou force is then given by

$$\boldsymbol{\lambda}_{Ci} = \begin{pmatrix} \lambda_{T1} \\ \lambda_{T2} \\ \lambda_\tau \end{pmatrix}. \quad (6.62)$$

The associated relative velocity  $\boldsymbol{\gamma}_{Ci}$  is obtained by projecting the velocity at the contact points  $S_i$  into the corresponding direction, i.e.

$$\boldsymbol{\gamma}_{Ci} = \begin{pmatrix} \mathbf{t}_1^\top \mathbf{v}_{S_i} \\ \mathbf{t}_2^\top \mathbf{v}_{S_i} \\ \mathbf{n}^\top \boldsymbol{\Omega} \end{pmatrix} = \mathbf{W}_{Ci}^\top(\mathbf{q})\mathbf{u} \quad (6.63)$$

To determine the matrix of generalized force directions  $\mathbf{W}_{Ci}$  from equation (6.63), we need to express the velocities  $\mathbf{v}_{S_i}$  in term of the generalized velocity  $\mathbf{u}$ . Using the relation (4.76) and the rotational invariance of the cross product we get

$$\begin{aligned} \mathbf{v}_{S_i} &= \mathbf{v}_C + \mathbf{R}\tilde{\boldsymbol{\omega}}(\mathbf{R}^\top \mathbf{r}_{CS_i}) \\ &= \mathbf{v}_C + \underbrace{\mathbf{R}\tilde{\boldsymbol{\omega}}\mathbf{R}^\top}_{(=\mathbf{R}\boldsymbol{\omega})^\sim} \mathbf{r}_{CS_i} \\ &= \mathbf{v}_C - \tilde{\mathbf{r}}_{CS_i}\mathbf{R}\boldsymbol{\omega} \\ &= (\mathbf{I} \quad -\tilde{\mathbf{r}}_{CS_i}\mathbf{R})\mathbf{u}. \end{aligned} \quad (6.64)$$

Inserting (6.64) into (6.63) allows to identify the matrix of generalized force directions

$$\mathbf{W}_{Ci}(\mathbf{q}) = \begin{pmatrix} \mathbf{t}_1 & \mathbf{t}_2 & 0 \\ \mathbf{R}^\top \tilde{\mathbf{r}}_{CS_i} \mathbf{t}_1 & \mathbf{R}^\top \tilde{\mathbf{r}}_{CS_i} \mathbf{t}_2 & \mathbf{R}^\top \mathbf{n} \end{pmatrix} \quad (6.65)$$

of the Coulomb-Contensou force law. The complete equations of motion are then given by

$$\begin{aligned} \mathbf{M}\dot{\mathbf{u}} + \mathbf{H}(\mathbf{u})\mathbf{u} + \mathbf{F}^\top(\mathbf{q}) \left( \frac{\partial V}{\partial \mathbf{q}} \right)^\top - \sum_{i \in \mathcal{I}(\mathbf{q})} (\mathbf{w}_{Ni}(\mathbf{q})\lambda_{Ni} + \mathbf{W}_{Ci}(\mathbf{q})\boldsymbol{\lambda}_{Ci}) &= 0 \\ \dot{\mathbf{q}} &= \mathbf{F}(\mathbf{q})\mathbf{u} \\ i \in \mathcal{I}(\mathbf{q}) : \mathbf{w}_{Ni}(\mathbf{q})^\top \mathbf{u} &\in \mathcal{N}_{\mathbb{R}_0^-}(-\lambda_{Ni}) \\ i \in \mathcal{I}(\mathbf{q}) : \mathbf{W}_{Ci}(\mathbf{q})^\top \mathbf{u} &\in \mathcal{N}_{\mathcal{C}_F(\lambda_{Ni})}(-\boldsymbol{\lambda}_{Ci}). \end{aligned} \quad (6.66)$$

The unilateral constraints are formulated on velocity level using the index set of active contacts

$$\mathcal{I}(\mathbf{q}) = \{i \mid g_{Ni}(\mathbf{q}) = 0\}. \quad (6.67)$$

The convex set  $\mathcal{C}_F(\lambda_{Ni})$  is the force reservoir of the Coulomb-Contensou force law. One can use either the set  $\mathcal{B}_F$  or the ellipsoid based approximation  $\mathcal{E}_F$  as described in Section 3.7.5. The impact equations and impact laws are given by

$$\begin{aligned} \mathbf{M}(\mathbf{u}^+ - \mathbf{u}^-) - \sum_{i \in \mathcal{I}(\mathbf{q})} (\mathbf{w}_{Ni}(\mathbf{q})\Lambda_{Ni} + \mathbf{W}_{Ci}(\mathbf{q})\Lambda_{Ci}) &= 0 \\ i \in \mathcal{I}(\mathbf{q}) : \mathbf{w}_{Ni}(\mathbf{q})^\top (\mathbf{u}^+ + \varepsilon_{Ni}\mathbf{u}^-) &\in \mathcal{N}_{\mathbb{R}_0^-}(-\Lambda_{Ni}) \\ i \in \mathcal{I}(\mathbf{q}) : \mathbf{W}_{Ci}(\mathbf{q})^\top (\mathbf{u}^+ + \varepsilon_{Ci}\mathbf{u}^-) &\in \mathcal{N}_{\mathcal{C}_F(\Lambda_{Ni})}(-\Lambda_{Ci}). \end{aligned} \quad (6.68)$$

For the numerical integration we use the parameters and initial conditions listed in Table 6.5. The parameters have been taken from [47] and are originally based on [21]. The average friction radius  $\bar{R}_{1,2} = 2.945 \cdot 10^{-4}$  m corresponds to a circular contact area with radius  $R = 5 \cdot 10^{-4}$  m and a parabolic normal force distribution as described in [47]. For the quaternion based consistent integrator (5.141) the discrete derivative of the gap

