

## THÈSE

Pour obtenir le grade de

### DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : Mathématiques et Informatique

Arrêté ministériel : 7 Août 2006

Présentée par

**Alejandro Blumentals**

Thèse dirigée par **Bernard Brogliato**  
et codirigée par **Florence Bertails-Descoubes**

préparée au sein **Laboratoire Jean Kuntzmann**  
et de **EDMSTII**

## Numerical modelling of thin elastic solids in contact

Thèse soutenue publiquement le **03 Juillet 2017**,  
devant le jury composé de :

**Olivier Brüls**

Professeur, Université de Liège, Rapporteur

**Remco Leine**

Professeur, Université de Stuttgart, Rapporteur

**David Dureisseix**

Professeur, INSA Lyon, Examinateur

**Jean Claude Léon**

Professeur, Grenoble INP, Examinateur

**Bernard Brogliato**

Directeur de recherche, Inria Rhône Alpes, Directeur de thèse

**Florence Bertails-Descoubes**

Chargée de recherche, Inria Rhône Alpes, Co-Directrice de thèse



# Acknowledgements

I would first like to thank my thesis advisers Bernard Brogliato and Florence Bertails-Descoubes for their guidance during this endeavour. Bernard, I couldn't think of a better person with whom to learn mechanics with. And Florence, thank you for always having your door open for me and for always fuelling my scientific curiosity. I would also like to thank the members of my thesis committee for kindly accepting to evaluate my work and for their patient reading of my dissertation.

Thanks to my teachers and colleagues at Grenoble-INP Ensimag for motivating me to pursue scientific research and for making possible the wonderful teaching experience I had. Thanks to Valentin Sonnevile, Joachim Linn and Holger Lang for all the instructive exchanges we had concerning rods and shells. Thanks to all the members of our team at Inria, past and present, for the many interactions we had and for their help. Thanks Jerome, Pierre-Brice, Guillaume, Arnaud, Vincent and Dimitar: our many discussions, scientific and other, certainly shaped the course of my thesis. And thanks to Olivier, Romain, Gilles, Alexandre, Nestor, Nahuel, Jan and to all of those who made these years at Inria a real pleasure.

I would like to thank my family and friends for always encouraging me. And last but not least I would like to thank Jana, whose love and support is what makes my work possible.

# Contents

<b>General Introduction</b>	<b>5</b>
<b>I Lagrangian Mechanics of systems subject to sliding Coulomb's friction in bilateral and unilateral constraints</b>	<b>7</b>
<b>1 Introduction</b>	<b>8</b>
<b>2 Solvability of the contact problem for frictionless systems</b>	<b>13</b>
2.1 Bilaterally constrained systems . . . . .	13
2.2 Unilaterally constrained systems . . . . .	17
2.3 Unilaterally/bilaterally constrained systems . . . . .	20
<b>3 Lagrangian Mechanics of systems subject to unilateral and bilateral constraints with sliding Coulomb's friction</b>	<b>24</b>
3.1 Bilaterally constrained systems . . . . .	25
3.1.1 All frictional bilateral constraints . . . . .	26
3.1.2 Mixed frictional/frictionless contacts . . . . .	31
3.2 Unilaterally constrained systems . . . . .	33
3.3 Unilaterally/bilaterally constrained systems . . . . .	34
3.4 Conclusion . . . . .	38
<b>II The statics of elastic rods from an Optimal Control point of view</b>	<b>41</b>
<b>4 Introduction</b>	<b>42</b>
<b>5 An Optimal Control Interpretation of Rod Statics</b>	<b>44</b>
5.1 The continuum description of planar rods . . . . .	45
5.1.1 Planar Kirchhoff rods . . . . .	45
5.1.2 Planar Cosserat rods . . . . .	48
5.2 The continuum description of spatial rods . . . . .	49
5.2.1 Spatial Kirchhoff rods . . . . .	49
5.2.2 Spatial Cosserat rods . . . . .	51
5.3 The Kirchhoff Kinetic analogy . . . . .	53
5.4 Conclusion . . . . .	54
<b>6 A Hamilton-Pontryagin Approach for the statics of Kirchhoff rods</b>	<b>55</b>
6.1 Indirect Methods . . . . .	56
6.1.1 Derivation of the Kirchhoff Equations for rod statics . . . . .	57

6.1.2	Kirchhoff Indirect single shooting experiments . . . . .	60
6.1.3	Indirect methods for Euler's Elastica with fixed-fixed boundary conditions.	61
6.2	Discrete Kinematics . . . . .	66
6.2.1	On the choice of the integrator to discretize the Kinematics. . . . .	67
6.2.2	Runge-Kutta-Munthe-Kaas methods . . . . .	69
6.3	Direct Methods and the derivation of Strain based and mixed discretizations . .	76
6.3.1	Direct Single Shooting . . . . .	76
6.3.2	Direct Multiple Shooting . . . . .	80
6.3.3	Numerical evaluation . . . . .	85
6.4	Quasi static experiments . . . . .	87
6.4.1	Instability when varying $\bar{\kappa}_1$ . . . . .	87
6.4.2	Circle Packing . . . . .	87
6.4.3	Plectoneme Formation . . . . .	89
<b>III</b>	<b>Strain based dynamics of developable shells</b>	<b>96</b>
<b>7</b>	<b>Introduction</b>	<b>97</b>
<b>8</b>	<b>The Strain based approach to the mechanics of shells</b>	<b>100</b>
8.1	The first and second fundamental forms . . . . .	100
8.2	The fundamental theorem of surface theory . . . . .	103
8.3	Koiter Shells . . . . .	105
8.4	Strain based methods for shells . . . . .	106
<b>9</b>	<b>Strain based dynamics of developable shells</b>	<b>108</b>
9.1	Strain based kinematics of developable shells . . . . .	108
9.2	Dynamics of a developable shell patch with spatially constant curvatures . . . . .	110
9.2.1	Kinematics of a developable shell patch with spatially constant curvatures	110
9.2.2	Dynamics of the developable shell patch with spatially constant curvatures	113
9.2.3	Numerical solution of the equations of motion . . . . .	115
9.2.4	Implementation . . . . .	115
9.2.5	Difficulties . . . . .	118
9.3	Extension of the model . . . . .	118
9.3.1	Explicit solutions to the compatibility conditions . . . . .	118
9.3.2	Dynamics of a clothoidal shell patch . . . . .	120
9.4	Conclusion . . . . .	121
<b>Perspectives</b>		<b>123</b>
<b>A</b>	<b>Convex Analysis and Complementarity theory</b>	<b>125</b>
A.1	Some convex analysis and complementarity theory tools . . . . .	125
A.2	Theorem 3.1.7 in [40] (excerpts) . . . . .	126
A.3	Theorem 3.8.6 in [40] . . . . .	126
A.4	Theorems 2.8 and 2.11 in [123] . . . . .	126
<b>B</b>	<b>KKT system: solvability and solution uniqueness</b>	<b>127</b>
<b>C</b>	<b>Lie Groups</b>	<b>131</b>
C.1	Differential Geometry of $SO(3)$ and $SE(3)$ . . . . .	131

# General Introduction

Many objects around us, either natural or man-made, are slender deformable objects. Curve-like objects such as industrial cables, helicopter blades, plant stems and hair can be modelled as thin elastic rods. While surface-like objects such as paper, boat sails, leaves and cloth can be modelled as thin elastic plates and shells. The numerical study of the mechanical response of such structures is of the utmost importance in many applications of mechanical and civil engineering, bio-mechanics, computer graphics and other fields. In this dissertation we address several problems related to the numerical simulation of slender elastic structures subject to contact and friction constraints.

This thesis is organized in three parts and each part can be read independently. We have, however, in numerous occasions throughout this manuscript, underlined the similarities between the problems considered in each part.

Throughout this thesis we shall treat rods, plates and shells as finite dimensional multibody systems. When performing extensive numerical simulations of the dynamics of multibody systems subject to Coulomb's friction, it is common to encounter configurations where the Coulomb friction solver fails to compute the contact forces. While these failures may sometimes be due to implementation errors or to numerical ill-conditioning, it is also often the case that they are due to Painlevé paradoxes. Quite simply, in some configurations there may exist no contact force which can prevent the system from violating its contact constraints. This happens even for very simple mechanical systems, like that of a rigid rod sliding against a frictional plane, as Paul Painlevé first showed. In Part I of this manuscript we analyze the contact problem (whose unknowns are the accelerations and the contact forces) and we derive computable upper bounds on the friction coefficients at each contact, such that this problem is well-posed and Painlevé paradoxes are avoided. The obtained results apply not only to flexible systems, but to any general multibody system.

To approximate a rod or a shell as a multibody system, a spatial discretization procedure needs to be carried out. In Parts II and III we deal with the problem of how to carry out such spatial discretizations.

Thin rod-like structures which may easily bend and twist but hardly stretch and shear can be modelled as Kirchhoff rods. In the second part of this manuscript we consider the problem of computing the stable static equilibria of Kirchhoff rods subject to different boundary conditions and to frictionless contact constraints. We formulate the problem as an Optimal Control Problem, where the strains of the rod are interpreted as control variables and the position and orientation of the rod are interpreted as state variables. This leads us to a new understanding of the finite element discretization of Kirchhoff rods developed in [1] and to the proposal of new spatial discretization schemes for Kirchhoff rods. The proposed schemes are either of the strain-based type, where the main degrees of freedom are the strains of the rod, or of the mixed type, where the main degrees of freedom are both the strains and the generalized displacements.

Similarly to the case of Kirchhoff rods, thin surface-like structures such as paper can hardly

stretch or shear at all. One of the advantages of the strain based approach is that the no extension and no shear constraints of the Kirchhoff rod are handled intrinsically, without the need of stiff repulsion forces, or of further algebraic constraints on the degrees of freedom. In Part III of this manuscript we propose an extension of this approach to model the dynamics of inextensible and unshearable shells. We restrict our study to the case of a shell patch with a developable mid-surface. We use as primary degrees of freedom the components of the second fundamental form of the shell's mid-surface. This also leads to an intrinsic handling of the no shear and no extension constraints of the shell. To the best of our knowledge the strain based simulation of developable shells is a problem that had not been treated before. Our contributions represent a small step in this direction and we hope that it spurs further developments in the strain based approach to the numerical mechanics of shells.

The contents of Part I of the manuscript were communicated in [2] and [3]. The contents of Part II will be presented in the 8th Eccomas thematic conference on multibody dynamics in Prague, 2017. While those of Part III were communicated in [4].

## Part I

**Lagrangian Mechanics of systems  
subject to sliding Coulomb's friction  
in bilateral and unilateral constraints**

# Chapter 1

## Introduction

**Abstract** This Part deals with the existence and uniqueness of the acceleration and contact forces for Lagrangian systems subject to bilateral and/or unilateral constraints with or without sliding Coulomb's friction. Sliding friction is known to yield singularities in the system, such as Painlevé's paradox. Our work aims at providing sufficient conditions on the parameters of the system so that singularities are avoided (*i.e.*, the contact problem is at least solvable). To this end, the frictional problem is treated as a perturbation of the frictionless case. We provide explicit criteria, in the form of calculable upper bounds on the friction coefficients, under which the frictional contact problem is guaranteed to remain well-posed. The results exposed in this part correspond to the contents of our publication [3].

**Introduction.** Lagrangian systems subject to (frictional) bilateral and unilateral constraints are considered. Such systems, mathematically described in Equation (1.1) below, feature a very rich dynamics, because they are nonlinear, nonsmooth, and set-valued. A large number of studies have been led on their well-posedness. Their goal is to formulate conditions under which (1.1) possesses a solution, that is, a trajectory  $(q(\cdot), \dot{q}(\cdot))$  belonging to a certain functional space (with absolutely continuous positions and right-continuous velocities of local bounded variations), and satisfying (1.1) for all  $t \geq 0$ . Well-posedness also consists in determining whether a solution is unique for given initial data  $(q(0), \dot{q}(0))$ , whether solutions depend continuously on initial data or not, whether they converge to some equilibrium state (stability and control analysis [5]), *etc.*

Besides such analysis, one may in turn be interested in properties of least mathematical relevance, but of high interest for mechanical engineers. Typically, are the contact forces calculable during persistent motion phases ? Are they calculable in a unique way ? If some solvability results can be obtained in the absence of friction, what happens when Coulomb friction is added at some contacts ? More generally, how does the system evolve during persistent contact phases in the presence of friction, e.g., does the dynamics exhibit some singular states ? In multi-body systems dynamics, these questions are to be examined by studying the so-called *contact problem*, which assumes the mechanical state  $(q(t), \dot{q}(t))$  to be known at a given time  $t$ , and considers the acceleration  $\ddot{q}(t)$  and the contact forces at time  $t$  as unknowns. It thus results in a merely algebraic system, for which well-posedness remains a difficult question. Studying the contact problem is of high interest in multibody systems where one often wants to calculate the contact forces at a certain given time. Another motivation comes from event-driven numerical integration methods, where one has to solve the contact problem at a given time where possibly the state of the system may switch to another mode [6].

In this part our goal is to study conditions under which the contact problem is solvable. We first focus on frictionless systems, then on systems with sliding friction (single-valued law).

Indeed sliding friction is known to yield hard singularities like the Painlevé paradoxes [7, 8, 9, 10], while it has been shown in the seminal paper [11] that sticking modes can always, under some mild assumption, be continued in another contact mode (including detachment from the constraints).

For frictionless systems subject to bilateral constraints, the necessary and sufficient conditions for the existence and uniqueness of the acceleration in the presence of a singular mass matrix and redundant constraints are given in [12, 13]. The computation of constraint reactions in the redundant case is addressed via augmented Lagrangian methods in [14, 15, 16, 17], by a constraint elimination method combined with solvability analysis techniques in [18, 19, 20] and by a pseudo-inverse method in [21]. In the case of bilateral constraints with sliding Coulomb's friction, Matrosov and Finogenko derive an implicit criterion in [22, 23], which guarantees the uniqueness of the acceleration for small enough friction coefficients, however no explicit upper bounds are given. In [9] such an upper bound may be found but concerns only systems with a single contact point. In Chapter 3.1 we derive such an upper bound for the case of multiple contact points. For systems subject only to unilateral constraints and sliding Coulomb's friction an existence result based on small enough friction coefficients and complementarity theory is given in [24] and an explicit upper bound is established in [11]. The contact problem with mixed (bilateral and unilateral) constraints has received surprisingly less attention, its analysis in the frictionless case with redundant constraints and a singular mass matrix is given in [25]. In the frictional sliding case it is established in [26] that it becomes a mixed linear complementarity problem but no explicit condition for its solvability is given. In Chapter 3.3 we derive an upper bound on the friction coefficients guaranteeing the solvability of the sliding friction problem with mixed constraints.

## Dynamics of Lagrangian systems subject to bilateral, unilateral and frictional constraints

In a Lagrangian formalism such systems may be written generically as follows,

$$\left\{ \begin{array}{l} M(q)\ddot{q} + F(q, \dot{q}, t) = \nabla h_{n,b}(q)\lambda_{n,b} + \nabla h_{n,u}(q)\lambda_{n,u} + H_{t,b}(q)\lambda_{t,b} + H_{t,u}(q)\lambda_{t,u} \end{array} \right. \quad (1.1a)$$

$$\left\{ \begin{array}{l} \text{Sliding friction (Coulomb)} : \lambda_{t,i} = -\mu_i |\lambda_{n,i}| \frac{v_{t,i}}{\|v_{t,i}\|}, v_{t,i} \neq 0, 1 \leq i \leq m \end{array} \right. \quad (1.1b)$$

$$\left\{ \begin{array}{l} \text{Complementarity conditions} : 0 \leq h_{n,u}(q) \perp \lambda_{n,u} \geq 0 \end{array} \right. \quad (1.1c)$$

$$\left\{ \begin{array}{l} \text{Bilateral (holonomic) constraints} : h_{n,b}(q) = 0, \end{array} \right. \quad (1.1d)$$

where  $q \in \mathbb{R}^n$  is the vector collecting the generalized coordinates  $q_i$ ,  $1 \leq i \leq n$ , assumed to be independent when all the constraints are removed,  $\dot{q}$  is the vector of generalized velocities,  $M(q) = M(q)^T$  is the inertia matrix, always assumed to be at least positive semi-definite (it may be assumed non-singular in some cases),  $F(q, \dot{q}, t)$  collects internal forces (including forces deriving from a potential, plus Coriolis and centrifugal forces), as well as external actions on the system such as disturbances or control.

We consider  $m = m_u + m_b$  constraints consisting of  $m_u$  unilateral (inequality) constraints  $h_{n,u}(q) \in \mathbb{R}^{m_u}$  and  $m_b$  bilateral (equality) constraints  $h_{n,b}(q) \in \mathbb{R}^{m_b}$ . The matrix  $\nabla h_{n,u}(q)$  (respectively  $\nabla h_{n,b}(q)$ ) collects on each column the gradient for each unilateral constraint (for each bilateral constraint, respectively). The vectors of Lagrange multipliers  $\lambda_{n,u} \in \mathbb{R}^{m_u}$  and  $\lambda_{n,b} \in \mathbb{R}^{m_b}$  are associated with the unilateral and bilateral constraints, respectively. From a mechanical point of view,  $\lambda_{n,u}$  and  $\lambda_{n,b}$  correspond to the normal components of the contact forces in the unilateral and bilateral cases, and are obtained from the local contact kinematics [27, Chapter 10] [6, Chapter 3]. The unilateral constraints and their associated Lagrange multipliers are related through the complementarity condition (1.1c), which is to be understood

componentwise (per contact). It models the fact that for each contact  $i$ , the normal contact force should not act at a unilateral contact point if the contact is open (i.e.,  $h_{n,u,i}(q) > 0$  implies  $\lambda_{n,u,i} = 0$ ), and that  $\lambda_{n,u,i} > 0$  if and only if  $h_{n,u,i}(q) = 0$ . The matrices  $H_{t,b}(q)$  and  $H_{t,u}(q)$  map local tangential velocities to generalized velocities (see Section 3 for more details).

The coefficients of friction at each contact point  $i$  are  $\mu_i \geq 0$ . We consider in this study frictional contacts in a sliding mode only (non-zero relative tangential velocities). The contact problem consists in determining the acceleration and the contact forces given  $(q, \dot{q})$  at a given time instant. This means that by a proper choice of the local contact kinematics frames one has  $\frac{v_{t,i}}{\|v_{t,i}\|} = \text{sgn}(v_{t,i})$ , both for 2-dimensional and 3-dimensional friction. Also note that we do not consider systems subject to impacts in the formulation (1.1).

This part is organized as follows. In Chapter 2 we review in detail the frictionless case, where results for the bilaterally and unilaterally constrained cases are recalled. In Chapter 3, it is shown how the problem gets all the more complex as Coulomb's friction is considered and as mixed constraints (both bilateral and unilateral) are added. Some useful results are recalled in the Appendix A.

## Further notations

**Mathematical notations**  $\mathbb{R}^n$  is the set of  $n$ -vectors with real entries,  $\mathbb{R}_+^n$  is the set of  $n$ -vectors with non-negative entries. Let  $a_1, a_2, \dots, a_n$  be some given reals, then  $[a] = \text{diag}(a_i)$  is the  $n \times n$  diagonal matrix with entries  $a_i$ . Let  $A \in \mathbb{R}^{n \times n}$  be a real square matrix, its induced norm is  $\|A\|_2 = \max_{x \in \mathbb{R}^n, \|x\|_2=1} \|Ax\|_2$ , where  $\|x\|_2$  is the Euclidean norm on the vector space  $\mathbb{R}^n$ . This induced matrix norm is sometimes denoted as  $\|A\|_{2,2}$ . One has  $\|A\|_2 = \sigma_{\max}(A) = \sqrt{\lambda_{\max}(AA^T)}$  [28, Proposition 9.4.7], where  $\sigma_{\max}$  is the maximum singular value of  $A$ , and  $\lambda_{\min}(\cdot)$  and  $\lambda_{\max}(\cdot)$  are its smallest and largest eigenvalues, respectively [28, Proposition 9.4.7]. Moreover, if  $A$  is invertible,  $\sigma_{\min}(A) = \frac{1}{\sigma_{\max}(A^{-1})}$ , where  $\sigma_{\min}(A)$  is the smallest singular value of  $A$  [28, fact 6.3.21]. A positive definite (resp. semi definite) matrix is denoted  $A \succ 0$  (resp.  $A \succeq 0$ ), it may be non-symmetric. Let  $f : \mathbb{R}^n \mapsto \mathbb{R}^p$  be a differentiable function. Its Euclidean gradient is  $\nabla f(x) = (\nabla f_1(x) \ \nabla f_2(x) \ \dots \ \nabla f_p(x)) \in \mathbb{R}^{n \times p}$ , and its Jacobian  $\frac{\partial f}{\partial x}(x) = \nabla f(x)^T$ . The cardinality of a countable set  $\mathcal{I}$  is denoted as  $\text{card}(\mathcal{I})$ . Let  $S \subseteq \mathbb{R}^n$  be a set, its orthogonal complement is defined as  $S^\perp \triangleq \{x \in \mathbb{R}^n | x^T y = 0 \text{ for all } y \in S\}$  and is a subspace. Its boundary is denoted as  $\text{bd}(S)$ . One has  $\text{Ker}(A) = \text{Im}(A^T)^\perp$  for any matrix  $A$ .