Main sphere distance	$c_1 = 0.003$ m
Handle sphere distance	$c_2 = 0.016$ m
Main sphere radius	$r_1 = 0.015$ m
Handle sphere radius	$r_2 = 0.005$ m
Mass	$m = 0.006$ kg
Moments of inertia	$J_1 = 8 \cdot 10^{-7}$ kg m <sup>2</sup> $J_3 = 7 \cdot 10^{-7}$ kg m <sup>2</sup>
Friction coefficients	$\mu_{1,2} = 0.3$
Average friction radii	$\bar{R}_{1,2} = 2.945 \cdot 10^{-4}$ m
Restitution coefficients	$\varepsilon_{N1,2} = 0$ $\varepsilon_{C1,2} = 0$
Initial position	${}^I\mathbf{r}_{OC}(0) = (0 \ 0 \ 1.2015 \cdot 10^{-2})^\top$ m
Initial angle	$\varphi(0) = 0.1$ rad
Initial velocity	${}^I\mathbf{v}_C(0) = (0 \ 0 \ 0)^\top$ m s <sup>-1</sup>
Initial angular velocity	${}^I\boldsymbol{\omega}(0) = (0 \ 0 \ 180)^\top$ rad s <sup>-1</sup>

Table 6.5: Parameters and initial conditions.

function

$$\frac{Dg_{Ni}}{D\mathbf{q}}(\mathbf{a}, \mathbf{b}) = \frac{\partial g_{Ni}}{\partial \mathbf{q}} \left( \frac{\mathbf{a} + \mathbf{b}}{2} \right) \quad (6.69)$$

and the discrete derivative of the potential energy

$$\left( \frac{DV}{D\mathbf{q}} \right)^\top = \left( \frac{\partial V}{\partial \mathbf{q}} \right)^\top = \begin{pmatrix} \mathbf{n} m g \\ 0 \end{pmatrix} \quad (6.70)$$

are required. Both can be obtained directly from the partial derivative, as both functions are at most quadratic. In Figure 6.23, the inclination angle  $\varphi$  vs. time is shown for the

tippe top with friction laws based on the set  $\mathcal{E}_F$  and  $\mathcal{B}_F$ . Both curves have been obtained with Moreau's midpoint rule and a time step of  $\Delta t = 10^{-5}$  s. Of course, the results start to deviate with time because the system is very sensitive, but still the overall behavior is preserved with the approximation  $\mathcal{E}_F$  very well. The ellipsoid approximation can make use of the transformation to a sphere  $\mathcal{S}_F$  in the proximal point iteration. A MATLAB<sup>®</sup> implementation results in a fixed point iteration for the set  $\mathcal{S}_F$  which is about 10.5 times faster than with  $\mathcal{B}_F$ . Compared to the set  $\mathcal{E}_F$ , the transformed fixed point iteration using the set  $\mathcal{S}_F$  is about 2.7 times faster. It has to be noted, that the solution obtained with the set  $\mathcal{B}_F$  is only more correct with respect to the assumptions made about the contact area and the friction interaction. In general, the geometry of the contact area, the distribution of the normal force in the contact area and the properties of the friction interaction are subject to large uncertainties. In this situation, the level of model-detail of the ellipsoidal set can be more appropriate. The one-step or two-step quaternion based

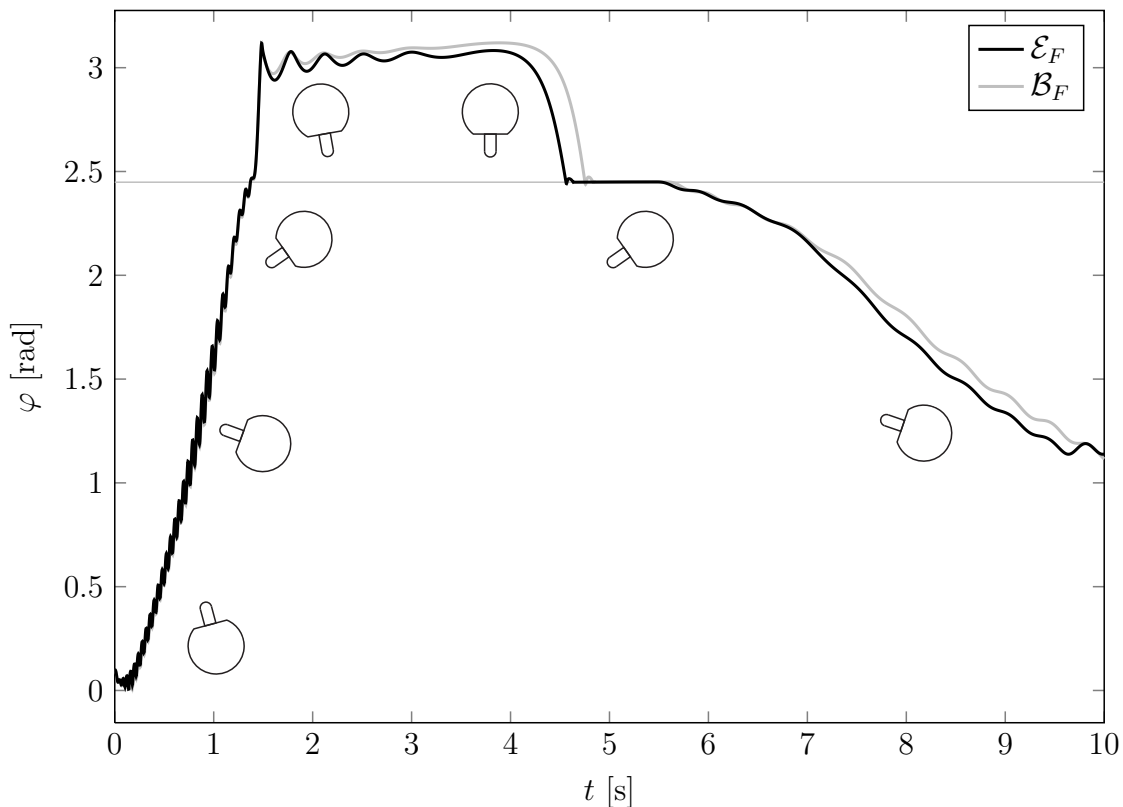


Figure 6.23: Tippe top inclination ( $\Delta t = 10^{-5}$  s).