**Mechanical notations** Since contacts may be frictionless, or in contrast may involve Coulomb friction, the following conventions shall be adopted:

- Bilateral contacts:  $1 \leq i \leq m_b$ , i.e.,  $i \in \mathcal{I}_b$ .
  - Frictional bilateral contacts (sliding):  $1 \leq i \leq m_b^\mu$ , i.e.,  $i \in \mathcal{I}_b^\mu$ .
  - Frictionless bilateral contacts:  $m_b^\mu + 1 \leq i \leq m_b$ , i.e.,  $i \in \mathcal{I}_b^0$ .
- Unilateral contacts:  $m_b + 1 \leq i \leq m$ , i.e.,  $i \in \mathcal{I}_u$ .
  - Frictional unilateral contacts (sliding):  $m_b + 1 \leq i \leq m_u^\mu$ , i.e.,  $i \in \mathcal{I}_u^\mu$ .
  - Frictionless unilateral contacts:  $m_u^\mu + 1 \leq i \leq m$ , i.e.,  $i \in \mathcal{I}_u^0$ .

We may therefore rewrite the first line in (1.1) as

$$\begin{aligned}
M(q)\ddot{q} + F(q, \dot{q}, t) &= \sum_{i \in \mathcal{I}_b^0} \nabla h_{n,b,i}^0(q) \lambda_{n,b,i}^0 + \sum_{i \in \mathcal{I}_u^0} \nabla h_{n,u,i}^0(q) \lambda_{n,u,i}^0 \\
&\quad + \sum_{i \in \mathcal{I}_b^\mu} \nabla h_{n,b,i}(q) \lambda_{n,b,i} + \sum_{i \in \mathcal{I}_b^\mu} H_{t,b,i}(q) \lambda_{t,b,i} \\
&\quad + \sum_{i \in \mathcal{I}_u^\mu} \nabla h_{n,u,i}(q) \lambda_{n,u,i} + \sum_{i \in \mathcal{I}_u^\mu} H_{t,u,i}(q) \lambda_{t,u,i}
\end{aligned} \tag{1.2}$$

## Mechanical systems

Throughout this part, examples will help illustrate the concepts. The different mechanical systems that will be considered are depicted in Figure 1.1.

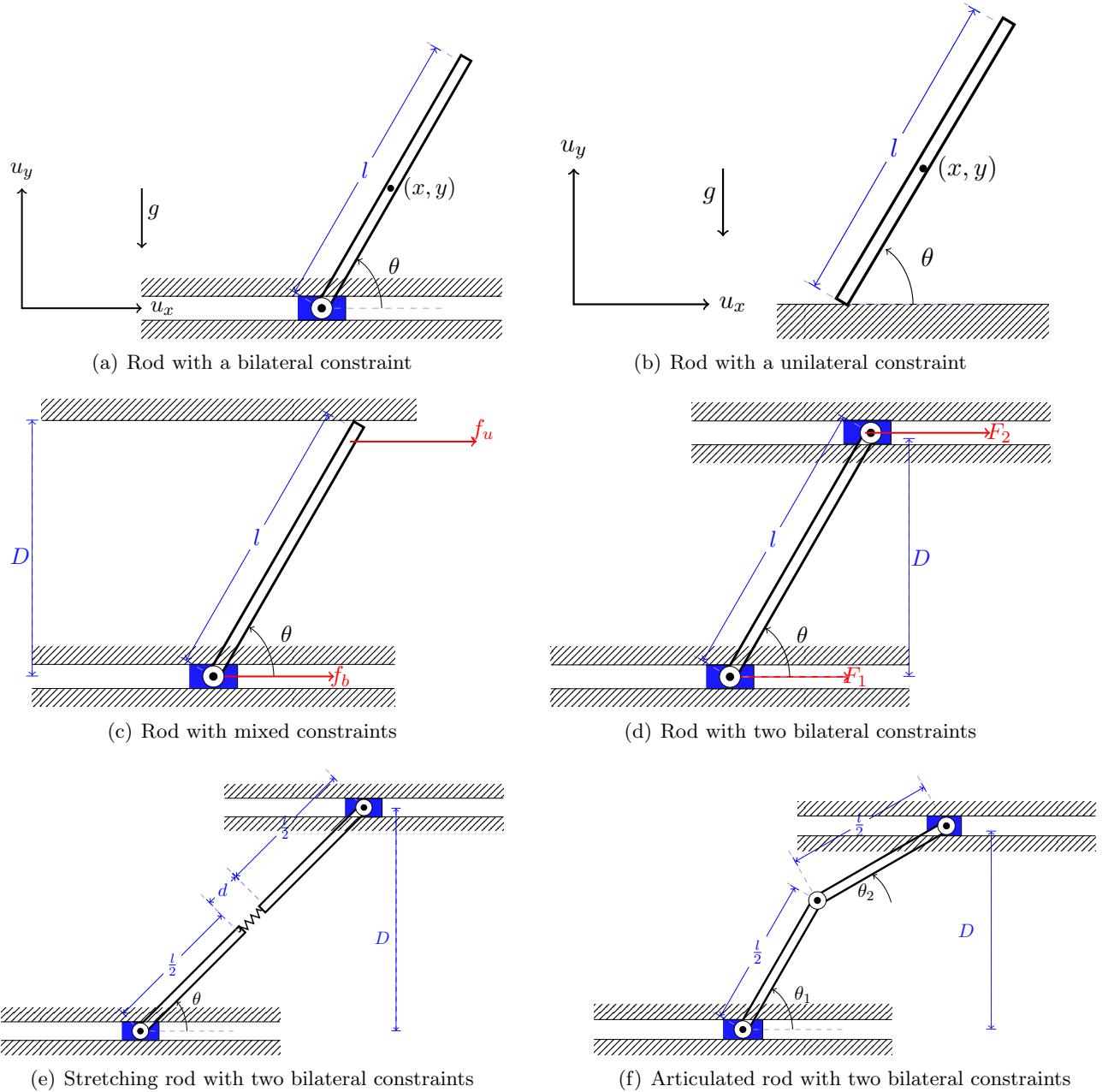


Figure 1.1: Mechanical systems with constraints.

# Chapter 2

## Solvability of the contact problem for frictionless systems

This chapter deals with the solvability of the contact problem for mechanical systems without friction. Three cases are considered. The case of bilaterally constrained systems of index 1 is treated in Section 2.1. The case of unilaterally constrained systems is treated in Section 2.2. The case of systems subject to both bilateral and unilateral constraints is treated in Section 2.3. In all cases we study the problem of existence and uniqueness of the contact forces and the acceleration at a given time, for a given position and velocity. In the frictionless case difficulties arise mainly when the mass matrix is singular or when the constraints are redundant. We give a thorough exposition of the known results on this topic which were initially scattered throughout the literature. Understanding the structure of these problems is crucial for the study of sliding friction case which is carried out in Chapter 3. As we will later see, friction adds a perturbation responsible for the loss of structure which may lead to the unsolvability of the contact problem, even in the case of a non-singular mass matrix and non-redundant constraints.

### 2.1 Bilaterally constrained systems

Let us assume that all constraints are bilateral and frictionless. The Lagrangian system (1.1) boils down to a differential-algebraic equation (DAE) of index 3 [29, 30],

$$\begin{cases} M(q)\ddot{q} + F(q, \dot{q}, t) = \nabla h_{n,b}(q)\lambda_{n,b} \\ h_{n,b}(q) = 0. \end{cases} \quad (2.1)$$

Reducing the index to 1 consists in differentiating twice the constraint  $h_{n,b}(q) = 0$  to obtain  $\ddot{h}_{n,b} = \nabla h_{n,b}(q)^T \ddot{q} + \frac{d}{dt}(\nabla h_{n,b}(q)^T) \dot{q} = 0$ , while keeping consistent initial conditions in position and velocity. This new equality combined with the dynamics yields the following linear system,

$$\underbrace{\begin{pmatrix} M(q) & -\nabla h_{n,b}(q) \\ \nabla h_{n,b}(q)^T & 0 \end{pmatrix}}_{\triangleq M_b(q)} \begin{pmatrix} \ddot{q} \\ \lambda_{n,b} \end{pmatrix} = \begin{pmatrix} -F(q, \dot{q}, t) \\ -\frac{d}{dt}(\nabla h_{n,b}(q)^T) \dot{q} \end{pmatrix}. \quad (2.2)$$

**Remark** Time-varying constraints  $h_{n,b}(q, t) = 0$  may be considered in the analysis. Then

$$\frac{d}{dt}h_{n,b}(q, t) = \nabla h_{n,b}(q, t)^T \dot{q}(t) + \frac{\partial h_{n,b}}{\partial t}(q, t)$$

and

$$\frac{d^2}{dt^2}h_{n,b}(q, t) = \nabla h_{n,b}(q, t)^T \ddot{q}(t) + \frac{d}{dt}(\nabla h_{n,b}(q, t)^T) \dot{q}(t) + \frac{\partial}{\partial q} \frac{\partial h_{n,b}}{\partial t}(q, t) \dot{q}(t) + \frac{\partial^2 h_{n,b}}{\partial t^2}(q, t).$$

If needed it is thus sufficient to add the missing terms to step from the time-invariant to the time-varying case.

Recall that  $M(q)$  is assumed to be symmetric positive semi-definite (with possibly  $\text{rank}(M(q)) < n$ ). The matrix in the left-hand side of (2.1), denoted as  $M_b(q)$  and often called the *DAE matrix*, is ubiquitous not only in the analysis and numerics of Lagrangian systems with holonomic constraints [29, 30, 31, 13, 32, 12, 33], but also in convex quadratic minimization. In the latter context, the  $M_b(q)$  matrix is referred to as the *KKT matrix* and system (2.2) corresponds to the *KKT system* of the following quadratic minimization problem [34, §10.1.1],

$$\begin{aligned} \ddot{q}(t) &= \operatorname{argmin}_x \left( \frac{1}{2} x^T M(q) x + F(q, \dot{q}, t)^T x \right) \\ \text{subject to } &\nabla h_{n,b}(q)^T x + \frac{d}{dt} (\nabla h_{n,b}(q)^T) \dot{q} = 0, \end{aligned} \quad (2.3)$$

which can be, in the case when  $M(q)$  is non-singular, interpreted mechanically as the Gauss principle of least constraints applied to a Lagrangian system subject to bilateral holonomic constraints [21].

The DAE or KKT matrix  $M_b(q)$  is a positive semi-definite matrix. The fact that  $M_b(q)$  has a skew-symmetric part in (2.2) is not intrinsic to the problem. Indeed it stems from an arbitrary choice in the way the bilateral constraints are introduced in the dynamics (1.1), and thus of the choice of the sign of the multiplier. By changing the sign of the multiplier, the problem may be analyzed equivalently without the minus sign in  $-\nabla h_{n,b}(q)$ , thus setting  $M_b(q)$  symmetric (meanwhile causing the loss of positiveness for  $M_b(q)$ ). This is actually the convention adopted in most of the DAE literature. The next proposition gathers solvability results for the system (2.2) from Optimization [34, §10.1.1] and Mechanics [12].

**Proposition 2.1.1** (Well-posedness of the bilateral frictionless contact problem). *Let all  $m_b$  constraints be bilateral and frictionless. Let the position and velocity  $(q, \dot{q})$  be given at a time instant  $t$ . Consider the KKT system in (2.2) with unknowns the acceleration  $\ddot{q}$  and the multipliers  $\lambda_{n,b} \in \mathbb{R}^{m_b}$ .*

- (i) *Let  $m_b < n$  and  $\nabla h_{n,b}(q)$  have full (column) rank  $m_b$ . Then  $M_b(q)$  is non-singular (or equivalently, given any arbitrary right-hand side vector of (2.2), there exists a unique solution  $(\ddot{q}, \lambda_{n,b})$  to (2.2)) if and only if*

$$\operatorname{Ker}(M(q)) \cap \operatorname{Ker}(\nabla h_{n,b}(q)^T) = \{0\}. \quad (2.4)$$

Moreover, if  $M(q) \succ 0$ , then the multiplier  $\lambda_{n,b}$  can be computed in closed form by solving the linear equation

$$A_{nb}(q) \lambda_{n,b} + w_b(q, \dot{q}, t) = 0, \quad (2.5)$$

where the symmetric positive definite matrix

$$A_{nb}(q) := \nabla h_{n,b}(q)^T M(q)^{-1} \nabla h_{n,b}(q) \quad (2.6)$$

is the Schur complement<sup>1</sup> of  $M(q)$  in the DAE matrix  $M_b(q)$ , and where

$$w_b(q, \dot{q}, t) := -\nabla h_{n,b}(q)^T M(q)^{-1} F(q, \dot{q}, t) + \frac{d}{dt} (\nabla h_{n,b}(q)^T) \dot{q}.$$

---

<sup>1</sup>The Schur complement of the invertible matrix  $A_{11}$  in the matrix  $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ , is the matrix  $A_{22} - A_{21} A_{11}^{-1} A_{12}$ . The Schur complement of the invertible matrix  $A_{22}$  in  $A$ , is the matrix  $A_{11} - A_{12} A_{22}^{-1} A_{21}$ .

- (ii) Let  $\nabla h_{n,b}(q)$  have arbitrary rank and satisfy the compatibility of constraints, i.e.,  $-\frac{d}{dt}(\nabla h_{n,b}(q)^T)\dot{q} \in \text{Im}(\nabla h_{n,b}(q)^T)$ . Then, given  $(q, \dot{q})$  and an arbitrary force vector  $F(q, \dot{q}, t)$ ,
  - A solution  $(\ddot{q}, \lambda_{n,b})$  of (2.2) exists, and
  - The acceleration  $\ddot{q}$  and the generalized contact force  $\nabla h_{n,b}(q)\lambda_{n,b}$  are unique,
if and only if condition (2.4) holds.

**Proof** Let us denote for brevity  $x := \ddot{q}$ ,  $\lambda := \lambda_{n,b}$ ,  $M := M(q)$ ,  $F := F(q, \dot{q}, t)$ ,  $J := \nabla h_{n,b}(q)^T$  and  $b := -\frac{d}{dt}(\nabla h_{n,b}(q)^T)\dot{q}$ . Since the mass matrix  $M$  is symmetric positive semi-definite the quadratic function  $f(x) := \frac{1}{2}x^T M x + F^T x$  is convex. The constraint set

$$K_b := \{x \in \mathbb{R}^n : Jx = b\} \quad (2.7)$$

is affine and hence convex. The convex quadratic program (2.3) reads

$$\begin{cases} \min_{x \in \mathbb{R}^n} & \frac{1}{2}x^T M x + F^T x \\ \text{s. t.} & Jx = b. \end{cases} \quad (2.8)$$

The KKT system (2.2) reads simply

$$\begin{bmatrix} M & -J^T \\ J & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} -F \\ b \end{bmatrix}. \quad (2.9)$$

In both items (i) and (ii) the compatibility of constraints  $b \in \text{Im}(J)$  holds. In (i) because  $J$  is full rank and in (ii) by hypothesis. The compatibility of constraints is precisely Slater's constraint qualification for the affine constraints  $Jx = b$ . Under these conditions (convex problem and qualified constraints) the KKT theorem [34, §5.5.3] states that  $x$  is a solution of the (QP) (2.8) if and only if there exists  $\lambda \in \mathbb{R}^{m_b}$  such that  $(x, \lambda)$  is a solution of the KKT system (2.9).

The solution  $x$  of the (QP) (2.8) exists and is unique if and only if  $\text{Ker}(M) \cap \text{Ker}(J) = \{0\}$  (that is if 2.4 holds).

To see this let us introduce a full rank matrix  $N$  whose columns span  $\text{Ker}(J)$  and a feasible point  $x_0$  such that  $Jx_0 = b$ . Then the constraint set can be parametrized as  $C = \{x_0 + Nz : z \in \mathbb{R}^d\}$ , where  $d = \dim \text{Ker}(J)$ . Then the (QP) (2.8) is equivalent to the unconstrained (QP) with objective function

$$g(z) := f(x_0 + Nz) = \frac{1}{2}z^T N^T M N z + (Mx_0 + F)^T Nz + f(x_0).$$

- Let us prove by contraposition that if the solution  $x$  of the (QP) (2.8) exists and is unique then  $\text{Ker}(M) \cap \text{Ker}(J) = \{0\}$ . If there exists  $v \in \text{Ker}(M) \cap \text{Ker}(J)$  such that  $v \neq 0$  then along that direction the (QP) (2.8) is unbounded by below. There exists a non zero  $z \in \mathbb{R}^d$  such that  $v = Nz$  and the objective function reduces to an affine function in  $z$ :  $g(z) = (Mx_0 + F)^T Nz + f(x_0)$ . Since  $F$  is arbitrary  $N^T(Mx_0 + F)$  is non zero and the objective is unbounded.
- Conversely, let us assume that  $\text{Ker}(M) \cap \text{Ker}(J) = \{0\}$ . Then the hessian  $N^T M N$  of  $g$ , which is already symmetric positive semi-definite, is positive definite since if there is  $z \in \mathbb{R}^d$  such that  $z^T N^T M N z = 0$  then  $Nz$  is in  $\text{Ker}(M)$  and in  $\text{Ker}(J)$  and hence is zero. Then the objective function  $g$  is strictly convex and continuous over  $\mathbb{R}^d$  and hence has a unique global minimum.

If the solution  $x$  of (2.9) exists and is unique then multipliers  $\lambda$  exist such that  $(x, \lambda)$  is a solution of (2.9). The uniqueness of generalized contact force  $J^T \lambda$  in item (ii) follows from the uniqueness of the acceleration  $x$ . If we assume that  $J^T$  is full column rank as in item (i) then the multiplier vector  $\lambda$  is also unique. If furthermore the mass matrix  $M$  is invertible then the linear equation for the multipliers (2.5) follows simply by inverting the mass matrix in the KKT system. ■

**Remark** The proof of the above proposition originally proposed in our article [3] relies on the material in Appendix B, where the solvability of the KKT system (2.2), and of other related systems, is treated from a linear algebra perspective.

Item (ii) obviously relies on weaker assumptions than item (i), since  $\nabla h_{n,b}(q)$  is not required to have full column rank  $m_b$ , but rather satisfies the compatibility of constraints, as in [12]. In particular, we consider in (ii) cases where  $m_b > n$ , i.e., where  $M_b(q)$  is necessarily a singular matrix, and where some constraints are redundant. Note that provided  $M(q) \succ 0$ , the Schur complement satisfies  $A_{nb}(q) = A_{nb}(q)^T \succeq 0$ . Furthermore, we have  $\text{Ker}(A_{nb}(q)) = \text{Ker}(\nabla h_{nb}(q))$ . In particular,  $A_{nb}(q)$  is non-singular if and only if  $\nabla h_{nb}(q)$  has full column rank  $m_b \leq n$ . Singular mass matrices and redundant constraints are common features of multibody dynamical systems subject to bilateral constraints, due to rotation parametrization [35], or redundant generalized coordinates [13, 36].

In the following example we compute explicitly the contact forces and the acceleration for a bilaterally constrained rigid rod. We will see in the next chapter that when sliding friction is added to the rod, the equation for the contact force is no longer linear but rather piecewise linear. It may then fail to have a unique solution or a any solution at all depending on the magnitude of the friction: this is known as a Painlevé-like paradox.