consistent integrator (5.141) with a time step  $\Delta t = 10^{-5}$  s yields a solution which is identical to the one shown for Moreau's midpoint rule and the set  $\mathcal{E}_F$  in Figure 6.23 up to the precision of the plot. An illustration of the dynamics of the tippe top is shown in Figure 6.24.

Integrating the dynamics of the tippe top with a larger time step leads to drift in the unilateral constraint for the one-step quaternion based consistent integrator or Moreau's



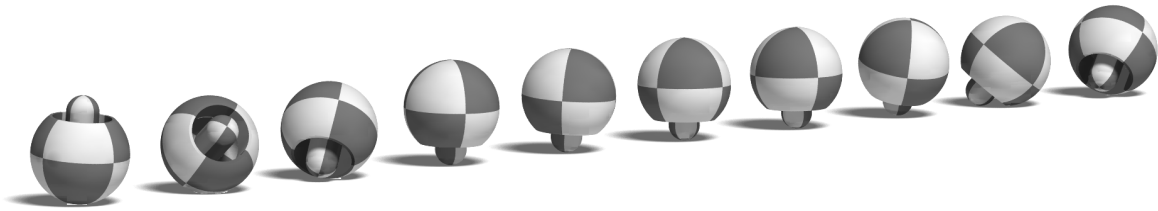
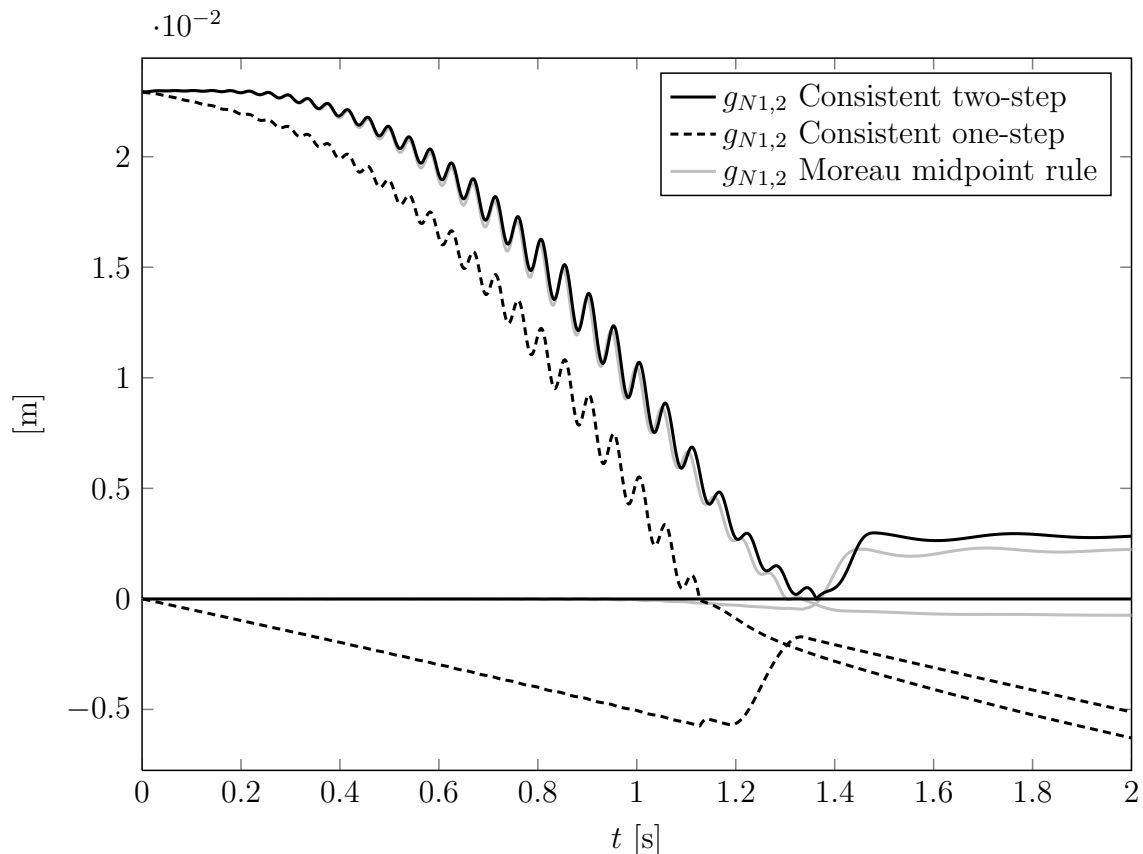


Figure 6.24: Tippe top dynamics.

midpoint rule. The development of the gap functions for the one-step and the two-step quaternion based consistent integrator as well as Moreau's midpoint rule is shown in Figure 6.25 for  $\Delta t = 10^{-3}$  s. As expected, the consistent two-step integrator is the only one that does not show any drift.

Figure 6.25: Tippe top gap functions ( $\Delta t = 10^{-3}$  s).

The drifting of Moreau's midpoint rule could be removed for this example by shifting the tippe top back to the surface of the floor after each integration step. Obviously this is not an option for the consistent one-step integrator, as this would change the potential energy. To improve the situation for the consistent one-step integrator one has to use

either a smaller time step, or a perfect elastic normal contact, i.e.  $\varepsilon_{N_i} = 1$ . This yields a unilateral contact that is consistent on velocity level for the consistent one-step integrator as discussed in Section 5.2.5. Of course changing  $\varepsilon_{N_i}$  modifies the dynamics of the model, at least to some extent.

The tippe top example has been implemented also with the decomposed mass matrix based two-step integrator from Section 5.2.4. The rigid body has been described as scalable body based DAE. For this example, the decomposed mass matrix based two-step integrator yields mostly the same results as the quaternion based two-step integrator. With larger time steps, the overall accuracy of the solution decreases earlier when compared with the quaternion based two-step integrator. At least some of this is caused by the explicit discretization of  $\mathbf{Q}$  and  $\mathbf{H}$  in the integrator from Section 5.2.4. The situation can be improved by switching to an implicit midpoint approximation of  $\mathbf{Q}$  and  $\mathbf{H}$ , but of course this also increases the numerical work slightly. Like the quaternion based two-step integrator, the decomposed mass matrix based two-step integrator does not show any drift in the unilateral contacts.

## Conclusions

---

In this thesis, different variants of consistent integrators for the dynamics of non-smooth mechanical systems have been developed. The integrators rely on the discrete derivative approximation of partial derivatives to achieve consistency with respect to the total energy and to the kinematics of unilateral and bilateral constraints.

In a first part, some approaches for the formulation of consistent integrators for smooth multibody systems have been discussed. This included a consistent integrator based on equations of motion in differential-algebraic form, where the mass matrix is constant. The equations of motion have been extended with the induced constraints on velocity level and corresponding multipliers in order to enforce the kinematics of bilateral constraints on displacement and velocity level. For rigid body dynamics based on the affine body formulation this leads to a system of nonlinear equations with 36 unknowns per body. The number of unknowns could be reduced with another integration scheme where a coordinate dependence of the mass matrix is allowed. Unfortunately, the coordinate dependence of the mass matrix results in complicated expressions for the discretization of the inertial forces which are not efficient to evaluate in general. Beside this, minimal parameterizations of the degrees of freedom of a rigid body with Euler or Kardan angles also have the singularity problem.

The number of unknowns for a rigid body can be reduced from 36 to 16 by formulating the equations of motion in differential-algebraic form using the scalable body with a quaternion based parametrization instead of the affine body based approach. The derivation of the equations of motion for a scalable body and its reduction to a rigid body has been discussed in detail. The mass matrix of scalable body based rigid body formulation is still a function of the generalized coordinates, but it can be decomposed into a constant matrix and a transformation of the generalized velocities. For equations of motion with an analytical mass matrix decomposition, a consistent integrator with support for bilateral constraints has been formulated. In a next step, the decomposed mass matrix based consistent integrator has been extended for perfect unilateral constraints. The extension to the more general set-valued force laws of normal cone type and dissipative impacts has been achieved with the formulation of a consistent two-step integrator. The two-step integrator separates the discretization of the impacts and impact free parts of the dy-

namics, but it is still an event-capturing approach. The consistent integrators based on constant and decomposed mass matrices are not limited to rigid body systems. The last consistent integrator that has been discussed is a quaternion based formulation using a direct discretization of the Newton-Euler equations. This formulation is limited to rigid body systems.

The consistent integrators for non-smooth mechanical systems yield a system of coupled nonlinear equations and normal cone inclusions for each time step, in general. To solve this one-step problem a projected Newton iteration has been proposed. In certain cases one gets only linear equations together with the normal cone inclusions, as for example when the constraint equations and potential forces are linear in the generalized coordinates. In these cases also the more efficient projected Jacobi or projected Gauss-Seidel iterations can be used. The linear transformation of normal cone inclusions and proximal point iterations as well as the recovering strategy for proximal point iterations have been developed in this work. These techniques can be used to improve the efficiency of the iteration methods for solving normal cone inclusions on multidimensional convex sets.

The main contributions of this thesis can be summarized as follows:

- Consistent integrators for non-smooth multibody dynamics have been developed and demonstrated for some example systems. The integrators are based on one- or two-step schemes and allow for inertial and potential forces, bilateral constraints and set-valued force laws of normal cone type. The set-valued force laws can be equipped with impact laws of Newton-type. The integrators are consistent with respect to the total energy, the bilateral constraints on both kinematic levels and the unilateral constraints on velocity level.
- The scalable body model associated with the kinematics of an unconstrained quaternion has been introduced. The equations of motion of a scalable body have been derived in different forms directly from the principle of virtual work without making any rigid body assumptions. A quaternion based rigid body formulation in differential-algebraic form has been derived by adding a mechanical quaternion unit length constraint to the scalable body.
- The projected Newton iteration method has been proposed, which can be used for solving systems of coupled nonlinear equations and normal cone inclusion. The method has been applied successfully to some examples.
- Improved methods for the solution of normal cone inclusions on multidimensional convex sets have been developed. The transformation technique for proximal point iterations can be applied in cases where the linear transformation of the convex set leads to a simplification of the associated proximal point function. The recovering strategy for proximal point iterations can be used to improve the convergence rate for badly scaled multidimensional normal cone inclusions.