**Example: 2D bilaterally constrained rigid rod.** Let us study the sliding rigid rod depicted in Figure 1.1(a), which can be seen as a rigid pendulum with a frictionless sliding base. The only forces applied here are gravity (oriented downwards) and the net (normal) contact force  $\lambda_{n,b}$  (taken positive by convention when oriented upwards). Let  $m > 0$  be the mass of the rod and  $l > 0$  its total length. Let us choose  $q = (x, y, \theta)^T$ , where  $x$  and  $y$  are the coordinates of the center of mass of the rod, and  $\theta \in [0, \pi]$  is the angle between the horizontal line and the main axis of the rod. The dynamic equations read

$$\begin{cases} m\ddot{x} &= 0 \\ m\ddot{y} &= -mg + \lambda_{n,b} \\ \frac{ml^2}{12}\ddot{\theta} &= -\frac{l}{2}\cos\theta\lambda_{n,b} \\ h_{n,b}(q) &= y - \frac{l}{2}\sin\theta = 0, \end{cases} \quad (2.10)$$

where the last equality expresses the prismatic constraint applied onto the bottom tip of the rod. Note that the multiplier  $\lambda_{n,b} \in \mathbb{R}$  may take positive (upward net force) or negative (downward net force) values. Formulating (2.10) as the canonical system (2.2), one obtains  $M(q) = M = \text{diag}(m, m, \frac{ml^2}{12})$ ,  $\nabla h_{n,b}(q) = (0, 1, -\frac{l}{2}\cos\theta)^T$ ,  $F(\dot{q}, q, t) = (0, mg, 0)^T$ . Observe that both  $M$  and  $\nabla h_{n,b}(q)$  have full rank, hence this case study falls into the first (the most classical) category (i) examined in Proposition 2.1.1.  $M$  being full rank, the Schur complement of the DAE matrix is well-defined and can be computed as the scalar number  $A_{nb}(q) = \nabla h_{nb}(q)^T M^{-1} \nabla h_{nb}(q) = \frac{1+3\cos^2\theta}{m}$ . The reduced equation for  $\lambda_{n,b}$  then boils down to a linear scalar equation

$$A_{nb}(q)\lambda_{n,b} + w_b(q, \dot{q}) = 0,$$

where  $w_b(q, \dot{q}) = \frac{d}{dt} (\nabla h_{n,b}(q)^T) \dot{q} - g = \frac{l}{2} \dot{\theta}^2 \sin \theta - g$ . As expected,  $A_{nb}(q)$  is non-singular and thus the solution for  $\lambda_{n,b}$  exists and is unique. Similarly, from the dynamic equation (2.10) there exists a unique solution for the generalized acceleration  $\ddot{q}$  of the system. Finally we get

$$\lambda_{n,b} = m \left( \frac{\frac{l}{2} \dot{\theta}^2 \sin \theta + g}{1 + 3 \cos^2 \theta} \right) \quad \ddot{q} = \left( 0, -g + \frac{\frac{l}{2} \dot{\theta}^2 \sin \theta + g}{1 + 3 \cos^2 \theta}, -\frac{6}{l} \cos \theta \left( \frac{\frac{l}{2} \dot{\theta}^2 \sin \theta + g}{1 + 3 \cos^2 \theta} \right) \right)^T.$$

## 2.2 Unilaterally constrained systems

Now we assume that all contacts are unilateral and frictionless. The Lagrangian system (1.1) boils down to the complementarity system

$$\begin{cases} M(q)\ddot{q} + F(q, \dot{q}, t) = \nabla h_{n,u}(q)\lambda_{n,u} \\ 0 \leq \lambda_{n,u} \perp h_{n,u}(q) \geq 0. \end{cases} \quad (2.11)$$

**Construction of the contact MLCP.** In the following we carry out a procedure analogous to index reduction to construct the so-called *contact MLCP*, which we subsequently analyze.

**Proposition 2.2.1.** *Let  $h(\cdot)$  and  $\lambda(\cdot)$  be two scalar functions of time, and let  $0 \leq h(t) \perp \lambda(t) \geq 0$  for all  $t$ . Assume that  $h(\cdot)$  is absolutely continuous, that  $\dot{h}(\cdot)$  is absolutely continuous in a right neighbourhood of  $t$ , and that  $\ddot{h}(\cdot)$  and  $\lambda(\cdot)$  are right-continuous at time  $t$ . (i) Let  $h(t) = 0$ , then  $0 \leq \dot{h}(t) \perp \lambda(t) \geq 0$ . (ii) Let  $h(t) = 0$  and  $\dot{h}(t) = 0$ , then  $0 \leq \ddot{h}(t) \perp \lambda(t) \geq 0$ .*

**Proof:** (i) For any  $t' \geq t$  one has  $h(t') - h(t) = \int_t^{t'} \dot{h}(s)ds$ . Hence  $h(t') = \int_t^{t'} \dot{h}(s)ds$  since  $h(t) = 0$ .

- First we prove that  $\dot{h}(t) \geq 0$ . Suppose that  $\dot{h}(t) < 0$ . Since  $\dot{h}(\cdot)$  is right-continuous, there exists  $\epsilon > 0$  such that  $\dot{h}(s) < 0$  for all  $s \in [t, t + \epsilon]$ . Thus for any  $t' \in [t, t + \epsilon]$  one has  $h(t') < 0$  which is impossible. Hence  $\dot{h}(t) \geq 0$ .
- Secondly, we prove that if  $\dot{h}(t) > 0$  then  $\lambda(t) = 0$ . Let us assume that  $\dot{h}(t) > 0$ . Then there exists  $\epsilon > 0$  such that for all  $t' \in (t, t + \epsilon)$ , we have  $\dot{h}(t') > 0$ . Consequently for all  $t' \in (t, t + \epsilon)$ ,  $h(t') > 0$  and  $\lambda(t') = 0$ . Then by right continuity of  $\lambda(\cdot)$ , we conclude that  $\lambda(t) = 0$ .
- Finally, if  $\lambda(t) > 0$  then  $\dot{h}(t) = 0$ . Indeed, if  $\lambda(t) > 0$  and  $h(t) = 0$  then assuming that  $\dot{h}(t) > 0$  leads to a contradiction: we would have that for all  $t' \in (t, t + \epsilon)$   $h(t') > 0$  and  $\lambda(t') = 0$ , hence the contradiction  $\lambda(t) = 0$ . Since  $\dot{h}(t)$  is non negative but not strictly positive we conclude that  $\dot{h}(t) = 0$ .

Part (ii) is proved in a similar way. ■

Assume that  $q(\cdot)$  and  $h_{n,u}(\cdot)$  are continuous, while  $\dot{q}(\cdot)$ ,  $\ddot{q}(\cdot)$  and  $\lambda_{n,u}(\cdot)$  are right-continuous. In fact since we disregard impacts we may even assume that  $\dot{q}(\cdot)$  is continuous. Replacing  $h(t)$  with  $h_{n,u,i}(q(t))$  and  $\lambda(t)$  with  $\lambda_{n,u,i}(t)$  allows one to assert that  $h_{n,u,i}(q(t)) = 0$  implies that

$$0 \leq \nabla h_{n,u,i}(q(t))^T \dot{q}(t) \perp \lambda_{n,u,i}(t) \geq 0,$$

while  $h_{n,u,i}(q(t)) = 0$  and  $\nabla h_{n,u,i}(q(t))^T \dot{q}(t) = 0$  implies

$$0 \leq \nabla h_{n,u,i}(q(t))^T \ddot{q}(t) + \frac{d}{dt} (\nabla h_{n,u,i}(q)^T) \dot{q} \perp \lambda_{n,u,i}(t) \geq 0.$$

In view of this, the acceleration  $\ddot{q}$  and the multipliers  $\lambda_{n,u}$  are solutions of the *mixed LCP* (*MLCP*)

$$\begin{cases} M(q)\ddot{q} + F(q, \dot{q}, t) = \nabla h_{n,u}(q)\lambda_{n,u} \\ 0 \leq \lambda_{n,u} \perp \nabla h_{n,u}(q)^T \ddot{q} + \frac{d}{dt}(\nabla h_{n,u}(q)^T) \dot{q} \geq 0, \end{cases} \quad (2.12)$$

where  $\nabla h_{n,u}(q)$  now collects only the gradients of the active constraints at  $q$ . The MLCP (2.12) is the counterpart of (2.2) for unilateral constraints.

**Analysis of the contact MLCP.** Let us assume that the constraint set

$$K_u := \{x \in \mathbb{R}^n : \nabla h_{n,u}(q)^T x + \frac{d}{dt}(\nabla h_{n,u}(q)^T) \dot{x} \geq 0\} \quad (2.13)$$

is non empty. Then Slater's constraint qualification for linear inequalities holds. Hence, by the KKT theorem and the convexity of the function

$$f(x) := \frac{1}{2}x^T M(q)x + F(q, \dot{q}, t)^T x$$

it holds that  $\ddot{q}$  is a solution of the (QP)

$$\begin{cases} \min_{x \in \mathbb{R}^n} & \frac{1}{2}x^T M(q)x + F(q, \dot{q}, t)^T x \\ \text{s. t.} & \nabla h_{n,u}(q)^T x + \frac{d}{dt}(\nabla h_{n,u}(q)^T) \dot{x} \geq 0 \end{cases} \quad (2.14)$$

if and only if there exists  $\lambda_{n,u}$  such that  $(\ddot{q}, \lambda_{n,u})$  is a solution of the MLCP (2.12). Claiming that the acceleration solves the QP (2.14) is a form of Gauss' principle of mechanics extended to frictionless unilateral constraints.

In the case of a positive definite mass matrix the existence and uniqueness of the acceleration which solves the QP (2.14) and hence the MLCP (2.12) follows naturally without any further assumption on  $\nabla h_{n,u}(q)$  nor on  $F(q, \dot{q}, t)$ . This fact was known to Moreau [37, 38]. The case of a singular mass matrix is more complicated and less well studied. In [25] a sufficient criterion to guarantee the solvability of the MLCP (2.12) is derived. It is analogous to the condition in (2.4) for bilateral systems. We collect these results in the following proposition.

**Proposition 2.2.2** (Well-posedness of the frictionless unilateral contact problem). *Let the position and velocity  $(q, \dot{q})$  be given at a time instant  $t$ . Let all  $m_u$  constraints be unilateral, frictionless and active at  $q$ . Let us assume that the constraint set  $K_u$  is non empty. Consider the MLCP in (2.12) with unknowns the acceleration  $\ddot{q}$  and the multipliers  $\lambda_{n,u} \in \mathbb{R}^{m_u}$ .*

- (i) *If  $M(q)$  is positive definite then there exists a unique acceleration  $\ddot{q}$  and there exists  $\lambda_{n,u}$  which solve the frictionless unilateral MLCP (2.12). The generalized contact force  $\nabla h_{n,u}(q)\lambda_{n,u}$  is unique. The multipliers  $\lambda_{n,u}$  are furthermore unique if and only if  $\nabla h_{n,u}(q)$  is of full column rank  $m_u$ .*
- (ii) *If the mass matrix is only positive semi definite ( $M(q) \succeq 0$ ) and if the gap functions  $h_{n,u,i}(q)$ ,  $1 \leq i \leq m_u$  satisfy the Mangasarian-Fromovitz constraint qualification then the following statement holds.*

*The MLCP in (2.12) is solvable if*

$$T_{\Phi_u}(q) \cap \text{Ker}(M(q)) = \{0\},$$

*where  $\Phi_u := \{q \in \mathbb{R}^n | h_{n,u}(q) \geq 0\}$  and  $T_{\Phi_u}(q) := \{v \in \mathbb{R}^n | \nabla h_{n,u}(q)^T v \geq 0\}$ .*

*Proof.* Under the conditions in the proposition the equivalence of the MLCP (2.12) and the QP (2.14) holds. To prove item (i) it suffices to show that the QP (2.14) has a unique solution. This is indeed the case. The objective function  $f$  is continuous and also 0-coercive (infinite at infinity) when  $M(q) \succ 0$ . Furthermore the constraint set  $K_u$  of the QP is closed and non empty. This guarantees the existence of a solution  $\ddot{q}$  of the QP. Uniqueness of  $\ddot{q}$  holds since the constraint set  $K$  is convex and since  $f$  is strictly convex when  $M(q) \succ 0$ . The existence of the acceleration  $\ddot{q}$  solving the QP then implies the existence of a multiplier  $\lambda_{n,u}$  such that  $(\ddot{q}, \lambda_{n,u})$  solves the MLCP (2.12). The uniqueness of the generalized contact force  $\nabla h_{n,u}(q)\lambda_{n,u}$  then follows from the uniqueness of the acceleration since  $\nabla h_{n,u}(q)\lambda_{n,u} = M(q)\ddot{q} + F(q, \dot{q}, t)$ . Furthermore  $\lambda_{n,u}$  is unique if and only if  $\nabla h_{n,u}(q)$  is injective or, equivalently, of full column rank.

Item (ii) is proved in [25].  $\square$

**Remark** In practice, to numerically compute the acceleration and the contact forces one can use an MLCP solver to solve (2.12) (see for example [39]). However to better exploit the structure of the problem it is preferable to use a QP solver to solve (2.14). Active set and interior point QP solvers would compute not only the acceleration but also the multipliers. Open source implementations of QP solvers are also more widely available. In the case where the mass matrix is positive definite a third option is available since the contact forces can be computed by solving a Linear Complementarity Problem (LCP) and the accelerations determined explicitly from the contact forces.

**Construction of the contact LCP** In the case that  $M(q) \succ 0$ , the MLCP (2.12) is easily transformed by elimination of  $\ddot{q}$  to construct the *frictionless contact LCP*. The acceleration  $\ddot{q}$  is explicitly determined by the contact forces by setting

$$\ddot{q} = M(q)^{-1}(\nabla h_{n,u}(q)\lambda_{n,u} - F(q, \dot{q}, t)).$$

By inserting the latter expression for  $\ddot{q}$  in the second line of the MLCP (2.12) we obtain the LCP

$$0 \leq \lambda_{n,u} \perp A_{nu}(q)\lambda_{n,u} + w_u(q, \dot{q}, t) \geq 0, \quad (2.15)$$

where

$$A_{nu}(q) := \nabla h_{n,u}(q)^T M(q)^{-1} \nabla h_{n,u}(q)$$

is the *Delassus matrix*, and

$$w_u(q, \dot{q}, t) := -\nabla h_{n,u}(q)^T M(q)^{-1} F(q, \dot{q}, t) + \frac{d}{dt}(\nabla h_{n,u}(q)^T) \dot{q}.$$

The well-posedness of the frictionless unilateral contact problem can then also be studied by analyzing the LCP (2.15) and relying on results of Linear Complementarity theory [40]. This path is taken in [41], [42]. Most notably there exists a unique contact force  $\lambda_{n,u}$  solving the LCP for any vector  $w_u(q, \dot{q}, t)$  if and only if  $A_{n,u}(q)$  is a P-matrix. Since  $A_{n,u}(q)$  is automatically symmetric positive semi-definite, it is a P-matrix if and only if it is non singular, which holds if and only if the constraints are linearly independent.

**Remark** The LCP in (2.15) is solvable if and only if the acceleration constraint set  $K_u$  is non empty. If  $K_u$  is non empty then the existence of  $\lambda_{n,u}$  holds by Proposition 2.2.2 item (i). If the LCP(2.15) has a solution  $\lambda_{n,u}$  then  $M(q)^{-1}(\nabla h_{n,u}(q)\lambda_{n,u} - F(q, \dot{q}, t))$  is in  $K_u$  and hence  $K_u$  is non empty.

**Remark** It is certainly not the case that any LCP( $A, w$ ) with a symmetric positive semi-definite matrix  $A$  is solvable for any given vector  $w$ . The necessary and sufficient conditions for such an LCP to be solvable are that for any  $\lambda \geq 0$  in  $\text{Ker}(A)$ ,  $w^T \lambda \geq 0$ . Applying this result from complementarity theory for  $A = A_{n,u}(q)$  and  $w = w_u(q, \dot{q}, t)$  we arrive to the result stated in [3] (Proposition 3(iv)) and in [42, pp.211-212] that the LCP in (2.15) is solvable if and only if the implication  $\lambda_{n,u} \geq 0, \nabla h_{n,u}(q) \lambda_{n,u} = 0 \Rightarrow \lambda_{n,u}^T \frac{d}{dt}(\nabla h_{n,u}(q)^T) \dot{q} \geq 0$  holds. However, we know by the remark above the latter condition is always verified. This was not clear in [3] nor in [42]. The LCP (2.15) is not just any LCP, it is the KKT system associated to the dual problem of the QP (2.14). Since the QP (2.14) has an optimal solution if  $K_u$  is non empty and  $M(q) \succ 0$ , the dual function

$$g(\lambda) := -\frac{1}{2} \lambda^T A_{n,u}(q)^{-1} \lambda - w_u(q, \dot{q}, t)^T \lambda - \frac{1}{2} F(q, \dot{q}, t)^T M(q)^{-1} F(q, \dot{q}, t)$$

is upper bounded for  $\lambda \geq 0$ .

In the study of the sliding friction case we will no longer be able to rely on results of convex optimization. We will see that friction adds a perturbation to the Delassus matrix which destroys the convexity of the problem. We will have to rely on complementarity theory results to analyze when it is that the perturbed Delassus operator remains a P-matrix.

In the following example we derive the contact LCP for a unilaterally constrained rigid rod sliding without friction against a plane and deduce the existence and uniqueness of the contact force. In the next chapter we will consider this same example with sliding Coulomb's friction and see that when the friction is too high the existence and uniqueness of the contact force no longer holds: this is known as the *Painlevé paradox*.

**Example: Unilaterally constrained rigid rod** Consider the system in Figure 1.1(b) with a unilateral constraint. The rod is allowed to slide on the horizontal plane but it may also take off (note that  $\theta$  is assumed to belong to  $[0, \pi]$  so that the whole rod should remain in the upper half space). This example corresponds to the classical example of Painlevé, albeit without friction. The dynamics of the system reads

$$\begin{cases} m\ddot{x} \\ m\ddot{y} \\ I\ddot{\theta} \\ 0 \leq h_{nu}(\theta) = y - \frac{l}{2} \sin(\theta) \end{cases} \begin{aligned} &= 0 \\ &= -mg + \lambda_{n,u} \\ &= -\frac{l}{2} \cos \theta \lambda_{n,u} \\ &\perp \lambda_{n,u} \geq 0 \end{aligned}$$

where the complementarity expresses the unilateral constraint applied onto the bottom tip of the rod. In contrast with the bilateral case, the multiplier  $\lambda_{n,u} \in \mathbb{R}^+$  can take positive values only, meaning that the contact force should always be oriented upwards. Derivations of  $M$ ,  $F$ , and  $\nabla h_{n,u}(q)$  are identical to the example in Section 2.1 (replacing  $\nabla h_{n,b}(q)$  with  $\nabla h_{n,u}(q)$ ), and the LCP in  $\lambda_{n,u}$  reads

$$0 \leq A_{n,u}(q) \lambda_{n,u} + w_u(q, \dot{q}) \perp \lambda_{n,u} \geq 0$$

with  $A_{n,u}(q) = \frac{1+3 \cos^2 \theta}{m}$  and  $w_u(q, \dot{q}) = \frac{l}{2} \dot{\theta}^2 \sin \theta - g$ . Since  $A_{n,u}(q) \succ 0$ , the solution for  $\lambda_{n,u}$  (and thus for  $\dot{q}$ ) exists and is unique. This is in accordance with Proposition 2.2.2(i), which applies here since  $\nabla h_{n,u}(q)$  has full column rank.

## 2.3 Unilaterally/bilaterally constrained systems

We now consider the case when both bilateral and unilateral constraints are involved. Strangely enough, it is only recently that the analysis of this case has received attention [43, 25], though

it may represent the most common case in practice [44]. Let us write the bilateral constraints on the acceleration level. The MLCP in (2.12) is augmented as follows:

$$\left\{ \begin{array}{l} (a) \quad M(q)\ddot{q} + F(q, \dot{q}, t) = \nabla h_{n,b}(q)\lambda_{n,b} + \nabla h_{n,u}(q)\lambda_{n,u} \\ (b) \quad \nabla h_{n,b}(q)^T \ddot{q} + \frac{d}{dt}(\nabla h_{n,b}(q)^T) \dot{q} = 0 \\ (c) \quad 0 \leq \lambda_{n,u} \perp \nabla h_{n,u}(q)^T \ddot{q} + \frac{d}{dt}(\nabla h_{n,u}(q)^T) \dot{q} \geq 0. \end{array} \right. \quad (2.16)$$

The acceleration is constrained to be both in the affine space

$$K_b = \left\{ x \in \mathbb{R}^n : \nabla h_{n,b}(q)^T x + \frac{d}{dt}(\nabla h_{n,b}(q)^T) \dot{x} = 0 \right\}$$

and in the convex polyhedral set

$$K_u = \left\{ x \in \mathbb{R}^n : \nabla h_{n,u}(q)^T x + \frac{d}{dt}(\nabla h_{n,u}(q)^T) \dot{x} \geq 0 \right\}.$$

Again under the assumption that the constraint set for the acceleration  $K_b \cap K_u$  is non empty Slater's linear constraint qualification holds. We can then deduce by the KKT theorem for convex optimization problems that the acceleration  $\ddot{q}$  is a solution of the QP

$$\left\{ \begin{array}{ll} \min_{x \in \mathbb{R}^n} & \frac{1}{2} x^T M(q)x + F(q, \dot{q}, t)^T x \\ \text{s. t.} & \nabla h_{n,b}(q)^T x + \frac{d}{dt}(\nabla h_{n,b}(q)^T) \dot{x} = 0 \\ & \nabla h_{n,u}(q)^T x + \frac{d}{dt}(\nabla h_{n,u}(q)^T) \dot{x} \geq 0 \end{array} \right. \quad (2.17)$$

if and only if there exists multipliers  $(\lambda_{n,b}, \lambda_{n,u})$  such that  $(\ddot{q}, \lambda_{n,b}, \lambda_{n,u})$  is a solution of the MLCP (2.16). The claim that the acceleration of the mechanical system is a solution of the QP (2.17) constitutes a form of Gauss' principle of mechanics for systems subject to both unilateral and bilateral frictionless constraints.

The next proposition summarizes well-posedness results for the unilateral/bilateral frictionless contact problem.

**Proposition 2.3.1** (Well-posedness of the frictionless contact problem with bilateral and unilateral constraints). *Let the position and velocity  $(q, \dot{q})$  be given at a time instant  $t$ . Let there be  $m_u$  unilateral, frictionless and active constraints at  $q$ , as well as  $m_b$  bilateral constraints. Let us assume that the constraint set  $K_b \cap K_u$  is non empty. Consider the MLCP in (2.12) with unknowns the acceleration  $\ddot{q}$  and the multipliers  $(\lambda_{n,b}, \lambda_{n,u}) \in \mathbb{R}^{m_b+m_u}$ .*

- (i) *If  $M(q)$  is positive definite then there exists a unique acceleration  $\ddot{q}$  and there exists  $(\lambda_{n,b}, \lambda_{n,u})$  which solve the MLCP (2.16). The generalized contact force  $\nabla h_{n,b}(q)\lambda_{n,b} + \nabla h_{n,u}(q)\lambda_{n,u}$  is unique. The multipliers  $(\lambda_{n,b}, \lambda_{n,u})$  are furthermore unique if and only if  $(\nabla h_{n,b}(q), \nabla h_{n,u}(q))$  is of full column rank  $m_b + m_u$ .*
- (ii) *If the mass matrix is only positive semi definite ( $M(q) \succeq 0$ ) and if the functions  $h_{n,b}(q)$  and  $h_{n,u}(q)$  satisfy the Mangasarian-Fromovitz constraint qualification then the following statement holds.*

*The MLCP in (2.16) is solvable if*

$$\text{Ker}(\nabla h_{n,b}(q)) \cap T_{\Phi_u}(q) \cap \text{Ker}(M(q)) = \{0\}.$$

The proof of item (i) is identical to that of Proposition 2.2.2, while item (ii) is proved in [25].

**Contact LCP with a bilateral distortion.** For the study of the contact problem with sliding friction in both bilateral and unilateral constraints we will no longer be able to rely on convex optimization arguments. We will have to study how the LCP for the contact forces is modified by the addition of friction. In the following we construct the LCP obeyed by  $\lambda_{n,u}$ . Let us assume that  $K_b \cap K_u$  is non empty, that  $M(q) \succ 0$  and that  $\nabla h_{n,b}(q)$  has full column rank. The acceleration can be expressed as a function of the contact forces via equation (2.16)(a) as

$$\ddot{q} = M(q)^{-1} (\nabla h_{n,b}(q)\lambda_{n,b} + \nabla h_{n,u}(q)\lambda_{n,u} - F(q, \dot{q}, t)).$$

Inserting the latter expression into equation (2.16)(a,b) we find that the contact forces are solutions of the MLCP

$$\begin{aligned} A_{n,b}(q)\lambda_{n,b} + A_{nbnu}(q)\lambda_{n,u} + w_b(q, \dot{q}, t) &= 0 \\ 0 \leqslant A_{nbnu}(q)^T\lambda_{n,b} + A_{n,u}(q)\lambda_{n,u} + w_u(q, \dot{q}, t) \perp \lambda_{n,u} &\geqslant 0, \end{aligned}$$

where the matrix

$$A_{nbnu}(q) := \nabla h_{n,b}(q)^T M(q)^{-1} \nabla h_{n,u}(q)$$

is responsible for the coupling of bilateral and unilateral contact forces. Since  $\nabla h_{n,b}(q)$  is of full rank  $A_{n,b}(q)$  is invertible and  $\lambda_{n,b}$  can be expressed explicitly as a function of  $\lambda_{n,u}$  as

$$\lambda_{n,b} = -A_{n,b}(q)^{-1}(A_{nbnu}(q)\lambda_{n,u} + w_b(q, \dot{q}, t)).$$

Finally the LCP obeyed by the unilateral contact forces  $\lambda_{n,u}$  reads

$$0 \leqslant A_c(q)\lambda_{n,u} + w_c(q, \dot{q}, t) \perp \lambda_{n,u} \geqslant 0, \quad (2.18)$$

where

$$A_c(q) := A_{n,u}(q) - A_{nbnu}(q)^T A_{n,b}(q)^{-1} A_{nbnu}(q)$$

is the Delassus matrix with distortion and the vector  $w_c$  reads

$$w_c(q, \dot{q}, t) := -A_{nbnu}(q)^T A_{n,b}(q)^{-1} w_b(q, \dot{q}, t) + w_u(q, \dot{q}, t).$$

**Remark** The matrix  $A_c(q)$  is by construction symmetric positive semi-definite. It is the Schur complement of  $A_{n,b}(q)$  in the matrix

$$S := (\nabla h_{n,b}(q) \ \nabla h_{n,u}(q))^T M(q)^{-1} (\nabla h_{n,b}(q) \ \nabla h_{n,u}(q)) = \begin{bmatrix} A_{n,b}(q) & A_{nbnu}(q) \\ A_{nbnu}(q)^T & A_{n,u}(q) \end{bmatrix}.$$

Since the matrix  $S$  is positive semi definite and  $A_{n,b}(q) \succ 0$  the Schur complement of  $A_{n,b}(q)$  in  $S$  is also positive semi-definite [34, A.5.5].

Given the hypotheses made in this paragraph the LCP (2.18) is solvable. Its solution is also unique if furthermore  $(\nabla h_{n,b}(q), \nabla h_{n,u}(q))$  is of full column rank (or equivalently if  $A_c(q) \succ 0$ ). In the next chapter we will study how the addition of sliding friction adds a perturbation which destroys the symmetry of the matrix  $A_c(q)$ . We will expose a limit on the magnitude of the friction which guarantees that the perturbed matrix remains a P-matrix such as to avoid Painlevé-type paradoxes.

In the next example we consider a rigid rod subject to both a bilateral and a unilateral constraint. For this system, in some configurations the contact forces are non unique.

**Example: The rod with one bilateral and one unilateral constraint** Consider the system in Figure 1.1(c), where the upper tip of the rod is unilaterally constrained and the lower tip is bilaterally constrained. The external forces are horizontal pulling forces  $f_b, f_u$  at each tip. One obtains the matrices involved in the dynamics:  $M = \text{diag}(m, m, \frac{ml^2}{12})$ ,  $\nabla h_{n,b}(q) = (0, 1, -\frac{l}{2} \cos(\theta))^T$ ,  $\nabla h_{n,u}(q) = (0, -1, -\frac{l}{2} \cos(\theta))^T$ ,  $F = (-f_b - f_u, 0, -\frac{l}{2} (f_b - f_u) \sin(\theta))^T$ ,  $w_c(q, \dot{q}, t) = -\frac{3 \sin(\theta) \cos(\theta)}{m(3 \cos^2(\theta) + 1)} (2(f_b - f_u) + lm \cos(\theta) \dot{\theta}^2)$ , and  $A_c(q) = \frac{12 \cos^2(\theta)}{m(3 \cos^2(\theta) + 1)}$ . The LCP in (2.18) then becomes :

$$0 \leq \lambda_{n,u} \perp \frac{12 \cos^2(\theta)}{m(3 \cos^2(\theta) + 1)} \lambda_{n,u} - \frac{3 \sin(\theta) \cos(\theta)}{m(3 \cos^2(\theta) + 1)} (2(f_b - f_u) + lm \cos(\theta) \dot{\theta}^2) \geq 0.$$

It has a unique solution for all  $\theta \neq \frac{\pi}{2}$ . If  $\theta = \frac{\pi}{2}$  then  $(\nabla h_{n,b}(q), \nabla h_{n,u}(q))$  is not of full rank and the LCP becomes  $0 \leq \lambda_{n,u} \perp 0 \geq 0$  so any non negative  $\lambda_{n,u}$  is a solution.

## Conclusion

In this chapter we analyzed the frictionless contact problems for mechanical systems subject to bilateral and unilateral constraints. We have summarized well-posedness results for these problems in propositions 2.1.1, 2.2.2 and 2.3.1. In the frictionless case, non existence or non uniqueness issues occur only when the mass matrix is singular or when the constraints aren't linearly independent.

When the mass matrix is non singular, the mere existence of a kinematically admissible acceleration ( $K_b \cap K_u \neq \emptyset$ ) guarantees the existence of the acceleration (an essential prerequisite to time step the mechanical system forward in time), as well as its uniqueness. When furthermore the constraints are linearly independent the contact forces are also unique.

In the following chapter we will study how the problems to compute the contact forces get all the more complicated as Coulomb's friction is considered.

## Chapter 3

# Lagrangian Mechanics of systems subject to unilateral and bilateral constraints with sliding Coulomb's friction

We now turn our attention to systems that are subject to sliding Coulomb's friction.

**Overview.** This chapter is organized in three sections, similarly to chapter 2. We analyze respectively in each section systems for which the constraints are only bilateral, only unilateral and finally both unilateral and bilateral. We will see that even in the case of a non singular mass matrix and of linearly independent constraints, the contact problem may be ill-posed. Our goal is to define bounds on the friction coefficients such that this does not occur. The bounds we derive in the bilateral and in the mixed case were previously unknown in the literature, while the bound in the unilateral case may be found in [11].

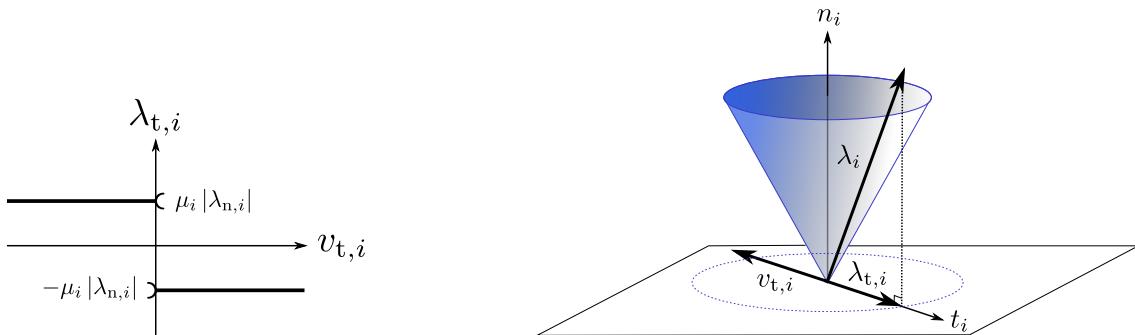


Figure 3.1: Sliding Coulomb friction can be expressed compactly in 2D as  $\lambda_{t,i} = -\mu_i |\lambda_{n,i}| \xi_i$  where  $\xi_i \in \{-1; 1\}$  is the (monovalued) sign of  $v_{t,i}$ . In 3D, the same law holds when expressed in the local frame  $(n_i, t_i)$ , where  $t_i$  is a tangent vector oriented along the sliding line (with arbitrary direction).

**The Lagrange dynamics with sliding friction** We shall assume in the sequel that all frictional contact points lie in the sliding mode of Coulomb friction. In 2D and 3D, this allows one to build for each contact point a local kinematical frame  $(n_i, t_i)$ , where  $n_i$  is the normal

at contact and  $t_i$  is oriented along the sliding line (see Figure 3.1 for an illustration). Using suitable local kinematics frames (see for instance [6, §3.3]), one can define the (scalar) tangential velocity  $v_{t,i} \neq 0$  and formulate the sliding Coulomb's friction as

$$\lambda_{t,i} = -\mu_i |\lambda_{n,i}| \xi_i \quad \text{where } \xi_i \in \{-1; 1\} \text{ is univoquely given by the sign of } v_{t,i}.$$

Let  $v_{t,b}$  be the vector that gathers local tangential velocities for the bilaterally constrained contact points, and  $v_{t,u}$  its unilateral counterpart. Then the operators  $H_{t,b}(q)$  and  $H_{t,u}(q)$  define a mapping between local and generalized tangential velocities, written as  $v_{t,b} = H_{t,b}(q)^T \dot{q}$  and  $v_{t,u} = H_{t,u}(q)^T \dot{q}$ . Let  $F_{t,b}$  denote the generalized force resulting from the contribution of the frictional (tangential) forces acting on the system. From the principle of virtual works it follows that  $v_{t,b}^T \lambda_{t,b} = \dot{q}^T H_{t,b}(q) \lambda_{t,b} = \dot{q}^T F_{t,b}$ . Hence  $F_{t,b} = H_{t,b}(q) \lambda_{t,b}$ , and similarly for the unilateral constraints. Then, using (1.1a) and (1.1b), one obtains the force balance equation for systems with sliding friction

$$M(q)\ddot{q} + F(q, \dot{q}, t) = \nabla h_{n,b}(q) \lambda_{n,b} + \nabla h_{n,u}(q) \lambda_{n,u} - H_{t,b}(q) [\mu_b \xi_b] |\lambda_{n,b}| - H_{t,u}(q) [\mu_u \xi_u] \lambda_{n,u}. \quad (3.1)$$

The matrix  $[\mu_b \xi_b]$  (respectively  $[\mu_u \xi_u]$ ) denotes the diagonal matrix with entries  $\mu_{b,i} \xi_{b,i}$  for  $i$  in  $1, \dots, m_b$  (respectively  $\mu_{u,i} \xi_{u,i}$  for  $i$  in  $1, \dots, m_u$ ). The vector  $|\lambda_{n,b}| := (|\lambda_{n,b,1}| \ |\lambda_{n,b,2}| \dots |\lambda_{n,b,m_b}|)^T \in \mathbb{R}_+^{m_b}$  collects the absolute values of the normal components of the bilateral contact forces. Note that unilateral normal forces  $\lambda_{n,u}$  do not need to be expressed in terms of absolute value since they should always remain positive due to the complementarity constraint (1.1c). Equation (3.1) serves as a starting point to analyze (1.1) with sliding Coulomb's friction.

**Remark** It is noteworthy however that Coulomb's law cannot be written as an *associated* law<sup>1</sup> [45, §4]. This means that the right-hand side of (3.1) together with the complementarity conditions, cannot be written compactly as the normal cone to some convex set. De Saxcé's bipotential function [46] allows one to recover an associated form at the local kinematics level, however at the price of using a modified tangential velocity [6, §3.9.2].

### 3.1 Bilaterally constrained systems

In this section we analyze the case where all  $m_b$  constraints are bilateral. We assume that the mass matrix of the system is positive definite ( $M(q) \succ 0$ ) and that all constraints are linearly independent. Recall that under these conditions the frictionless contact problem is well posed and that the contact force  $\lambda_{n,b}$  is the unique solution of the linear equation

$$A_{n,b}(q) \lambda_{n,b} + w_b(q, \dot{q}, t) = 0.$$

We will see in Section 3.1.1 how the equation for the contact forces becomes a piecewise linear equation rather than a linear one when all constraints are subject to sliding Coulomb's friction. We will then proceed to derive bounds on the bilateral friction coefficients which guarantee the existence and uniqueness of the contact force and hence of the acceleration. In Section 3.1.2 the result is extended to systems where some constraints may be frictional and others frictionless. We illustrate our results on the so called Painlevé-Klein example and on two other related mechanical systems.

---

<sup>1</sup>A contact law is associated if it can be expressed as an inclusion in the subdifferential of a convex, proper function, *i.e.* it admits a convex pseudo-potential.

### 3.1.1 All frictional bilateral constraints

Suppose that all  $m_b$  constraints are bilateral and that all contacts are sliding:  $\xi_{b,i} := \operatorname{sgn}(v_{t,b,i}) \in \{-1, 1\}$ . Using (3.1) the extension of the KKT system in (2.2) is:

$$\underbrace{\begin{pmatrix} M(q) & -\nabla h_{n,b}(q) + H_{t,b}(q)[\mu_b \xi_b] [\operatorname{sgn}(\lambda_{n,b})] \\ \nabla h_{n,b}(q)^T & 0 \end{pmatrix}}_{\triangleq M_{b\mu\xi}(q)} \begin{pmatrix} \ddot{q} \\ \lambda_{n,b} \end{pmatrix} = \begin{pmatrix} -F(q, \dot{q}, t) \\ -\frac{d}{dt}(\nabla h_{n,b}(q)^T) \dot{q} \end{pmatrix} \quad (3.2)$$

and we may denote the matrix in the left-hand side of (3.2) as  $M_{b\mu\xi}(q)$ . Comparing with (2.2) one sees that friction modifies the system's matrix, not its right-hand side. However the matrix  $M_{b\mu\xi}(q)$  depends non linearly on  $\lambda_{n,b}$  and hence equation (3.2) is not a linear equation in  $(\ddot{q}, \lambda_{n,b})$  as it was in the frictionless case. Hence relying on results of linear algebra is not possible. Bilateral friction introduces a non-linearity or rather a piecewise linearity in the variable  $\lambda_{n,b}$ .

A direct way of analyzing the existence and uniqueness of solutions  $(\ddot{q}, \lambda_{n,b})$  of (3.2) is to form the analog of (2.5),

$$A_{nb}(q)\lambda_{n,b} - A_{tb}(q)[\mu_b \xi_b] |\lambda_{n,b}| + w_b(q, \dot{q}, t) = 0, \quad (3.3)$$

where the absolute value is meant componentwise and  $A_{tb}(q) \triangleq \nabla h_{n,b}(q)^T M(q)^{-1} H_{t,b}(q)$  is a Delassus-like matrix coupling normal and tangential frictional effects. By inspecting each of the  $2^{m_b}$  possible signs of  $\lambda_{n,b}$  one obtains  $2^{m_b}$  linear systems. One may solve each such system and obtain a candidate solution  $\lambda_{n,b}$ . If the signs of the newly computed  $\lambda_{n,b}$  coincide with those of the assumption, a solution has indeed been found, otherwise it is rejected. If no solution or several solutions are found, the system is said to be wedged or jammed. Such an exhaustive procedure would yield necessary and sufficient conditions on the parameter values for a unique solution of problem (3.3) to exist (see [9] in the case of a single contact). However, in practice, it becomes intractable for systems with more than a few contact points.

In the rest of this section we present two alternative ways to derive sufficient conditions for equation (3.2) to have a unique solution. Problem (3.3) can be treated as a perturbation of the frictionless problem (2.5) and the non-linearity introduced by bilateral friction can be treated via a fixed point argument or via complementarity theory.

**Proposition 3.1.1.** *Let  $(q, \dot{q})$  be given at a time instant  $t$ . Let  $A_{nb}(q)$  be positive definite (or equivalently  $\nabla h_{n,b}(q)$  of full rank). Suppose that all contacts are sliding ( $v_{t,b,i} \neq 0$ ) and that the friction coefficients  $\mu_i$  satisfy*

$$\max_{1 \leq i \leq m} \mu_i < \mu_{\max}^b(q), \quad (3.4)$$

where we define the bound  $\mu_{\max}^b(q)$  as

$$\mu_{\max}^b(q) := \frac{\sigma_{\min}(A_{nb}(q))}{\sigma_{\max}(A_{tb}(q))}. \quad (3.5)$$

Then the bilateral sliding friction problem (3.2) has a unique solution  $(\ddot{q}, \lambda_{n,b})$ .

**Proof:** Under the rank assumption made in the proposition, system (3.2) is equivalent to

$$\begin{cases} \ddot{q} = M(q)^{-1} \nabla h_{n,b}(q) \lambda_{n,b} - M(q)^{-1} H_{t,b}(q)[\mu_b \xi_b] |\lambda_{n,b}| - M(q)^{-1} F(q, \dot{q}, t) \\ A_{nb}(q)\lambda_{n,b} - A_{tb}(q)[\mu_b \xi_b] |\lambda_{n,b}| + w_b(q, \dot{q}, t) = 0 \end{cases} \quad (3.6)$$

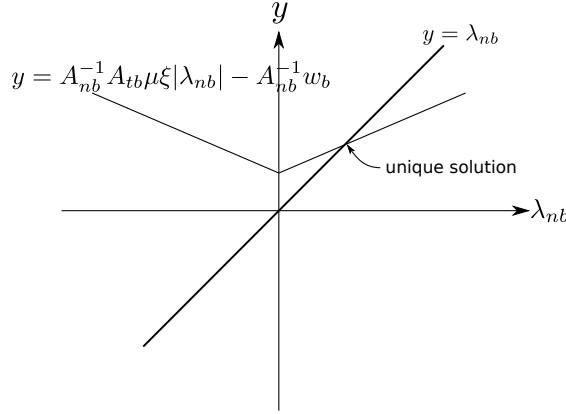


Figure 3.2: Illustration of the fixed point argument in the scalar case.