The consistent integration schemes presented in this work are robust for large time steps, as is to be expected for a completely implicit and energy consistent integrator. The complete implicit discretization increases the numerical work per time step when compared

with Moreau's midpoint rule which uses only a partially implicit discretization. Moreau's midpoint rule yields a one-step problem consisting of linear equations and normal cone inclusions. For the consistent integrators one has to solve a system of coupled nonlinear equations and normal cone inclusions for each time step. For the consistent two-step integrators there is additionally the impact inclusion problem that has to be solved. Also the discrete derivative leads to a higher complexity of the expressions used in the one-step problem of a consistent integrator in general. It is difficult to gain any overall performance in the integration process with a consistent integrator, unless the system contains very fast dynamics that force integrators like Moreau's midpoint rule to use very small time steps, in order to remain stable.

Moreau's midpoint rule can preserve the total energy for certain special cases as well. This can be seen from a comparison with the consistent integrators, or by deriving the energy balance with the same approach as for consistent integrators. For example, the total energy is preserved in Moreau's midpoint rule for systems with a constant mass matrix, constant potential forces and scleronomic perfect elastic unilateral contacts. Two dimensional granular media models are examples of such systems, although the discretization of the non-smooth forces is usually not consistent with respect to the total energy.

Integrators formulated from equations of motion in differential-algebraic form are well suited for systems with many degrees of freedom and only few constraints. A differential-algebraic equations formulation is required for systems with kinematic loops. For systems where most degrees of freedom initially considered by a set of generalized coordinates, are removed with bilateral constraints, a formulation of the equations of motion in minimal coordinates (with respect to bilateral constraints) can be more efficient in terms of the number of equations. Unfortunately, a minimal coordinates based approach is difficult to realize in an efficient way for consistent integrators.

The consistent integrators developed in this work do not strictly enforce the unilateral contacts on displacement level. Instead, only the increasing of the interpenetration is prevented strictly once the contact is closed. An energy consistent strict enforcing of the unilateral constraint on displacement level is difficult with a fixed time step integrator. The difficulties are related to the fact, that the complementarity between gap function and contact force integral is lost when time intervals are considered during which the contact closes or opens. An event driven integration scheme could make sure that contacts open or close only at boundaries of a time step, but event driven integrators introduce other problems.

To achieve the reduced form of consistency for unilateral contacts discussed in this work, the formulation of a gap function is required. For Moreau's midpoint rule or consistent integrators without the consistency of unilateral contacts, only an index function and a generalized force direction has to be provided. The index function has to determine whether the contact is closed or not on displacement level. In contrast to this, the gap function has to provide a measure how far open the contact is as well.

Not all possible combinations of consistent integrators and non-smooth extensions are discussed in this work. For example, one could extend the consistent integrator based on a general coordinate dependent mass matrix with non-smooth forces as well. For this, one can apply the formulation and discretization of the non-smooth forces that have been

described in detail for the decomposed mass matrix based integrator. Another possibility is to combine the non-consistent one-step discretization of the non-smooth forces from Moreau's midpoint rule with the consistent integration schemes. The consistency of the resulting integrators would be limited to inertial and potential forces as well as bilateral constraints. For time steps where the non-smooth forces are zero, one would still get exact consistency of the total energy. In the same spirit also other elements like kinematic excitation or other forces (for example dissipation) can be added to the model. Some of the robustness properties should still be preserved even when an exact consistency of the total energy is lost or given only when certain forces are zero.

There are still many topics for further research. The problem of momentum consistency in non-smooth consistent integrators has not been addressed in this thesis. Work on additional consistency with respect to momentum and first integrals arising from symmetry could be started with a Hamiltonian approach and generalized momenta as variables or a rigid body based approach parametrized with momentum and spin. The consistency on both kinematic levels for unilateral contacts could be interesting as well, even for integrators without consistency of the total energy. Also convergence of non-smooth consistent integrators presented in this work should be investigated further. The consistency on both kinematic levels of bilateral constraints has been achieved by adding a set of artificial multipliers to the equations of motion. Alternatively, an approach using a projection of the kinematic equation could be further investigated. Other interesting approaches include using a semi-implicit discretization of the smooth forces, where a linear approximation around the beginning of a time step is discretized implicitly. This should enhance the robustness of the integration scheme without increasing the numerical costs per time step too much, since the resulting one-step problem would be still a system of coupled linear equations and normal cone inclusions. The two-step integrators discussed in this work or non-smooth integrators based on a time-splitting method could be subject for further research as well. To increase the performance of non-smooth integrators in general, one could try to reduce the coupling horizon for the contacts in a numerical approximation, for example by employing time sequential impact laws or inaccurate solutions of the one-step inclusion problem (while still keeping the scheme convergent). Beside this, any improvement for the solution of normal cone inclusion problems with linear or nonlinear equations could be valuable for non-smooth integrators.

# Bibliography

---

- [1] ACARY, V., BONNEFON, O., AND BROGLIATO, B. Time-stepping numerical simulation of switched circuits within the nonsmooth dynamical systems approach. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on* 29, 7 (2010), 1042–1055.
- [2] ACARY, V., AND BROGLIATO, B. Numerical time integration of higher order dynamical systems with state constraints. In *ENOC-2005, Eindhoven, Netherlands, 7-12 August (2005)*.
- [3] ACARY, V., AND BROGLIATO, B. *Numerical Methods for Nonsmooth Dynamical Systems: Applications in Mechanics and Electronics*, vol. 35 of *Lecture Notes in Applied and Computational Mechanics*. Springer, Berlin, Heidelberg, 2008.
- [4] ACARY, V., BROGLIATO, B., AND GOELEVELN, D. Higher order Moreau’s sweeping process: Mathematical formulation and numerical simulation. *Mathematical Programming* 113 (2008), 133–217.
- [5] AEBERHARD, U. *Geometrische Behandlung idealer Stösse*. PhD thesis, ETH Zürich, 2008.
- [6] ALART, P., AND CURNIER, A. A mixed formulation for frictional contact problems prone to Newton like solution methods. *Computer Methods in Applied Mechanics and Engineering* 92, 3 (1991), 353–375.
- [7] ALTMANN, S. L. *Rotations, Quaternions and Double Groups*. Clarendon Press, Oxford, 1986.
- [8] ANITESCU, M., POTRA, F., AND STEWART, D. Time-stepping for three-dimensional rigid body dynamics. *Computer Methods in Applied Mechanics and Engineering* 177, 3 (1999), 183–197.
- [9] BARAFF, D. Fast contact force computation for nonpenetrating rigid bodies. *Computer Graphics Proceedings, Annual Conference Series 23-24* (1994).
- [10] BETSCH, P. The discrete null space method for the energy consistent integration of constrained mechanical systems. Part I: Holonomic constraints. *Computer Methods in Applied Mechanics and Engineering* 194, 50-52 (2005), 5159–5190.

- [11] BETSCH, P., AND LEYENDECKER, S. The discrete null space method for the energy consistent integration of constrained mechanical systems. Part II: Multibody dynamics. *International Journal for Numerical Methods in Engineering* 67 (2006), 499–552.
- [12] BETSCH, P., AND SIEBERT, R. Rigid body dynamics in terms of quaternions: Hamiltonian formulation and conserving numerical integration. *International Journal for Numerical Methods in Engineering* 79 (2009), 444–473.
- [13] BETSCH, P., AND STEINMANN, P. Constrained integration of rigid body dynamics. *Computer Methods in Applied Mechanics and Engineering* 191, 3-5 (2001), 467–488.
- [14] BETSCH, P., AND STEINMANN, P. Conservation properties of a time FE method - part III: Mechanical systems with holonomic constraints. *International Journal for Numerical Methods in Engineering* 53 (2002), 2271–2304.
- [15] BREMER, H. *Dynamik und Regelung mechanischer Systeme*. Teubner, Stuttgart, 1988.
- [16] BROGLIATO, B., TEN DAM, A. A., PAOLI, L., GÉNOT, F., AND ABADIE, M. Numerical simulation of finite dimensional multibody nonsmooth mechanical systems. *ASME Applied Mechanics Reviews* 55 (2002), 107–150.
- [17] CATALDI-SPINOLA, E. *Curve Squealing Mechanism of Railway Vehicles*. PhD thesis, ETH Zürich, 2007.
- [18] COTTLE, R. W., AND DANTZIG, G. B. Complementary pivot theory of mathematical programming. *Linear Algebra and Its Applications* 1 (1968), 103–125.
- [19] COTTLE, R. W., PANG, J.-S., AND STONE, R. E. *The Linear Complementarity Problem*. Academic Press, London, 1992.
- [20] DE SAXCÉ, G., AND FENG, Z.-Q. The bipotential method: A constructive approach to design the complete contact law with friction and improved numerical algorithms. *Mathematical and Computer Modelling* 28 (1998), 225–245.
- [21] FRIEDL, C. Der Stehaufkreisel. Universität Augsburg, 1997. Master’s thesis.
- [22] GEAR, C. W., LEIMKUHNER, B., AND GUPTA, G. K. Automatic integration of Euler-Lagrange equations with constraints. *Journal of Computational and Applied Mathematics* 12-13 (1985), 77–90.
- [23] GLOCKER, CH. *Dynamik von Starrkörpersystemen mit Reibung und Stößen*, vol. 18, no. 182 of *VDI-Fortschrittberichte Mechanik/Bruchmechanik*. VDI-Verlag, Düsseldorf, 1995.
- [24] GLOCKER, CH. Formulation of spatial contact situations in rigid multibody systems. *Computer Methods in Applied Mechanics and Engineering* 177 (1999), 199–214.



- [25] GLOCKER, CH. *Set-Valued Force Laws: Dynamics of Non-Smooth Systems*, vol. 1 of *Lecture Notes in Applied Mechanics*. Springer, Berlin, Heidelberg, 2001.
- [26] GLOCKER, CH. Models of non-smooth switches in electrical systems. *Int. J. Circuit Theory Appl.* 33 (2005), 205–234.
- [27] GLOCKER, CH. An introduction to impacts. In *Nonsmooth Mechanics of Solids*, J. Haslinger and G. Stavroulakis, Eds., vol. 485 of *CISM Courses and Lectures*. Springer, Wien, New York, 2006, pp. 45–102.
- [28] GLOCKER, CH. Reduction techniques for distributed set-valued force laws. In *Proceedings of the 2nd International Conference on Nonsmooth/Nonconvex Mechanics with Applications in Engineering* (Thessaloniki, Greece, July 2006), C. C. Baniotopoulos, Ed., Editions Ziti, Thessaloniki, pp. 173–180.
- [29] GLOCKER, CH. Simulation von harten Kontakten mit Reibung: Eine iterative Projektionsmethode. In *Schwingungen in Antrieben 2006* (Fulda, 2006), VDI-Berichte Nr. 1968, VDI-Verlag, Düsseldorf, pp. 19–44.
- [30] GLOCKER, CH., AND STUDER, C. Formulation and preparation for numerical evaluation of linear complementarity systems in dynamics. *Multibody System Dynamics* 13 (2005), 447–463.
- [31] GONZALEZ, O. *Design and analysis of conserving integrators for nonlinear Hamiltonian systems with symmetry*. PhD thesis, Stanford University, 1996.
- [32] GONZALEZ, O. Time integration and discrete Hamiltonian systems. *Journal of Nonlinear Science* 6 (1996), 449–467.
- [33] GONZALEZ, O. Mechanical systems subject to holonomic constraints: Differential-algebraic formulations and conservative integration. *Physica D: Nonlinear Phenomena* 132, 1-2 (1999), 165–174.
- [34] HAMEL, G. *Theoretische Mechanik*. Springer, Berlin, 1967.
- [35] HAMILTON, W. R. On quaternions; or on a new system of imaginaries in algebra. *Philosophical Magazine* 25-36 (1844-1850).
- [36] HESCH, C. *Mechanische Integratoren für Kontaktvorgänge deformierbarer Körper unter großen Verzerrungen*. PhD thesis, Universität Siegen, 2007.
- [37] HOWE, R. D., AND CUTKOSKY, M. R. Practical force-motion models for sliding manipulation. *The International Journal of Robotics Research* 15, 6 (1996), 557–572.
- [38] JEAN, M. Unilateral contact and dry friction: time and space variables discretization. *Archives of Mechanics* 40 (1988), 677–691.
- [39] JEAN, M. The non-smooth contact dynamics method. *Computer Methods in Applied Mechanics and Engineering* 177 (1999), 235–257.