Since  $A_{nb}(q)$  can be inverted,  $\lambda_{n,b}$  is a solution of (3.2) if and only if

$$\lambda_{n,b} = A_{nb}^{-1} A_{tb} [\mu_b \xi_b] |\lambda_{n,b}| - A_{nb}^{-1} w_b.$$

Let  $T(x) := A_{nb}^{-1} A_{tb} [\mu \xi] |x| - A_{nb}^{-1} w_b$ , we will show that the mapping  $T(\cdot)$  is contracting under (3.4). Indeed one has:

$$\begin{aligned} \|T(x) - T(y)\|_2 &= \|A_{nb}^{-1} A_{tb} [\mu_b \xi_b] (|x| - |y|)\|_2 \leq \|A_{nb}^{-1} A_{tb} [\mu_b \xi_b]\|_2 \| |x| - |y| \|_2 \\ &\leq \|A_{nb}^{-1} A_{tb} [\mu_b \xi_b]\|_2 \|x - y\|_2 \end{aligned} \tag{3.7}$$

for any  $x, y \in \mathbb{R}^{m_b}$ . By the property of induced matrix norms [28, Proposition 9.4.3], and the fact that  $\|[a]\|_2 = \max_i |a_i|$ , we have  $\max_{1 \leq i \leq m} \mu_i < \frac{1}{\|A_{nb}^{-1}\|_2 \|A_{tb}\|_2}$  which implies  $\|A_{nb}^{-1} A_{tb} [\mu \xi]\|_2 < 1$ . Hence  $T(\cdot)$  is contracting. Using the Banach fixed point theorem one concludes that a unique solution  $\lambda_{n,b}$  to (3.6) exists. Existence and uniqueness of the acceleration in (3.6) follows from that of  $\lambda_{n,b}$ . ■

In Figure 3.2 we illustrate the fixed point argument of the proof in the scalar case. In the scalar case the operator  $T$  is simply of the form  $T(\lambda_{n,b}) = a|\lambda_{n,b}| + b$  for some given scalars  $a, b$ . The graph of the operator  $T$  describes the boundary of a cone, which is oriented upwards or downwards depending on the sign of  $a$ . The equation for the bilateral contact forces admits a unique solution if the graph of  $T$  intersects the line  $y = \lambda_{n,b}$  only once. This clearly happens for any value of  $b$  if  $|a| < 1$ .

The upper bound for the friction coefficients in Proposition 3.1.1 is necessary and sufficient in the case of a single contact. In the general case (several contact points) it is only sufficient. Depending on the mechanical system under consideration the friction bound may or may not be conservative. Note however, that the actual critical friction bound over which Painlevé-like paradoxes occur may, in fact, be arbitrarily small [10].

The following result is similar to Proposition 3.1.1. It has a stronger requirement on the friction bounds, but the proof uses complementarity theory rather than a fixed point argument and the decomposition introduced is key in decoupling frictional and frictionless parts in the mixed unilateral/bilateral case studied in Section 3.3.

**Proposition 3.1.2.** *Let  $(q, \dot{q})$  be given at a time instant  $t$ . Let  $A_{nb}(q)$  be positive definite (or equivalently  $\nabla h_{n,b}(q)$  of full rank). Suppose that all contacts are sliding ( $v_{t,b,i} \neq 0$ ) and that the friction coefficients  $\mu_i$  satisfy*

$$\max_{1 \leq i \leq m} \mu_i < \frac{1}{3} \mu_{max}^b(q)$$

where  $\mu_{\max}^b(q)$  is defined in (3.5). Then the bilateral sliding friction problem (3.2) has a unique solution  $(\dot{q}, \lambda_{n,b})$ .

**Proof:** The proof consists of three steps: reformulating (3.3) as a horizontal LCP (hLCP<sup>2</sup>), casting the hLCP matrix as a perturbation of a positive definite matrix and applying Corollary A.4.2.1 in Appendix A.4. Following [26] let us introduce  $\lambda^+ \triangleq \frac{|\lambda_{n,b}| + \lambda_{n,b}}{2}$  and  $\lambda^- \triangleq \frac{|\lambda_{n,b}| - \lambda_{n,b}}{2}$ . Solving (3.3) boils down to find  $\lambda^+$  and  $\lambda^-$  such that

$$\begin{cases} (A_{nb}(q) - A_{tb}(q)[\mu_b \xi_b])\lambda^+ - (A_{nb}(q) + A_{tb}(q)[\mu_b \xi_b])\lambda^- + w_b(q, \dot{q}, t) = 0 \\ 0 \leq \lambda^+ \perp \lambda^- \geq 0 \end{cases} \quad (3.8)$$

As pointed out in [26], a solution to the hLCP (3.8) exists and is unique if and only if  $A_{nb} - A_{tb}[\mu_b \xi_b]$  is invertible and  $(A_{nb} - A_{tb}[\mu_b \xi_b])^{-1}(A_{nb} + A_{tb}[\mu_b \xi_b])$  is a P-matrix. The upperbound on the friction coefficient assures

$$\|A_{nb}^{-1} A_{tb}[\mu_b \xi_b]\|_2 \leq \|A_{nb}^{-1} A_{tb}\|_2 \|[\mu_b]\|_2 \leq \frac{1}{3} < 1,$$

securing that  $I - A_{nb}^{-1} A_{tb}[\mu_b \xi_b]$  is non-singular and that its inverse can be written using Taylor series expansion. The hLCP matrix, whose positive definiteness we want to enforce, can be rewritten as

$$\begin{aligned} (A_{nb} - A_{tb}[\mu_b \xi_b])^{-1}(A_{nb} + A_{tb}[\mu_b \xi_b]) &= (I - A_{nb}^{-1} A_{tb}[\mu_b \xi_b])^{-1} A_{nb}^{-1} A_{nb}(I + A_{nb}^{-1} A_{tb}[\mu_b \xi_b]) \\ &= (\sum_{k=0}^{+\infty} (A_{nb}^{-1} A_{tb}[\mu_b \xi_b])^k)(I + A_{nb}^{-1} A_{tb}[\mu_b \xi_b]) \\ &= I + 2 \underbrace{\sum_{k=1}^{+\infty} (A_{nb}^{-1} A_{tb}[\mu_b \xi_b])^k}_{\triangleq K_\mu(q, \dot{q})} \end{aligned}$$

From Corollary A.4.2.1 in Appendix A.4, it follows that if  $\|K_\mu(q, \dot{q})\|_2 < 1$  then  $I + K_\mu(q, \dot{q}) \succ 0^3$ . The upperbound  $\frac{1}{3}\mu_{\max}^b(q)$  is chosen specifically so that this condition holds:  $\|K_\mu(q, \dot{q})\| \leq 2 \sum_{k=1}^{+\infty} \|A_{nb}^{-1} A_{tb}\|^k \|[\mu_b]\|^k < 2 \sum_{k=1}^{+\infty} \|A_{nb}^{-1} A_{tb}\|^k \frac{1}{3^k \|A_{nb}^{-1} A_{tb}\|^k} < 1$ . ■

**Example: A bilaterally constrained system with two frictional contacts.** We consider now the so-called Painlevé-Klein system depicted in Figure 1.1(d), where a rod is subjected to two bilateral frictional constraints with friction coefficients  $\mu_1, \mu_2$ . Proposition 3.1.1 can be used to study this system. It has been long known [8] that the contact LCP of this system has a unique solution for the multipliers, if and only if  $|\mu_1 - \mu_2| < \frac{2}{\tan \theta}$ ; this necessary and sufficient condition can be thought of as defining a region in the  $(\mu_1, \mu_2)$  parameter plane and it is depicted in dashes in Figure 3.3. In contrast, the merely sufficient condition of Proposition 3.1.1 defines a smaller region in the friction parameter plane (the dotted square in Figure 3.3). The Proposition states that if  $\max(\mu_1, \mu_2) < \frac{\sigma_{\min}(A_{nb}(\theta))}{\sigma_{\max}(A_{tb}(\theta))}$ , then the multiplier vector exists and is unique. Explicitly computing the singular values for a rod of length 1m and mass 1kg yields

$$\max(\mu_1, \mu_2) < \mu_{\max}^b(\theta) \triangleq \frac{2}{3\sqrt{-\cos(4\theta) + 1}} \sqrt{-(3\cos^2(\theta) + 1)|3\cos^2(\theta) - 1| + 9\cos^4(\theta) + 1},$$

<sup>2</sup>A horizontal LCP is just an MLCP with the particular structure of (3.8), see [6, Definition 12.23].

<sup>3</sup>A positive definite matrix is also a P-matrix

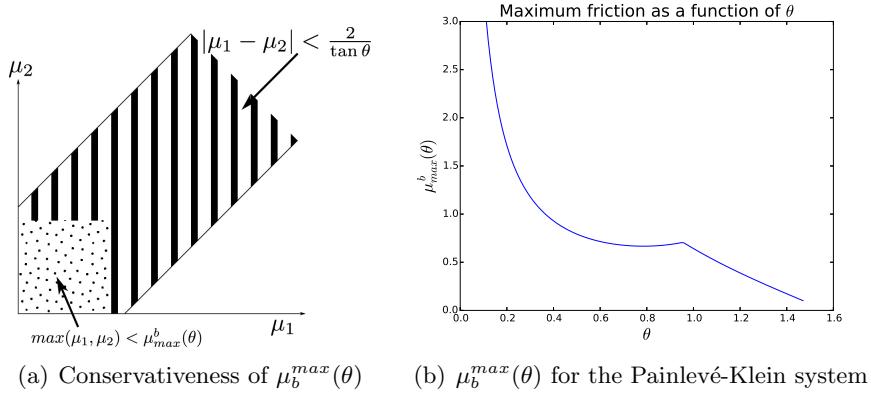


Figure 3.3: Friction bounds for the Painlevé-Klein system with two frictional guides.

and it can be checked numerically that  $\mu_b^{max}(\theta) < \frac{2}{\tan \theta}$  for all  $\theta \in ]0; \pi/2[$ . In Figure 3.3(b)  $\mu_b^{max}(\theta)$  is plotted as a function of  $\theta$ . If the system is at a certain state  $\theta$  and the friction at each guide is less than the value  $\mu_b^{max}(\theta)$  on the curve, then the contact force exists and is unique.

**Flexible Painlevé-Klein system** In this example we consider two rods of same mass and length coupled by a linear spring, constrained to be aligned and subject to the same two bilateral, frictional constraints as in the previous example, see Figure 1.1(e). The additional degree of freedom is  $d$  (the distance between the two rods), and  $q = (x, y, d, \theta)^T$ . To see how this change in the kinematics affects the friction bound of Proposition 3.1.1,  $\mu_b^{max}(q)$  is numerically computed for different configurations, see Figure 3.4. The matrices which are needed to compute the friction bound are:

$$M(q) = \begin{bmatrix} m & 0 & \frac{m}{2} \cos(\theta) & -\frac{m}{2}(d+l)\sin(\theta) \\ 0 & m & \frac{m}{2}\sin(\theta) & \frac{m}{2}(d+l)\cos(\theta) \\ \frac{m}{2}\cos(\theta) & \frac{m}{2}\sin(\theta) & \frac{m}{2} & 0 \\ -\frac{m}{2}(d+l)\sin(\theta) & \frac{m}{2}(d+l)\cos(\theta) & 0 & \frac{m}{96}(48d^2 + 72dl + 32l^2) \end{bmatrix},$$

$$\nabla h_{n,b}(q) = \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 0 & \sin(\theta) \\ 0 & (l+d)\cos(\theta) \end{bmatrix} \text{ and } H_{t,b}(q) = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 0 & \cos(\theta) \\ 0 & -(l+d)\sin(\theta) \end{bmatrix}.$$

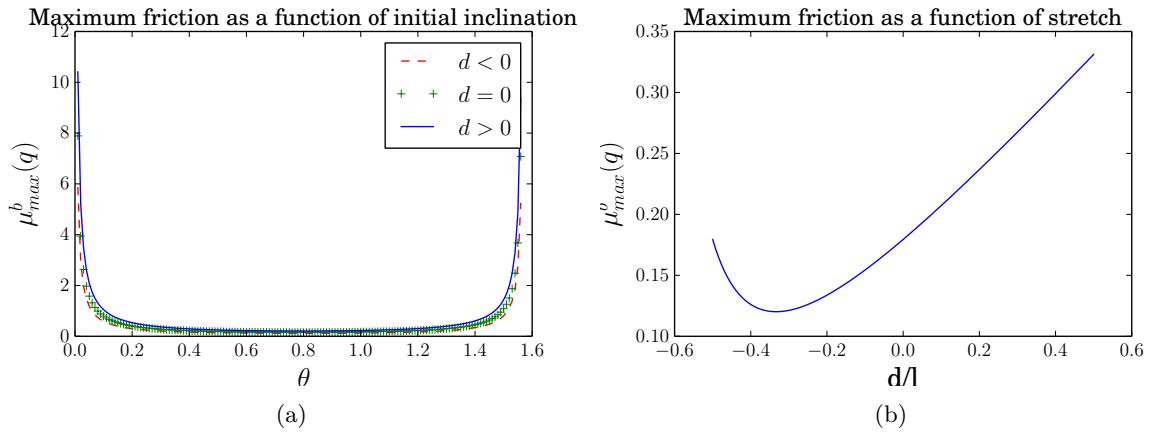


Figure 3.4: (a)  $d = -1.2l, l, 1.2l$ , (b)  $\theta = \pi/3$

Figure 3.4 illustrates the fact that a change in  $\theta$  or  $d$  influence the friction bound's conservativeness. One can also study the system in reduced coordinates and obtain a closed form expression for the necessary and sufficient conditions of existence and uniqueness of  $\lambda_{n,1}$ . The multiplier  $\lambda_{n,1}$  exists and is unique if and only if  $\mu_1 < \frac{4-3\cos^2\theta}{3\sin\theta\cos\theta}$ . Therefore, just as in the classic Painlevé example, no singularities occur when the friction coefficient is under 4/3 [7].

**Articulated Painlevé-Klein system** We consider an articulated Painlevé-Klein mechanism where both ends are subject to bilateral constraints with Coulomb friction. The rigid rod is replaced by two rods of same size and mass, linked together by a rotational joint, as in Figure 1.1(f). The friction bound of Proposition 3.1.1 is computed to understand how the coupling of the constraints plays a role on the criterion's conservativeness. Here  $q = (x_1, y_1, x_2, y_2, \theta_1, \theta_2)^T$ . The matrices which are needed to compute the friction coefficient upper bound are

$$M(q) = \begin{pmatrix} m & 0 & -\frac{3l}{8}m\sin(\theta_1) & -\frac{lm}{8}\sin(\theta_2) \\ 0 & m & \frac{3l}{8}m\cos(\theta_1) & \frac{lm}{8}\cos(\theta_2) \\ -\frac{3l}{8}m\sin(\theta_1) & \frac{3l}{8}m\cos(\theta_1) & \frac{l^2m}{6} & \frac{l^2m}{16}\cos(\theta_1 - \theta_2) \\ -\frac{lm}{8}\sin(\theta_2) & \frac{lm}{8}\cos(\theta_2) & \frac{l^2m}{16}\cos(\theta_1 - \theta_2) & \frac{l^2m}{24} \end{pmatrix},$$

$$\nabla h_{n,b}(q) = \begin{pmatrix} 0 & 0 \\ 1 & 1 \\ 0 & \frac{l}{2}\cos\theta_1 \\ 0 & \frac{l}{2}\cos\theta_2 \end{pmatrix} \text{ and } H_{t,b}(q) = \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 0 & -\frac{l}{2}\sin\theta_1 \\ 0 & -\frac{l}{2}\sin\theta_2 \end{pmatrix}. \text{ We numerically computed}$$

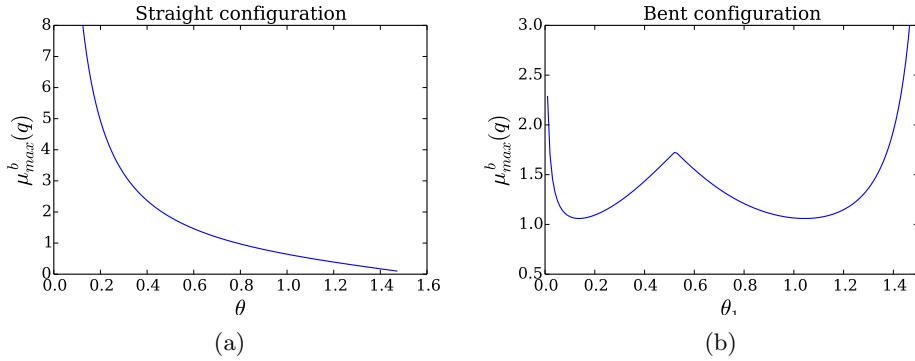


Figure 3.5: Evolution of  $\mu_{max}^b(q)$  for (a)  $q = (0, 0, \theta, \theta)$ , (b)  $q = (0, 0, \theta_1, \theta_2 = \arcsin(2D/l - \sin\theta_1))$

$\mu_{max}^b(q)$  for a variety of cases. Figure 3.5(a) depicts how  $\mu_{max}^b(q)$  changes for a set of configurations corresponding to the case where the rod is perfectly straight, keeping every other parameter fixed. We observe that as the system approaches a vertical configuration the normal couplings become stronger and the admissible set of friction values decreases. Figure 3.5(b) depicts how  $\mu_{max}^b(q)$  changes for initial configurations corresponding to a bent rod. In contrast with the classical Painlevé-Klein system, the relative orientation of the rods with respect to the guides may evolve during the motion. This is also the case of the unilateral sliding rod system, and it may create additional dynamical features [7].

### 3.1.2 Mixed frictional/frictionless contacts

Let us now assume that all  $m_b$  constraints are bilateral and that  $m_b^0$  of them are frictionless while  $m_b^\mu$  of them are subject to sliding friction. We rewrite (3.1) in its index 1 form:

$$\begin{cases} M(q)\ddot{q} + F(q, \dot{q}, t) = \nabla h_{n,b}^0(q)\lambda_{n,b}^0 + \nabla h_{n,b}^\mu(q)\lambda_{n,b}^\mu + H_{t,b}^\mu(q)[\mu_b][\xi_b][\text{sgn}(\lambda_{n,b}^\mu)]\lambda_{n,b}^\mu \\ \nabla h_{n,b}^0(q)^T\ddot{q} + \frac{d}{dt}(\nabla h_{n,b}^0(q)^T)\dot{q} = 0, \quad \nabla h_{n,b}^\mu(q)^T\ddot{q} + \frac{d}{dt}(\nabla h_{n,b}^\mu(q)^T)\dot{q} = 0 \end{cases} \quad (3.9)$$

In a matrix form one gets the extension of (3.2):

$$\begin{pmatrix} M(q) & -\nabla h_{n,b}^0(q) & -\nabla h_{n,b}^\mu(q) + H_{t,b}^\mu(q)[\mu_b][\xi_b][\text{sgn}(\lambda_{n,b}^\mu)] \\ \nabla h_{n,b}^0(q)^T & 0 & 0 \\ \nabla h_{n,b}^\mu(q)^T & 0 & 0 \end{pmatrix} \begin{pmatrix} \ddot{q} \\ \lambda_{n,b}^0 \\ \lambda_{n,b}^\mu \end{pmatrix} = \begin{pmatrix} -F(q, \dot{q}, t) \\ -\frac{d}{dt}(\nabla h_{n,b}^0(q)^T)\dot{q} \\ -\frac{d}{dt}(\nabla h_{n,b}^\mu(q)^T)\dot{q} \end{pmatrix} \quad (3.10)$$