- [40] KELLEY, C. T. *Solving nonlinear equations with Newton's method*. Society for Industrial and Applied Mathematics, 2003.
- [41] KLARBRING, A., AND BJÖRKMAN, G. A mathematical programming approach to contact problems with friction and varying contact surface. *Computers & Structures* 30, 5 (1988), 1185–1198.
- [42] KUIPERS, J. B. *Quaternions and Rotation Sequences*. Princeton University Press, Princeton NJ, 1999.
- [43] LABUDDE, R. A., AND GREENSPAN, D. Energy and momentum conserving methods of arbitrary order for the numerical integration of equations of motion. Part I: Motion of a single particle. *Numerische Mathematik* 25 (1975), 323–346.
- [44] LABUDDE, R. A., AND GREENSPAN, D. Energy and momentum conserving methods of arbitrary order for the numerical integration of equations of motion. Part II: Motion of a system of particles. *Numerische Mathematik* 26 (1976), 1–16.
- [45] LE SAUX, C., LEINE, R. I., AND GLOCKER, CH. Dynamics of a rolling disk in the presence of dry friction. *Journal of Nonlinear Science* 15, 1 (2005), 27–61.
- [46] LEINE, R. I. On the stability of motion in non-smooth mechanical systems. ETH Zurich, 2006. Habilitation thesis.
- [47] LEINE, R. I., AND GLOCKER, CH. A set-valued force law for spatial Coulomb-Contensou friction. *European Journal of Mechanics A/Solids* 22 (2003), 193–216.
- [48] LEINE, R. I., AND NIJMEIJER, H. *Dynamics and Bifurcations of Non-smooth Mechanical Systems*, vol. 18 of *Lecture Notes in Applied and Computational Mechanics*. Springer, Berlin, 2004.
- [49] LEINE, R. I., AND VAN DE WOUW, N. Stability properties of equilibrium sets of non-linear mechanical systems with dry friction and impact. *Nonlinear Dynamics* 51 (2008), 551–583.
- [50] LENS, E., AND CARDONA, A. An energy preserving/decaying scheme for nonlinearly constrained multibody systems. *Multibody System Dynamics* 18 (2007), 435–470.
- [51] LENS, E., CARDONA, A., AND GÉRADIN, M. Energy preserving time integration for constrained multibody systems. *Multibody System Dynamics* 11 (2004), 41–61.
- [52] LEYENDECKER, S. *Mechanical integrators for constrained dynamical systems in exible multibody dynamics*. PhD thesis, University of Kaiserslautern, 2006.
- [53] LEYENDECKER, S., BETSCH, P., AND STEINMANN, P. The discrete null space method for the energy-consistent integration of constrained mechanical systems. Part III: Flexible multibody dynamics. *Multibody System Dynamics* 19 (2008), 45–72.

- [54] MACIEJEWSKI, A. J. Hamiltonian formalism for Euler parameters. *Celestial Mechanics* 37 (1985), 47–57.
- [55] MÖLLER, M. Lösung der Time-Stepping Gleichung mit der Augmented Lagrangian Methode. Semesterarbeit, ETH Zürich, 2004.
- [56] MÖLLER, M., AND GLOCKER, CH. Analogous non-smooth models of mechanical and electrical systems. In *IUTAM Symposium on Multiscale Problems in Multibody System Contacts* (Stuttgart, Germany, 2007), P. Eberhard, Ed., vol. 1, pp. 45–54.
- [57] MÖLLER, M., AND GLOCKER, CH. Non-smooth modelling of electrical systems using the flux approach. *Nonlinear Dynamics* 50 (2007), 273–295.
- [58] MÖLLER, M., AND GLOCKER, CH. Rigid body dynamics with a scalable body, quaternions and perfect constraints. *Multibody System Dynamics* (2011). to appear.
- [59] MÖLLER, M., LEINE, R. I., AND GLOCKER, CH. An efficient approximation of set-valued force laws of normal cone type. In *Proceedings of the 7th EUROMECH Solid Mechanics Conference (ESMC2009)* (Lisbon, Portugal, 2009).
- [60] MOREAU, J. J. Unilateral contact and dry friction in finite freedom dynamics. In *Non-Smooth Mechanics and Applications*, J. J. Moreau and P. D. Panagiotopoulos, Eds., vol. 302 of *CISM Courses and Lectures*. Springer, Wien, 1988, pp. 1–82.
- [61] MOREAU, J. J. Numerical aspects of the sweeping process. *Computer Methods in Applied Mechanics and Engineering* 177 (1999), 329–349.
- [62] MORTON, JR., H. S. Hamiltonian and Lagrangian formulations of rigid-body rotational dynamics based on the Euler parameters. *Journal of the Astronautical Sciences* 41, 4 (1993), 569–591.
- [63] MURTY, K. G. *Linear complementarity, linear and nonlinear programming*. Heldermann, Berlin, 1988.
- [64] NGUYEN, M. H. Zeitliche Integration von strukturvarianten dynamischen Systemen mit Time-Splitting-Methoden. ETH Zürich, 2009. Masterarbeit.
- [65] NIKRAVESH, P. E. *Computer-aided analysis of mechanical systems*. Prentice-Hall International, London, 1988.
- [66] NIKRAVESH, P. E., AND CHUNG, I. S. Application of Euler parameters to the dynamic analysis of three-dimensional constrained mechanical systems. *Journal of Mechanical Design* 104 (1982), 785–791.
- [67] PAOLI, L., AND SCHATZMAN, M. A numerical scheme for impact problems I: The one-dimensional case. *SIAM Journal on Numerical Analysis* 40, 2 (2002), 702–733.
- [68] PAOLI, L., AND SCHATZMAN, M. A numerical scheme for impact problems II: The multi-dimensional case. *SIAM Journal on Numerical Analysis* 40, 2 (2002), 734–768.