Similarly to the all frictional case, assuming  $M(q) \succ 0$  and that the non-frictional constraints are independent, one obtains a reduced system where the only unknown is  $\lambda_{n,b}^\mu$ . Under these assumptions, system (3.9) is equivalent to (the argument  $q$  is dropped):

$$\begin{cases} \ddot{q} = -M^{-1}F(q, \dot{q}, t) + M^{-1}\nabla h_{n,b}^0\lambda_{n,b}^0 + M^{-1}\nabla h_{n,b}^\mu\lambda_{n,b}^\mu - M^{-1}H_{t,b}^\mu[\mu_b\xi_b]\left|\lambda_{n,b}^\mu\right| \\ \lambda_{n,b}^0 = -(A_{nb}^{00})^{-1}A_{nb}^{0\mu}\lambda_{nb}^\mu + (A_{nb}^{00})^{-1}A_{tb}^{0\mu}[\mu_b\xi_b]\left|\lambda_{n,b}^\mu\right| + (A_{nb}^{00})^{-1}((\nabla h_{n,b}^0)^T M^{-1}F - \frac{d}{dt}((\nabla h_{n,b}^0)^T)\dot{q}) \\ \left(A_{nb}^{\mu\mu} - (A_{nb}^{0\mu})^T(A_{nb}^{00})^{-1}A_{nb}^{0\mu}\right)\lambda_{n,b}^\mu - \left(A_{tb}^{\mu\mu} - (A_{nb}^{0\mu})^T(A_{nb}^{00})^{-1}A_{tb}^{0\mu}\right)[\mu_b\xi_b]\left|\lambda_{n,b}^\mu\right| + \tilde{w}_b = 0, \end{cases} \quad (3.11)$$

where  $A_{nb}^{00}(q) = \nabla h_{n,b}^0(q)^T M(q)^{-1} \nabla h_{n,b}^0(q)$  is the matrix of normal/normal frictionless couplings,  $A_{nb}^{0\mu}(q) = \nabla h_{n,b}^0(q)^T M(q)^{-1} \nabla h_{n,b}^\mu(q)$  is the matrix of normal frictionless/normal frictional couplings,  $A_{tb}^{0\mu}(q) = \nabla h_{n,b}^\mu(q)^T M(q)^{-1} H_{t,b}^\mu(q)$  is the matrix of normal frictionless/tangential frictional couplings,  $A_{tb}^{\mu\mu}(q) = \nabla h_{n,b}^\mu(q)^T M(q)^{-1} H_{t,b}^\mu(q)$  is the matrix of normal frictional/tangential frictional couplings, and finally  $\tilde{w}_b(q, \dot{q}, t) \triangleq A_{nb}^{0\mu}(A_{nb}^{00})^{-1}((\nabla h_{n,b}^0)^T M^{-1}F - \frac{d}{dt}((\nabla h_{n,b}^0)^T)\dot{q}) - (\nabla h_{n,b}^\mu)^T M^{-1}F - \frac{d}{dt}((\nabla h_{n,b}^\mu)^T)\dot{q}$  is the vector collecting all remaining terms in the last equation. Hence the problem is reduced to studying essentially the same equation as (3.3) in the all frictional case (which was itself the extension of (2.5)):

$$\tilde{A}_{nb}(q)\lambda_{n,b}^\mu - \tilde{A}_{tb}(q)[\mu_b\xi_b]\left|\lambda_{n,b}^\mu\right| + \tilde{w}_b(q, \dot{q}, t) = 0, \quad (3.12)$$

where  $\tilde{A}_{nb} \triangleq A_{nb}^{\mu\mu} - (A_{nb}^{0\mu})^T(A_{nb}^{00})^{-1}A_{nb}^{0\mu}$  and  $\tilde{A}_{tb} \triangleq A_{tb}^{\mu\mu} - (A_{nb}^{0\mu})^T(A_{nb}^{00})^{-1}A_{tb}^{0\mu}$ . It is important to observe that  $\tilde{A}_{nb}(q)$  is nothing but the Schur complement of  $A_{nb}^{\mu\mu}(q)$  in the matrix

$$(\nabla h_{n,b}^0 \ \nabla h_{n,b}^\mu)^T M^{-1} (\nabla h_{n,b}^0 \ \nabla h_{n,b}^\mu).$$

So  $\tilde{A}_{nb}(q) \succ 0$  and  $A_{nb}^{00}(q) \succ 0$  if and only if  $(\nabla h_{n,b}^0 \ \nabla h_{n,b}^\mu)^T M^{-1} (\nabla h_{n,b}^0 \ \nabla h_{n,b}^\mu) \succ 0$ . Hence, if  $(\nabla h_{n,b}^0(q) \ \nabla h_{n,b}^\mu(q))$  is full rank then  $\tilde{A}_{nb}(q)$  is invertible. We deduce the following result.

**Proposition 3.1.3.** *Let  $(q, \dot{q})$  be given at a time instant  $t$ . Assume that the matrix*

$$(\nabla h_{n,b}^0(q) \ \nabla h_{n,b}^\mu(q))^T M(q)^{-1} (\nabla h_{n,b}^0(q) \ \nabla h_{n,b}^\mu(q)) \quad (3.13)$$

is positive definite. Suppose that all frictional contacts are sliding ( $v_{t,b,i} \neq 0$ ) and that the friction coefficients  $\mu_i$  satisfy

$$\max_{1 \leq i \leq m_b^\mu} \mu_i < \frac{\sigma_{\min}(\tilde{A}_{nb}(q))}{\sigma_{\max}(\tilde{A}_{tb}(q))}.$$

Then the solution  $(\ddot{q}, \lambda_{n,b}^0, \lambda_{n,b}^\mu)$  of system (3.9) exists and is unique.

The proof is almost the same as that of Proposition 3.1.1: existence and uniqueness are obtained by a fixed point argument, and the bound on the matrix norm guarantees that a piecewise affine operator is contracting. The assumption (3.13) is there to ensure that  $\tilde{A}_{nb}(q)$  is invertible.

**Example: The Painlevé-Klein system with one frictional guide and one frictionless guide.** Let us we study the classical example of Painlevé-Klein within our framework of frictional/frictionless bilateral constraints. The system consists of a rigid rod with two prismatic joints as in Figure 3.6. The lower guide is rough (it is acted on by Coulomb friction), the upper guide is frictionless. The external forces are horizontal pulling forces  $F_1$  and  $F_2$ . The dynamics (3.9) is given by:

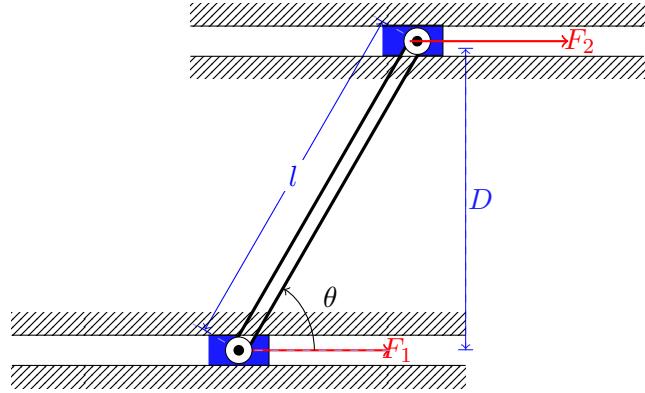


Figure 3.6: The Painlevé-Klein system

$$\begin{cases} m\ddot{x} + F_1 + F_2 &= -\mu_1 \xi_b |\lambda_{nb}^\mu| \\ m\ddot{y} &= \lambda_{n,b}^\mu + \lambda_{n,b}^0 \\ \frac{ml^2}{12}\ddot{\theta} + \frac{l}{2}\sin\theta(F_1 - F_2) &= \frac{l}{2}\cos\theta(\lambda_{n,b}^0 - \lambda_{n,b}^\mu) - \frac{l}{2}\sin\theta\mu_1\xi_b |\lambda_{nb}^\mu| \\ h_{n,b}^\mu(\theta) &= y - \frac{l}{2}\sin(\theta) \\ h_{n,b}^0(\theta) &= y + \frac{l}{2}\sin(\theta) - D. \end{cases}$$

The matrices involved in the analysis of the system are:  $M = \text{diag}(m, m, \frac{ml^2}{12})$ ,  $\nabla h_{n,b}^\mu(\theta) = (0, 1, -\frac{l}{2}\cos(\theta))^T$ ,  $\nabla h_{n,b}^0(\theta) = (0, 1, \frac{l}{2}\cos(\theta))^T$ ,  $H_{t,b}(\theta) = (1, 0, \frac{l}{2}\sin(\theta))^T$ ,  $\tilde{A}_{nb}(\theta) = \frac{12\cos^2(\theta)}{m(3\cos^2(\theta)+1)}$ ,  $\tilde{A}_{tb}(\theta) = \left[ \frac{6\sin(2\theta)}{m(3\cos(2\theta)+5)} \right]$ . Equation (3.12) boils down to a scalar equation of the form  $\tilde{A}_{nb}(\theta)\lambda_{n,b}^\mu - \tilde{A}_{tb}(\theta)\mu_1\xi_b |\lambda_{nb}^\mu| + \tilde{w}_b(q, \dot{q}, t) = 0$ . From Proposition 3.1.3 it follows that if  $\mu_1 < \frac{2}{\tan\theta}$  then  $(\ddot{x}, \ddot{y}, \ddot{\theta}, \lambda_{n,b}^\mu, \lambda_{n,b}^0)$  exists and is unique. The condition turns out to be not only sufficient but also necessary.

## 3.2 Unilaterally constrained systems

Let us now turn to the case where all constraints are unilateral and subject to sliding friction. The unilateral sliding friction problem then consists in finding the accelerations and contact forces  $(\ddot{q}, \lambda_{n,u})$  solutions of the MLCP:

$$\begin{cases} M(q)\ddot{q} + F(q, \dot{q}, t) = (\nabla h_{n,u}(q) - H_{t,u}(q)[\mu_u] [\xi_u])\lambda_{n,u} \\ 0 \leq \nabla h_{n,u}(q)^T \ddot{q} + \frac{d}{dt}(\nabla h_{n,u}(q)^T) \dot{q} \perp \lambda_{n,u} \geq 0. \end{cases} \quad (3.14)$$

If  $M(q) \succ 0$  then the problem is reduced to an LCP. Let us introduce the matrix:

$$A_{u,\mu,\xi}(q) := \nabla h_{n,u}(q)^T M(q)^{-1} (\nabla h_{n,u}(q) - H_{t,u}(q)[\mu_u] [\xi_u]) \quad (3.15)$$

$$= A_{nu}(q) - A_{tu}(q)[\mu_u \xi_u], \quad (3.16)$$

where  $A_{tu}(q) := \nabla h_{n,u}(q)^T M(q)^{-1} H_{t,u}(q)$ ,  $A_{nu}$  as defined in (2.15) and  $\xi_{u,i} = \text{sgn}(v_{t,u,i})$ ,  $v_{t,u,i} \neq 0$  for  $1 \leq i \leq m_u$ . Then the frictional contact LCP, which extends (2.15), is given by:

$$0 \leq \lambda_{n,u} \perp A_{u,\mu,\xi}(q)\lambda_{n,u} + w_u(q, \dot{q}, t) \geq 0 \quad (3.17)$$

with  $w_u(q, \dot{q}, t)$  as defined in (2.15).

**Proposition 3.2.1.** *Let  $(q, \dot{q})$  be given at some time instant  $t$ , that all  $m_u$  constraints are unilateral, active at  $t$ , and in the sliding mode of Coulomb's friction. Let  $A_{nu}(q)$  be positive definite and suppose that:*

$$\max_{1 \leq i \leq m_u} \mu_i < \mu_{\max}^u(q) := \frac{\sigma_{\min}(A_{nu}(q))}{\sigma_{\max}(A_{tu}(q))}. \quad (3.18)$$

*Then the matrix  $A_{u,\mu,\xi}(q)$  is positive definite (hence a P-matrix), and the contact LCP (3.17) always has a unique solution  $\lambda_{n,u}$  for any given  $w_u(q, \dot{q}, t)$ .*

*Proof.* Let the hypotheses in the proposition hold. Let  $x \in \mathbb{R}^{m_u}$  be non zero. We have that  $x^T(A_{tu}(q)[\mu_u \xi_u])x \leq \|\mu_u \xi_u\| \|A_{tu}(q)\| \|x\|^2$ , and thus

$$x^T(A_{tu}(q)[\mu_u \xi_u])x \leq \max_{1 \leq i \leq m_u} \mu_i \sigma_{\max}(A_{tu}(q)) \|x\|^2.$$

Combining the latter inequality with  $x^T A_{nu}(q)x \geq \sigma_{\min}(A_{nu}(q)) \|x\|^2$  we get that

$$x^T A_{u,\mu,\xi}(q)x \geq \left( \sigma_{\min}(A_{nu}(q)) - \max_{1 \leq i \leq m_u} \mu_i \sigma_{\max}(A_{tu}(q)) \right) \|x\|^2.$$

Clearly, the bound on the friction coefficients in (3.18) guarantees that for all non zero  $x \in \mathbb{R}^{m_u}$  the quadratic product  $x^T A_{u,\mu,\xi}(q)x$  is strictly positive, so that  $A_{u,\mu,\xi}(q)$  is positive definite (albeit non symmetric). Hence the matrix  $A_{u,\mu,\xi}(q)$  is a P-matrix and the contact LCP (3.17) has a unique solution  $\lambda_{n,u}$  for any given  $w_u(q, \dot{q}, t)$ .  $\square$

**Remark** The fact that (3.18) guarantees that  $A_{u,\mu,\xi}(q) \succ 0$ , may also be proved using Corollary A.4.2.1 in Appendix A.4, which was our original approach. The above proof is due to [11], we include it as it makes the bound on the friction coefficients appear more naturally.

The friction bound in (3.18) provides a sufficient condition for  $A_{u,\mu,\xi}(q)$  to be a P-matrix. The bound is easy to compute even for a system with a very large number of contacts. Indeed the largest and smallest singular values of a matrix may be computed efficiently using the Lanczos algorithm. This is in contrast with trying to determine if the matrix  $A_{u,\mu,\xi}(q)$  is a P-matrix by determining the sign of its  $2^{m_u} - 1$  principal minors [47], which would be computationally prohibitive for a large number of contacts.

Next we examine the classical example usually used to illustrate the Painlevé paradox.

**The Classical Example of Painlevé** We now add Coulomb's friction to the mechanical system in Figure 1.1(b) and point out how the equations of motion are affected. The system under consideration is the classical example of Painlevé of a rigid rod sliding against a frictional plane. Its dynamics is given by:

$$\begin{cases} m\ddot{x} = -\mu\xi\lambda_{nu} \\ m\ddot{y} = -mg + \lambda_n \\ \frac{ml^2}{12}\ddot{\theta} = -\frac{l}{2}\cos\theta\lambda_n - l/2\sin\theta\mu\xi\lambda_{nu} \\ 0 \leq h_{nu}(\theta) = y - \frac{l}{2}\sin(\theta)\perp\lambda_{nu} \geq 0 \end{cases}$$

The additional matrices due to friction are  $H_{tu}(\theta) = (1, 0, \frac{l}{2}\sin(\theta))^T$ ,  $A_{tu}(\theta) = \frac{-3\sin\theta\cos\theta}{m}$ , and  $A_{u,\mu,\xi}(\theta) = \frac{1+3\cos^2(\theta)}{m} + \frac{3\sin(\theta)\cos(\theta)}{m}\mu\xi$ . Hence the LCP in (3.17) reads:

$$0 \leq \left( \frac{1+3\cos^2(\theta)}{m} + \frac{3\sin(\theta)\cos(\theta)}{m}\mu\xi \right) \lambda_{nu} + \frac{l}{2}\dot{\theta}^2 \sin\theta - g\perp\lambda_{nu} \geq 0. \quad (3.19)$$

The necessary and sufficient condition for a solution to exist (for any value of  $\frac{l}{2}\dot{\theta}^2 \sin\theta - g$ ) is that the scalar  $A_{u,\mu,\xi}(\theta)$  be positive. If we consider values of  $\theta$  in  $]0, \frac{\pi}{2}[$  then one sees that for rightward sliding ( $\xi = 1$ )  $A_{u,\mu,\xi}(\theta)$  is always positive, regardless of the magnitude of friction. Whereas for leftward sliding ( $\xi = -1$ ),  $A_{u,\mu,\xi}(\theta)$  becomes negative for high values of friction. Indeed solving for  $\mu$  in the scalar inequality  $A_{u,\mu,\xi}(\theta) < 0$  one finds the critical friction value beyond which paradoxes occur, and it coincides with that of Proposition 3.2.1, namely

$$\mu_u^{max}(\theta) = \frac{1+3\cos^2\theta}{3\sin\theta\cos\theta}.$$

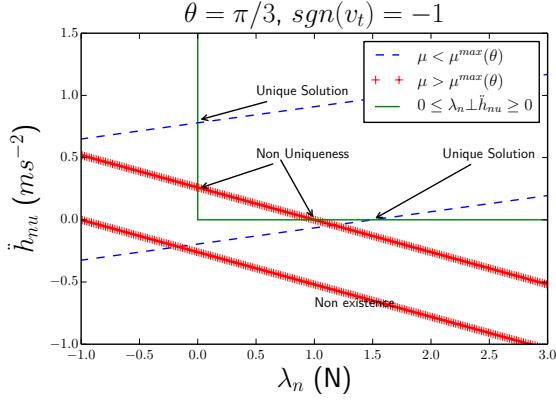
A graphical interpretation of these singular LCP situations is given on Figure 3.7(a), while the critical friction bound for the Painlevé system is plotted in Figure 3.7(b). It is seen that the frictional singularities occur only when the friction coefficient is larger than  $\frac{4}{3}$ . While some have argued that it is unrealistic to have such a large value of friction, it has been shown in [10] that the bound above which frictional singularities occur may in fact be arbitrarily low. The structure of the bound in (3.18) is in agreement with this as we can see that depending on the constraints and the mass matrix it may be made arbitrarily small.

### 3.3 Unilaterally/bilaterally constrained systems

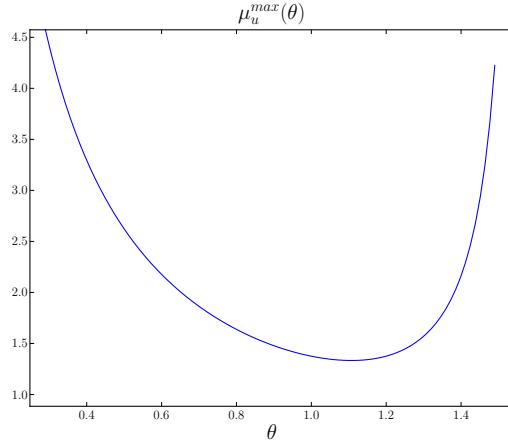
In this section we consider mechanical systems with both bilateral and unilateral constraints subject to sliding friction. Recall that in the frictionless case, if the mass matrix is invertible and the constraints linearly independent then the contact problem is well-posed and the unilateral contact force is the unique solution of the LCP (2.18). The matrix of the LCP (2.18),  $A_c(q)$  is symmetric positive definite under these conditions. We will see how the LCP matrix gets perturbed when friction is added and we will derive a bound on the friction coefficients so that the perturbed matrix remains a P-matrix.

#### All contacts with friction

Let us consider first the case where all contacts are frictional, i.e.,  $\mathcal{I}_b^0 = \mathcal{I}_u^0 = \emptyset$ . Let us consider (1.1) in its index 1 form, i.e., with the acceleration constraints  $\nabla h_{n,b}(q)^T \ddot{q} + \frac{d}{dt}(\nabla h_{n,b}(q)^T) \dot{q} = 0$



(a) Singular and non-singular situations for (3.19).



(b) Critical friction bound for the Painlevé example.

Figure 3.7: Critical friction and singular situations in the classic example of Painlevé

and  $0 \leq h_{n,u}(q) \perp \nabla h_{n,u}(q)^T \ddot{q} + \frac{d}{dt}(\nabla h_{n,u}(q)^T) \dot{q} \geq 0$ . The dynamics in (1.1) becomes :

$$\left\{ \begin{array}{l} M(q)\ddot{q} = \nabla h_{nb}(q)\lambda_{n,b} - H_{t,b}(q)[\mu_b \xi_b] |\lambda_{n,b}| + \nabla h_{nu}(q)\lambda_{n,u} - H_{t,u}(q)[\mu_u \xi_u] \lambda_{n,u} - F(q, \dot{q}, t) \\ \nabla h_{n,b}(q)^T \ddot{q} + \frac{d}{dt}(\nabla h_{n,b}(q)^T) \dot{q} = 0 \\ 0 \leq \lambda_{n,u} \perp \nabla h_{nu}(q)^T \ddot{q} + \frac{d}{dt}(\nabla h_{n,u}(q)^T) \dot{q} \geq 0. \end{array} \right. \quad (3.20)$$

By inverting the mass matrix,  $\ddot{q}$  can be expressed as a function of the multipliers. Introducing  $\lambda^+ = \frac{|\lambda_{n,b}| + \lambda_{n,b}}{2}$  and  $\lambda^- = \frac{|\lambda_{n,b}| - \lambda_{n,b}}{2}$  as in Proposition 3.1.2 casts the piecewise linearity induced by the absolute value into a complementarity formalism. The problem is thus transformed to determining whether for arbitrary parameter values and arbitrary vectors  $w_1, w_2$  the following

MLCP has a unique solution  $(\lambda^+(q, \dot{q}), \lambda^-(q, \dot{q}), \lambda_{n,u}(q, \dot{q}))$ <sup>4</sup>:

$$\begin{cases} (A_{nb} - A_{tb}[\mu_b \xi_b])\lambda^+ - (A_{nb} + A_{tb}[\mu_b \xi_b])\lambda^- + (A_{nbnu} - A_{nbtu}[\mu_u \xi_u])\lambda_{n,u} + w_1 = 0 \\ 0 \leq (A_{nunb} - A_{nutb}[\mu_b \xi_b])\lambda^+ - (A_{nunb} + A_{nutb}[\mu_b \xi_b])\lambda^- + (A_{nu} - A_{tu}[\mu_u \xi_u])\lambda_{n,u} + w_2 \perp \lambda_{n,u} \geq 0 \\ 0 \leq \lambda^+ \perp \lambda^- \geq 0, \end{cases} \quad (3.21)$$

where:

$$\begin{aligned} w_1(q, \dot{q}, t) &\triangleq -\nabla h_{n,b}(q)^T M(q)^{-1} F(q, \dot{q}, t) + \frac{d}{dt} (\nabla h_{n,b}(q)^T) \dot{q} \\ w_2(q, \dot{q}, t) &\triangleq -\nabla h_{n,u}(q)^T M(q)^{-1} F(q, \dot{q}, t) + \frac{d}{dt} (\nabla h_{n,u}(q)^T) \dot{q} \end{aligned}$$

possess the same structure as  $w_u(q, \dot{q}, t)$  in (2.2). Assume that the conditions of Proposition 3.1.2 hold. Then  $A_{nb}(q) - A_{tb}(q)[\mu_b \xi_b] \succ 0$  and the problem is to determine the existence and uniqueness of a solution  $(\lambda^-(q, \dot{q}), \lambda_{n,u}(q, \dot{q}))$  (for arbitrary  $q, \dot{q}$ ) of the following LCP (the argument  $q$  is dropped):

$$0 \leq \begin{pmatrix} \lambda^- \\ \lambda_{n,u} \end{pmatrix} \perp \underbrace{\begin{pmatrix} A^{-1}\bar{A} & -A^{-1}B \\ CA^{-1}\bar{A} - \bar{C} & D - CA^{-1}B \end{pmatrix}}_{\triangleq M_{ub}^\mu(q)} \begin{pmatrix} \lambda^- \\ \lambda_{n,u} \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \geq 0, \quad (3.22)$$

where  $A \triangleq A_{nb} - A_{tb}[\mu_b \xi_b]$ ,  $\bar{A} \triangleq A_{nb} + A_{tb}[\mu_b \xi_b]$ ,  $B \triangleq A_{nbnu} - A_{nbtu}[\mu_u \xi_u]$ ,  $C \triangleq A_{nunb} - A_{nutb}[\mu_b \xi_b]$ ,  $\bar{C} \triangleq A_{nunb} + A_{nutb}[\mu_b \xi_b]$ ,  $D \triangleq A_{nu} - A_{tu}[\mu_u \xi_u]$ ,  $z_1 \triangleq -A^{-1}w_1$  and  $z_2 \triangleq w_2 + Cz_1$ .

**Remark** When all the contacts are frictionless, the LCP in (3.22) has the form:

$$0 \leq \begin{pmatrix} \lambda^- \\ \lambda_{n,u} \end{pmatrix} \perp \underbrace{\begin{pmatrix} I & -A_{nb}^{-1}A_{nbnu} \\ 0 & A_{nu} - A_{nbnu}^T A_{nb}^{-1} A_{nbnu} \end{pmatrix}}_{\triangleq M_{ub}^0(q)} \begin{pmatrix} \lambda^- \\ \lambda_{n,u} \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \geq 0, \quad (3.23)$$

The second line is an LCP involving  $\lambda_{nu}$  alone and is the same as the LCP in (2.18). Thus its solution  $\lambda_{nu}$  exists and is unique provided that all the constraints are independent. Under this assumption it follows that  $\lambda^-$  also exists and is unique for any  $(z_1, z_2)$ : hence the matrix  $M_{ub}^0(q)$  of the LCP (3.23) is a P-matrix.

One has the following decomposition for the matrix in (3.22) into frictionless and frictional parts:  $M_{ub}^\mu(q) = M_{ub}^0(q) + P_\mu(q, \dot{q})$  where the frictional part  $\|P_\mu(q, \dot{q})\|$  tends to zero as the friction coefficient  $\mu_i$  goes to zero for each contact. Once again the matrix from the frictional LCP is obtained as a perturbation of the frictionless one (3.23). Here sufficiently small friction coefficients will guarantee that  $M_{ub}^\mu$  is a P-matrix. Let us provide some details on the matrix  $P_\mu(q, \dot{q})$ . We know from section 3.1 that  $A^{-1}\bar{A} = I + 2K_\mu(q, \dot{q})$  with  $K_\mu(q, \dot{q}) = \sum_{i=1}^\infty (A_{nb}^{-1}A_{tb}[\mu_b \xi_b])^i$ .

---

<sup>4</sup>Recall that  $A_{nunb}(q) \triangleq \nabla h_{n,u}(q)^T M(q)^{-1} \nabla h_{n,b}(q)$ ,  $A_{nutb}(q) \triangleq \nabla h_{n,u}(q)^T M(q)^{-1} H_{t,b}(q)$ ,  $A_{nbnu}(q) = A_{nunb}(q)^T$ ,  $A_{nbtu}(q) \triangleq \nabla h_{n,b}(q)^T M(q)^{-1} H_{t,u}(q)$ .

Using the same decomposition for the other blocks of  $M_{ub}^\mu(q)$ , one obtains

$$A^{-1}\bar{A} = I + 2K_\mu(q, \dot{q}) \quad (3.24)$$

$$-A^{-1}B = -A_{nb}^{-1}A_{nbnu} + (I + K_\mu(q, \dot{q}))A_{nb}^{-1}A_{nbtu}[\mu_u \xi_u] - K_\mu(q, \dot{q})A_{nb}^{-1}A_{nbnu} \quad (3.25)$$

$$CA^{-1}\bar{A} - \bar{C} = 0 - 2A_{nutb}[\mu_b \xi_b] + CK_\mu(q, \dot{q}) \quad (3.26)$$

$$\begin{aligned} D - CA^{-1}B &= A_{nu} - A_{nbnu}^T A_{nb}^{-1} A_{nbnu} - A_{tu}[\mu_u \xi_u] + A_{nutb}[\mu_b \xi_b] A_{nb}^{-1} A_{nbnu} \\ &\quad + C((I + K_\mu(q, \dot{q}))A_{nb}^{-1}A_{nbtu}[\mu_u \xi_u] - K_\mu(q, \dot{q})A_{nb}^{-1}A_{nbnu}) \end{aligned} \quad (3.27)$$

By Theorem A.4.1 in Appendix A.4, if

$$\|P_\mu(q, \dot{q})\|_2 < 1/\beta_2(M_{ub}^0(q))$$

then  $M_{ub}^\mu(q)$  is a P-matrix. One can therefore derive an implicit maximum value of all friction coefficients below which existence and uniqueness of a solution to problem (3.20) is guaranteed.

**Proposition 3.3.1.** *Let  $(q, \dot{q})$ ,  $F(q, \dot{q}, t)$  be given and  $v_{t,i} \neq 0$  for all  $i$ . Assume that  $M(q) \succ 0$  and  $A_{nb}(q) \succ 0$ . Suppose that all the constraints are independent, that the bilateral friction coefficients satisfy  $\max_{1 \leq i \leq m_b} \mu_i^b < \mu_{\max}^b(q)$ , and that  $\|P_\mu(q, \dot{q})\|_2 < \frac{1}{\beta_2(M_{ub}^0(q))}$ . Then  $M_{ub}^\mu(q)$  is a P-matrix, there exists a unique solution  $(\ddot{q}, \lambda_{nb}, \lambda_{nu})$  to the LCP in (3.22) for any  $z_1, z_2$ , and thus to the mixed sliding friction problem (3.20).*

One sees once again that there is no obvious quadratic problem that could be associated with the LCP in (3.22), because  $M_{ub}^\mu(q)$  is guaranteed to be neither symmetric nor  $\succeq 0$ .

**Example: A falling ladder.** The above Proposition 3.3.1 defines implicitly a friction bound  $\mu_{ub}^{\max}(q)$  below which frictional singularities are avoided. We compute it numerically for the example of a falling ladder in frictional bilateral contact with the wall and frictional unilateral contact with the ground as in Figure 3.8(a). The ladder is modeled as a rigid rod of length 2 m and of mass 1 kg. Figure 3.8(b) depicts the evolution of the friction bound as the rod slides from a vertical position to a horizontal one. It is seen for example that at the configuration  $\theta = \frac{\pi}{4}$  the friction bound is around 0.65. In Figure 3.8(c) we show, for the configuration  $\theta = \frac{\pi}{4}$ , the range of values  $(\mu_b, \mu_u)$  of the bilateral and unilateral friction coefficients for which the matrix  $M_{ub}^\mu(q)$  is a P-matrix. The black region of allowable friction coefficients is thus larger than the safety region defined by the proposition (the  $[0, 0.65]^2$  square). The criterion is thus conservative, as expected.

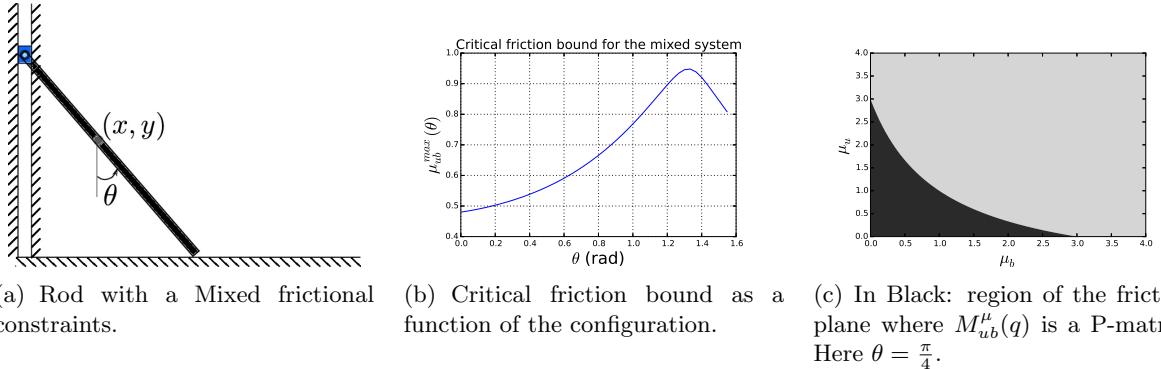


Figure 3.8: Critical friction bound for the falling ladder example.

It is important to mention that the numerical computation of the bound  $\mu_{ub}^{\max}(q)$  for a mixed system is not as straightforward as its purely unilateral or purely bilateral counterparts.

The quantity  $\beta_2(M_{ub}^0(q))$  is by definition the result of an optimization problem (see Theorem D.A.4.1 in the Appendix), for the sliding ladder example we solve it using a sequential least squares solver available in the SciPy library [48]. The matrix  $P_\mu$  is obtained as the difference  $M_{ub}^\mu - M_{ub}^0$ . Finally the bound  $\mu_{ub}^{max}(q)$  is computed by numerically solving for  $\mu$ , with a secant method, the scalar algebraic equation  $\|P_\mu(q, \dot{q})\|_2 = 1/\beta_2(M_{ub}^0(q))$ .

### Frictionless bilateral contacts

When the bilateral constraints are frictionless ( $\mu_{b,i} = 0$  for all  $0 \leq i \leq m_b$ ), one can easily express  $\lambda_{nb}$  as a function of  $\lambda_{nu}$  and equation (3.20) becomes

$$0 \leq \lambda_{nu} \perp (D - CA^{-1}B)\lambda_{nu} + z_2 \geq 0,$$

with  $z_2$  in (3.22),  $D - CA^{-1}B = A_{nu} - A_{nbnu}^T A_{nb}^{-1} A_{nbnu} + (A_{nbnu}^T A_{nb}^{-1} A_{nbtu} - A_{tu})[\mu_u \xi_u]$ , which is the sum of the matrix  $A_c(q)$  in (2.18), plus a frictional perturbation term. We know that  $A_c(q) \succ 0$  when all constraints are independent. The matrix  $D - CA^{-1}B$  is simpler than  $M_{ub}^\mu(q)$  in (3.22) since it is the perturbation of a symmetric positive definite matrix, rather than that of a P-matrix. One can use the results of Section 3.2.

**Proposition 3.3.2.** *Let  $(q, \dot{q})$ ,  $F(q, \dot{q}, t)$  be given and  $v_{t,i} \neq 0$  for all  $i$ . Suppose that all constraints are independent, that the bilateral constraints are frictionless ( $\mu_{b,i} = 0$  for all  $i$ ), and that*

$$\max_{1 \leq i \leq m_u} \mu_{u,i} < \frac{\sigma_{min}(A_{nu}(q) - A_{nbnu}(q)^T A_{nb}(q)^{-1} A_{nbnu}(q))}{\sigma_{max}(A_{tu}(q) - A_{nbnu}(q)^T A_{nb}(q)^{-1} A_{nbtu}(q))}. \quad (3.28)$$

*Then there exists a unique solution  $(\ddot{q}, \lambda_{nu}, \lambda_{nb})$  to the problem (3.20).*

It is noteworthy that (3.28) is the direct extension of (3.18), taking into account the distortion due to the presence of frictionless bilateral constraints. However, there is no clear relationship between  $\mu_{max}^u(q)$  in (3.18), and the upperbound in (3.28).

## 3.4 Conclusion

The main purpose of this part was to present in a progressive way the so-called contact problem (whose unknowns are the contact forces and the accelerations), in the frictionless and sliding Coulomb's friction cases. In both settings Lagrangian systems with bilateral, unilateral and mixed bilateral/unilateral constraints are analysed. Various points of view (linear complementarity problems, minimization problems, KKT conditions) are considered. The frictional case is treated as a perturbation of the frictionless one, and explicit criteria on the friction coefficients upper bound guaranteeing the contact problem's well-posedness, are given.

**Recapitulation** We now give a global view on the structures of the different contact problems we have been considering thus far. What we first wish to underline is that (for a non-singular mass matrix and full rank constraint Jacobian) the frictionless problems are well posed. The bilateral one as a linear system in (2.5)

$$A_{nb}(q)\lambda_{n,b} + w_b(q, \dot{q}, t) = 0,$$

the unilateral one as an LCP in (2.15)

$$0 \leq A_{nu}(q)\lambda_{n,u} + w_u(q, \dot{q}, t) \perp \lambda_{n,u} \geq 0,$$

and the mixed one as the unilateral LCP with a bilateral distortion in (2.18)

$$0 \leq A_c(q)\lambda_{n,u} + w_c(q, \dot{q})\perp\lambda_{n,u} \geq 0.$$

In the case of a full rank Jacobian and an invertible mass matrix all three matrices  $A_{nb}(q)$ ,  $A_{nu}(q)$ , and  $A_c(q)$  are symmetric positive-definite. One can solve uniquely for contact forces and deduce a unique acceleration. Then in the case of small friction this well posedness is preserved. The all sliding bilateral friction problem becomes the LCP

$$0 \leq (I + 2K_\mu(q, \dot{q}))\lambda^- + w_{bf}(q, \dot{q}, t)\perp\lambda^- \geq 0,$$

for some  $w_{bf}(q, \dot{q}, t)$ , where  $\|K_\mu(q, \dot{q}, t)\|$  tends to zero as the friction  $\mu$  goes to zero. For small enough friction the positivity of the identity matrix is preserved and the LCP is well posed. The unilateral all sliding friction problem becomes (3.17)

$$0 \leq A_{u,\mu,\xi}(q)\lambda_{n,u} + w_u(q, \dot{q}, t)\perp\lambda_{n,u} \geq 0.$$

For small friction  $\mu_u$  the well posedness of the contact problem is kept. As for the mixed all sliding friction problem one has an LCP of the form (3.22)

$$0 \leq (M_{ub}^0(q) + P_\mu(q, \dot{q}))\lambda + w(q, \dot{q}, t)\perp\lambda \geq 0,$$

where  $\|P_\mu(q, \dot{q})\|$  tends to zero as the friction  $\mu$  goes to zero. The matrix  $M_{ub}^0(q)$  is a P-matrix whose terms appear in the frictionless problems and the P-matrix property is also preserved when adding the small frictional perturbation term  $P_\mu(q, \dot{q})$ .

We have derived bounds on the friction coefficients which guarantee the well posedness of the frictional sliding contact problem. In the purely bilateral and purely unilateral problems the friction bounds are explicit and straightforward to compute. They require only the computation of some singular values of the system matrices. In contrast some more intricate numerical procedure is required to compute the bound in Proposition 3.3.1 for the mixed (unilateral/bilateral) case.

**Use of the friction bounds.** The derived critical friction bounds may be used for the explicit investigation of Painlevé like singularities in simple mechanical systems as those considered in the examples throughout this chapter. They may serve as well in the investigation of the dynamics of these systems as carried out in [7] and [10]. In mechanism design the friction bounds could be used as a certificate that the contact forces can be uniquely determined. They could perhaps be useful in the design of mechanical systems which explicitly avoid inconsistent configurations. Furthermore they may be used in diagnosing issues in the event-driven numerical simulation of larger multibody systems. In an event-driven method, if an event is detected for some configuration  $(q, \dot{q})$  of the mechanical system at time  $t$  the event is to be handled by computing the contact force and the acceleration of the system so as to advance the simulation in time. If the handling of the event fails then one can compute the friction bound for the given configuration and test whether the actual friction coefficients are above or below the bound. If the actual friction coefficients are below the bound then one can be sure that the failing of the numerical procedure is not due to the appearance of a friction induced singularity. If the actual friction is above the bound then in general one cannot conclude, as the bound yields only a sufficient condition for the existence and uniqueness of the contact force.

**Handling frictional singularities in numerical simulations.** As the Painlevé example and all the other examples considered in this part show, frictional singularities occur even in the simplest of systems and are by no means exceptional. In the presence of these singularities, event-driven methods fail to evolve the system in time, as any value of the acceleration and the contact forces will inevitably lead to the violation of the constraints. What one should do in such a scenario is to use a time-stepping scheme [49], rather than an event driven scheme. Although this is obvious to many experts in the field, event driven schemes are widespread in industry and further collaboration between academia and industry is required on this topic. An example of such a collaboration (between our team and Ansys) is the thesis of Haddouni [50], where a mixed event-driven/time-stepping integration scheme was proposed. The justification of why time stepping methods would work better in such inconsistent configurations is rooted in the work of Moreau [49], Stewart [51] and Genot and Brogliato [7]. In [49] it was empirically shown that time-stepping schemes are successful in the simulation of Painlevé paradoxes, which are referred to as *frictional paroxysms* by Moreau. In [51] Stewart shows that, for the Painlevé example, although the forces may not be computed in inconsistent configurations, an impulsive solution to the equations of motion does exist. In [7] a detailed dynamical study of the Painlevé system near a singular point was carried out and it was shown that while the contact force may blow up and be unbounded, its integral (the impulse) remains bounded. Since in time-stepping schemes the unknowns are the impulses rather than the forces it is normal that time stepping schemes are better at simulating mechanical systems which in continuous time are subject to frictional paroxysms. Let us note however that the use of a time stepping scheme doesn't solve all solvability issues as it shown in the examples of [52] and [53].

**Outlook.** Perhaps some similar bounds can be derived on the friction coefficients which guarantee the existence and uniqueness of the velocities and the contact impulses in the time-stepping discretization of a mechanical system. It would be highly desirable to apply the methods developed in this part to the analysis of the discrete time frictional contact problem.

## **Part II**

# **The statics of elastic rods from an Optimal Control point of view**

# Chapter 4

## Introduction

**Overview.** This part of the thesis deals with the statics of thin elastic rods. In Chapter 5 we present the continuous problem of rod statics. In Chapter 6 we address the problem of computing stable static equilibria of Kirchhoff rods under different boundary conditions and possibly subject to contact constraints. Our proposed approach relies on formulating the continuous problem as an Optimal Control Problem (OCP) on the Lie group  $SE(3)$  and discretizing it using direct methods of numerical optimal control. We will see that the Optimal Control point of view provides a very useful theoretical and numerical framework to conceive, analyze and implement strain based and mixed finite element discretizations of rod statics under constraints. The applications considered include the packing of an elastic ring and the formation of plectonemes.

**Motivation.** Slender elastic filament-like structures are ubiquitous in nature and in man made objects. Examples in nature, at our human scale, include tree branches, plant stems, hair and fur: the numerical simulation of the visually rich dynamics of these objects is of high interest in computer animation [1, 54]. At a smaller scale, the flagella of swimming micro-organisms [55] and even DNA molecules [56] are examples of slender structures studied in biomechanics. While examples in man made objects, of obvious interest in mechanical engineering, include industrial cables [57], marine cables [58], columns and helicopter blades [59] just to name a few. It is of the utmost importance in many applications to be able to predict the mechanical behaviour of such structures. This includes predicting their equilibrium configuration given some external loads, boundary conditions and constraints; predicting whether the equilibria are stable or unstable. Whether the structure will buckle or not; what dynamical evolution will the structure undergo, what internal forces will the structure experience, and so on.