- [69] PAPANASTAVRIDIS, J. G. *Analytical Mechanics: A Comprehensive Treatise on the Dynamics of Constrained Systems; for Engineers, Physicists, and Mathematicians*. Oxford University Press, 2002.
- [70] PAYR, M. *An Experimental and Theoretical Study of Perfect Multiple Contact Collisions in Linear Chains of Bodies*. PhD thesis, ETH Zürich, 2008.
- [71] PFEIFFER, F., AND GLOCKER, CH. *Multibody Dynamics with Unilateral Contacts*. John Wiley & Sons, New York, 1996.
- [72] ROCKAFELLAR, R. T. *Convex Analysis*. Princeton University Press, Princeton, New Jersey, 1970.
- [73] SCHWARZ, H. R. *Numerische Mathematik*. Teubner, Stuttgart, 1997.
- [74] SIMO, J. C., AND LAURSEN, T. A. An augmented lagrangian treatment of contact problems involving friction. *Computers & Structures* 42, 1 (1992), 97–116.
- [75] SIMO, J. C., AND WONG, K. K. Unconditionally stable algorithms for rigid body dynamics that exactly preserve energy and momentum. *International Journal for Numerical Methods in Engineering* 31, 1 (1991), 19–52.
- [76] STEWART, D., AND TRINKLE, J. An implicit time-stepping scheme for rigid body dynamics with inelastic collisions and coulomb friction. *International Journal For Numerical Methods In Engineering* 39, 15 (1996), 2673–2691.
- [77] STÖCKLI, F. *Stossgesetze und Energetische Konsistenz*. ETH Zürich, 2008. Masterarbeit.
- [78] STUDER, C. *Augmented time-stepping integration of non-smooth dynamical systems*. PhD thesis, ETH Zürich, 2008.
- [79] STUDER, C., AND GLOCKER, CH. Simulation of non-smooth mechanical systems with many unilateral constraints. *Proceedings ENOC-2005, Eindhoven* (2005).
- [80] STUDER, C., AND GLOCKER, CH. Representation of normal cone inclusion problems in dynamics via non-linear equations. *Archive of Applied Mechanics* 76, 5-6 (2006), 327–348.
- [81] STUDER, C., LEINE, R. I., AND GLOCKER, CH. Step size adjustment and extrapolation for time-stepping schemes in non-smooth dynamics. *International Journal for Numerical Methods in Engineering* 76 (2008), 1747–1781.
- [82] TRANSETH, A. A., LEINE, R. I., GLOCKER, CH., AND PETTERSEN, K. Y. 3D Snake robot motion: Non-smooth modeling, simulations, and experiments. *IEEE Transactions on Robotics* 24, 2 (2008), 361–376.
- [83] UHLAR, S. *Energy Consistent Time-Integration of Hybrid Multibody Systems*. PhD thesis, Universität Siegen, 2009.

- 
- [84] VADALI, S. R. On the Euler parameter constraint. *Journal of the Astronautical Sciences* 36, 3 (1988), 259–265.
- [85] WARD, J. P. *Quaternions and Cayley Numbers*. Kluwer Academic Publishers, Dordrecht, 1997.
- [86] WELGE-LÜSSEN, T. *Design of a Passively Actuated Robot Manipulator*. PhD thesis, ETH Zürich, 2008.



# Curriculum Vitae

---

Michael Heiner Möller  
born June 11, 1980  
Citizen of Zurich (Switzerland) and Germany

---

## Education

8/1995 – 1/2000	Kantonsschule Kreuzlingen, Matura Typus C
10/2000 – 4/2005	Studies on Mechanical Engineering at ETH Zurich with specialization on mechanics and control systems, graduation as Dipl. Masch.-Ing. ETH
4/2005 – 6/2011	Doctoral student at the Institute of Mechanical Systems, Center of Mechanics, ETH Zurich

---

## Professional Experience

4/2000 – 10/2000	Freelance work at Leicom AG, Winterthur (process control, communication techniques and building automation)
8/2003 – 10/2003	Internship at EADS Astrium GmbH, Friedrichshafen (modeling and control)
4/2005 – 6/2011	Research and teaching assistant at the Institute of Mechanical Systems, Center of Mechanics, ETH Zurich

---

## Award

11/2005	Willi-Studer-Preis 2005 for the best diploma exam in Mechanical and Process Engineering
---------	---