**Rod models.** Slender elastic filament-like structures have in common that they are three dimensional elastic bodies where one dimension (the length) is much larger than the other two dimensions (the cross sections). They can be idealized and compactly described by mechanical models for rods (or beams). There is a wealth of rod models available, each relevant or valid under certain assumptions. Rod models are distinguished on the basis of at least two factors. The first is the degrees of freedom which the model allows, that is the kinematics. For example, does the model account for stretching, shearing, bending or twisting? And does it account for large displacements and finite rotations? The second, is on the underlying physics of the phenomenon it tries to capture, that is whether it captures, for instance, elasticity, plasticity, viscosity etc. The applications we consider involve the deformation of very thin rods under moderate loads and undergoing large deflections and rotations. Therefore we base our work in the setting of geometrically exact rod theory and make use of the Kirchhoff and Cosserat theories which we present in Chapter 5. Both theories account for small strains, large displacements

and finite rotations. The kinematics of a Kirchhoff rod allows only elastic bending and twisting while that of a Cosserat rod additionally allows for stretching and shearing.

**Numerical discretization of rod models.** To predict the mechanical behaviour of a rod-like structure, it suffices to solve the equilibrium equations of the rod. This is however no easy task. The equations have in general no closed form solution and analytical methods are of reduced applicability. One must then rely on a numerical solution of the equilibrium equations. The dynamic equilibrium equation of a Kirchhoff or of a Cosserat rod is a non linear hyperbolic partial differential equation (PDE) with one spatial and one temporal dimension. While some authors propose to discretize in both space and time simultaneously [60], such an approach is not convenient for our purposes, or more generally for purposes of flexible multibody dynamics. In the simulations of hair dynamics performed in our team [1, 61, 54] a single rod models a hair strand and is just part of a larger multibody system. To handle contact and friction interactions between rods, or to handle the coupling of rods with rigid bodies it is more convenient to discretize the rods spatially first. This way the whole set of rods and other bodies are described as a single Lagrangian mechanical system evolving in time, as that encountered in Part I of this thesis. The procedure of first discretizing in space and then in time is often referred to as a *method of lines*. For the computation of the static equilibrium of a rod only a spatial discretization needs to be carried out. In both the static and dynamic equilibrium cases the spatial discretization of the rod is key and it is precisely the issue which we address.

**Our contributions.** We propose methods for the computation of the stable static equilibria of Kirchhoff rods under constraints. The methods are fast and robust. They can handle boundary conditions, equality constraints and inequality constraints efficiently. They yield good qualitative results even at very coarse discretization levels, a property which is much cherished in computer graphics applications but also in fast prototyping design. Although we only deal with the problem of statics in this Part II of the manuscript, the spatial discretizations used in our methods can also be applied in a method of lines discretization of the dynamics. The applicability of the method for Cosserat rods is straightforward and is discussed. Our proposed methods, in the smooth case, can be seen as a family of integrators [62] previously unexplored for the statics of Kirchhoff rods. This fact alone motivates their development and investigation. The proposed methods rely on an alternative (Optimal Control) formulation of the Kirchhoff kinetic analogy and on a discretization of the kinematics using Runge-Kutta-Munthe-Kaas methods. We will see that our proposed methods fall into the category of *strain based* and *mixed* methods which generalize the construction of the Super-Helix element for Kirchhoff rods proposed in [1]. Our implementation methodology is very simple and we hope that this feature will lead to a wider use of strain based methods in applications.

## Chapter 5

# An Optimal Control Interpretation of Rod Statics

**Overview.** We present the Optimal Control formulation for the statics of Kirchhoff and Cosserat rods. To the best of our knowledge, the Optimal Control interpretation of the statics of Cosserat rods hasn't been exposed before. Although our simulations involve only Kirchhoff rods, we also present the Cosserat rod model as it is closely related and much of the literature on the numerics of Cosserat rods is relevant to our work. We also wish to highlight the applicability of our methods for Cosserat rods. In Section 5.1 we treat the case of planar rods, while in Section 5.2 we treat the case of spatial rods.

**On the Optimal Control interpretation.** We will see that it is possible to interpret the problem of rod statics as an Optimal Control Problem where the strains of the rod are interpreted as control variables and the position and orientation of the cross sections are interpreted as state variables. The Optimal Control formulation of the statics of rods is essentially a mixed variational formulation where the strains, the displacements and the orientation of the rod are taken as a priori independent variables coupled by kinematic constraints. A mixed variational formulation was already present in the work of Hodges [63] and in the work of Borri and Botasso [64] for Cosserat rods, although the latter uses stresses rather than the strains as independent variables. Recently Lang et al. [65] also used a mixed kinematic description for the dynamics of planar Kirchhoff rods. However the Optimal Control interpretation of Kirchhoff rod statics is not widespread. The only place where we have seen it explicitly stated is in the book of Jurdjevic 'Geometric Control Theory' [66], where it appears as an application of the theory of Optimal Control on Lie groups developed in the book.

In our opinion the Optimal Control formulation has the advantage of being the most explicit formulation, all variables and constraints are included. In this sense it constitutes an interesting alternative form of the Kirchhoff kinetic analogy. Kirchhoff figured out that the problem of finding static equilibria of inextensible and unshearable rods was remarkably similar to that of the dynamics of a heavy top. This analogy is very useful, as one can then draw conclusions on the theory and numerics of rods based on the enormous literature on the classical mechanics of the heavy top. However, as pointed out by Kehrbaum and Maddocks in [67], many authors are quick to overlook, or to be discrete about, the quirks in the Kirchhoff kinetic analogy. The most notable difference is that rod statics is a Boundary Value problem (BVP) instead of an Initial Value Problem (IVP). There is no counterpart to the position of the rod centerline in the dynamics of the top. Furthermore the role of gravity, a known parameter in the the top dynamics, is played by the internal force (an unknown) in the Kirchhoff rod.

Secondly, from a numerical and from an implementation point of view, the optimal control

interpretation is quite fruitful. It provides a very clear language to classify different discretization approaches, and suggests ways to combine Automatic differentiation (AD) tools, advanced numerical integrators for ODE's/DAE's and Non-Linear Program (NLP) solvers to assemble and solve the discretized OCPs. Nowadays, non-linear programming solvers are readily available, they get better by the day as the mathematical optimization community is very active; and some solvers have decades of heuristics and robustness built in to them. Environments like Casadi[68], intended for optimal control, bring all these different components together (AD, ODE/DAE solvers, NLP solvers), making it extremely easy to implement different formulations and discretizations of OCPs. An interesting feature of the direct methods, as we will see, is that even at a very coarse level of discretization the qualitative behavior of the solutions remains accurate. We will also see how direct methods of optimal control allow for a natural and efficient handling of frictionless inequality constraints.

**On the Lie group formalism.** Another feature of our formulation is that we employ a Lie group formalism. Specifically we take the configuration space of a rod to be the special Euclidean group  $SE(d)$ , where  $d = 2$  for planar rods and  $d = 3$  for spatial rods. We will see that a rod can be described by a point  $r(s) \in \mathbb{R}^d$  which tracks the position of the rod's centerline, and by an orthonormal frame  $R(s) \in SO(d)$  which tracks the orientation of the rod's cross sections. Our previous point of view was that the configuration space of a Kirchhoff rod was  $\mathbb{R}^d \times SO(d)$  and that the no stretch and no shear conditions were to be treated as separate constraints. After our reading of [69], the special Euclidean structure of the kinematics of Kirchhoff rods became apparent and we adopted it in our formulation. It may at first seem to the reader that using such a Lie group formalism is unnecessary, since after all the treatments of rods by Euler, Kirchhoff and the Cosserats did not require it. However we find it convenient to use the Lie group formalism since this language allows for a more unified treatment of planar and spatial Cosserat and Kirchhoff rods. Furthermore taking the special Euclidean Lie group structure into account leads to a natural formulation of strain measures and to the avoidance of 'locking problems' which are typical in classical finite element formulations [70]. Formulating rod kinematics in the special Euclidean groups is not yet standard practice in computational elasticity, although important exceptions include [69, 71, 72, 57]. In some papers it is used implicitly, more focus is given to the interpolation of rotations, interpolating positions being seen as a secondary problem. Locking problems then typically arise (in the Cosserat case) when positions are interpolated independently of the rotations or when strain measures aren't 'objective' (invariant with respect to rigid body motions). For example in [73] 'geodesic finite elements' for Cosserat rods are developed, however the authors use  $SO(3) \times \mathbb{R}^3$  as the configuration space leading to an incorrect interpolation of the position variables and to shear locking. The discretizations we propose avoid this problem if applied to Cosserat rods.

We include in Appendix C.1 background material on the Lie groups  $SO(3)$  and  $SE(3)$  for the convenience of the reader.

## 5.1 The continuum description of planar rods

### 5.1.1 Planar Kirchhoff rods

**Definition.** We call *rod* a three dimensional elastic body with two dimensions (the cross section) much smaller than the third (the length). A planar Kirchhoff rod, illustrated in Figure 5.1, has a centerline restricted to move in the plane and is characterized by small strains (linear elasticity), large displacements and finite rotations (geometrical non linearities). It is assumed to be perfectly inextensible, and undergoes only pure bending deformations in the plane.

**Some History.** The problem of finding the static equilibria of such an elastic structure was first formally formulated by James Bernoulli in 1691 who also published the first solution in 1694. The first full mathematical treatment using the calculus of variations was done by Euler who classified and illustrated the different solutions, which are to date referred to as Euler's elastic curves or Elastica. The reader interested in the beautiful history of the Elastica can consult for example [74].

**Kinematics.** The centerline of a planar Kirchhoff rod of length  $L$  is parametrized by a curve  $r$  mapping the arclength parameter  $s \in [0, L]$  to  $\mathbb{R}^2$ . The orientation of the cross sections is given by a material frame varying along the curve  $r$  and represented here as a rotation matrix  $R(s) \in SO(2)$ . The curvature of the centerline is given by scalar valued function which we denote  $\kappa$ .

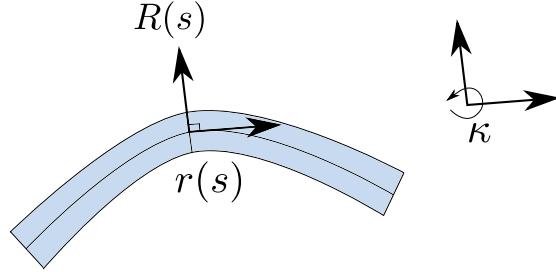


Figure 5.1: Centerline, material frame and curvature of a planar Kirchhoff rod.

The kinematics of the material frame reads  $R' = R\hat{\kappa}$ , where  $\hat{\kappa}$  is the skew symmetric matrix  $\hat{\kappa} = \begin{bmatrix} 0 & -\kappa \\ \kappa & 0 \end{bmatrix}$  and prime(') denotes derivation with respect to the arclength  $s$ . The Euler-Bernoulli constraint reads  $r' = Re_1$ , where  $e_1 = (1, 0)^T$ ; it encodes the incompressibility and no shear conditions, and couples the frame to the centerline. The frame is said to be adapted to the centerline. Because of this coupling we can't consider the frame  $R$  and the centerline  $r$  as independent variables. We thus take the configuration space of the planar Kirchhoff rod to be the Special Euclidean group  $SE(2)$  so that  $(R, r) : [0, L] \rightarrow SE(2)$ . The  $SE(2)$  structure of the problem becomes apparent by rewriting the frame kinematics and Euler-Bernoulli constraint as the  $SE(2)$  reconstruction equation,

$$\frac{d}{ds} \begin{bmatrix} R(s) & r(s) \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R(s) & r(s) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\kappa} & e_1 \\ 0 & 0 \end{bmatrix}. \quad (5.1)$$

We will denote the orthonormal column vectors of the matrix  $R$  as  $d_1$  and  $d_2$ , these are nothing but the tangent and normal vectors to the centerline. In terms of mechanics the *strains* of the beam can be defined as follows. The stretching strain is  $\epsilon_1 := r' \cdot d_1$ , it is unit because of the incompressibility condition :  $\epsilon_1 = d_1 \cdot d_1 = 1$ . The shearing strain is defined as  $\epsilon_2 := r' \cdot d_2$  and it is zero because of the no shear condition. The bending strain is just the curvature  $\kappa$  of the centerline.

**Energy of a planar Kirchhoff rod** The elastic bending energy density of a planar Kirchhoff rod is

$$U_b(s) = \frac{EI}{2}(\kappa(s) - \bar{\kappa}(s))^2,$$

where  $E$  is Young's modulus,  $I$  is the inertia of the cross section and  $\bar{\kappa}$  is the natural curvature of the rod. One can also take into account the potential energy due to gravity. Denoting by  $\rho$  the mass density of the rod,  $S$  the area of the cross section and  $g$  the gravitational constant, the expression of the gravitational potential energy density  $U_g$  of the rod is

$$U_g(s) = \rho S g r_2(s).$$

**Optimal control formulation.** Finding the stable static equilibria of a planar Kirchhoff rod subject to boundary conditions and frictionless contact constraints under gravitational forces can then be formulated as an Optimal Control Problem (OCP) where the states are  $(R, r)$  and the controls  $\kappa$ , which reads

$$\begin{aligned} \min_{R, r, \kappa} \quad & \int_0^L \mathcal{L}(R(s), r(s), \kappa(s)) ds \\ \text{s. t.} \quad & R' = R\hat{\kappa} \\ & r' = Re_1 \\ & g_I(R, r, \kappa) \leq 0 \\ & g_{bd}(R(0), r(0), R(L), r(L)) = 0, \end{aligned} \tag{5.2}$$

where the 'Lagrange cost' is the sum of bending and gravitational potential energy densities  $\mathcal{L} = U_b + U_g$ ,  $g_I$  models inequality constraints and  $g_{bd}$  encodes the boundary conditions. Boundary conditions may be for example fixed-fixed, fixed-free, periodic or even something more exotic. Note that the representation of  $SO(2)$  we're using is that of the set of 2 by 2 matrices  $R$  such that  $R^T R = I$  and  $\det(R) = 1$ . These two equality constraints are implicit in the frame reconstruction equation of (5.2). Provided that the initial frame  $R(0)$  is indeed in  $SO(2)$  and that  $R'(s) = R(s)\hat{\kappa}(s)$  then it holds that  $R(s)$  is in  $SO(2)$  for all  $s \in [0, L]$ .

### Alternative formulations

Several equivalent formulations of the statics of planar Kirchhoff rods can be derived, depending on the choice of the different possible representations of the configuration space  $SE(2)$  and on the handling of the Euler-Bernoulli constraint. The different formulations can be classified according to which subset of the configuration variables  $(R, r, \kappa)$  is used and on how  $R$  is represented. Formulations using  $(R, r)$  are called *absolute*. Those using only  $\kappa$  are called *intrinsic* or *strain based*. Those using all  $(R, r, \kappa)$  are called *mixed*. In [75] the author discretizes the strain based formulation by introducing a piecewise linear continuous strain field. The optimal control formulation proposed above is of the *mixed* kind and will lead in the discrete settings to both mixed and strain based numerical schemes in Chapter 6.

It is noteworthy that in [65] all three formulations are considered for the dynamic simulation of planar Kirchhoff rods with an extensional degree of freedom. Therein the rotations  $R$  are represented using unit complex numbers, indeed one has the isomorphism  $SO(2) \approx S^1 \approx \{z \in \mathbb{C} : \|z\| = 1\}$ . It is shown in [65] that strain based and mixed coordinates allow for larger time steps than absolute coordinates but that for a large number of elements the dense structure of the mass matrix in the strain formulation is prohibitive. In Chapter 6 a similar conclusion is reached for the statics of Kirchhoff rods, where the Hessian of the gravitational potential energy is dense for strain based discretizations.

**Angle formulation.** Nowadays the most common formulation of the Elastica [76, 77] is what we will call the 'angle formulation'. It was unknown to Euler and first discovered by Kirchhoff. It consists in representing  $SO(2)$  as the unit circle  $S^1$ , which can be parametrized by an angle  $\theta$ . The angle  $\theta(s)$  is the angle between the constant  $e_1$  direction and the tangent vector

$r'(s)$ . Note that this parametrization doesn't introduce singularities contrary to the situation for  $SO(3)$  and Euler angles. The material frame is then  $R(s) = \begin{bmatrix} \cos \theta(s) & -\sin \theta(s) \\ \sin \theta(s) & \cos \theta(s) \end{bmatrix}$ . The curvature is  $\kappa(s) = \theta'(s)$ . The expressions for the centerline are  $r_1(s) = \int_0^s \cos \theta(u) du$  and  $r_2(s) = \int_0^s \sin \theta(u) du$ . The problem of finding stable static equilibria of a planar Kirchhoff rod without gravity, without contact constraints and given boundary conditions  $r(0) = 0$ ,  $\theta(0) = \theta_0$  and  $r(L) = (x_L, y_L)$ ,  $\theta(L) = \theta_L$  then reads

$$\begin{aligned} \min_{\theta} \quad & \int_0^L \frac{EI}{2} \theta'(s)^2 ds \\ \text{s. t.} \quad & \int_0^L \cos \theta(s) ds = x_L \\ & \int_0^L \sin \theta(s) ds = y_L \\ & \theta(0) = \theta_0, \theta(L) = \theta_L. \end{aligned} \quad (5.3)$$

The problem in (5.3) is a problem of the Calculus of Variations with isoperimetric constraints [78], the mathematical theory of which was already available at the time of Kirchhoff whereas Optimal Control was not. The Euler-Lagrange equations of (5.3) are of the same form as those for the pendulum, this remark being known as the planar Kirchhoff kinetic analogy. We will use the angle formulation (5.3) in Chapter 6 mainly to obtain reference solutions.

**Position Formulation.** It is possible also to use only the position of the centerline  $r$  as an unknown and imposing  $\|r'(s)\|^2 = 1$ . The resulting problem is again a problem of the Calculus of Variations with isoperimetric like constraints. For a full account of this approach, including the quadrature of resulting the Euler-Lagrange equations in terms of Jacobi elliptic functions, the interested reader can consult [79]. The Euler-Lagrange equations for this formulation result in a 4th order non linear BVP for  $r$  which is hard to tackle numerically because of its stiffness.

### 5.1.2 Planar Cosserat rods

**Kinematics** Planar Cosserat rods [80, 81, 82] are like planar Kirchhoff rods which are allowed to stretch and shear. The planar Cosserat rod is still described as a framed curve in  $SE(2)$ , having a centerline  $r$  and a material frame  $R$ . However, the tangent vector  $r'$  needs not be unit (thus accounting for stretch), nor aligned with the director  $d_1$  (thus accounting for shear). The extra degrees of freedom are then the stretching strain  $v_1 := r' \cdot d_1$  and the shearing strain  $v_2 := r' \cdot d_2$ .

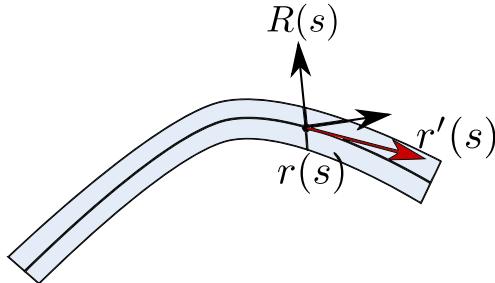


Figure 5.2: Centerline and material frame of a planar Cosserat rod. The tangent vector  $r'(s)$  isn't unit (accounting for stretch). The cross sections don't remain orthogonal to the centerline (accounting for shear).

Note however that, although the Euler-Bernoulli constraint has been relaxed, the frame and the centerline of a planar Cosserat rod are still coupled. The kinematics of the frame and the centerline follow the  $SE(2)$  reconstruction equation

$$\frac{d}{ds} \begin{bmatrix} R(s) & r(s) \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R(s) & r(s) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\kappa} & v \\ 0 & 0 \end{bmatrix}. \quad (5.4)$$

**Energy.** The planar Cosserat rod model has the same energy as the planar Kirchhoff rod, with the addition of an extra term in the elastic energy which accounts for extension and shearing  $U_{es}$ . Denoting  $S$  the area of the cross section,  $k$  the Timoshenko shear correction factor and  $\mu$  the shear modulus, the extension and shearing term in the elastic energy reads

$$U_{es} = \frac{1}{2}(v - e_1)^T C(v - e_1),$$

where  $C = \text{diag}(ES, \mu k S)$ .

**Optimal Control formulation.** The problem of finding stable static equilibria of planar Cosserat rods can then be formulated as the Optimal Control Problem

$$\begin{aligned} \min_{R, r, \kappa, v} \quad & \int_0^L \mathcal{L}(R(s), r(s), \kappa(s), v(s)) ds \\ \text{s. t.} \quad & R' = R\hat{\kappa} \\ & r' = Rv \\ & g_I(R, r, \kappa, v) \leq 0 \\ & g_{bd}(R(0), r(0), R(L), r(L)) = 0, \end{aligned} \quad (5.5)$$

where now  $\mathcal{L} = U_b + U_{es} + U_g$ , the states are  $(R, r)$  and the controls are  $(\kappa, v)$ .

## 5.2 The continuum description of spatial rods

### 5.2.1 Spatial Kirchhoff rods

**Kinematics.** The centerline of a spatial Kirchhoff rod [83, 84, 77, 85] of length  $L$  is parametrized by a curve  $r$  mapping the arclength parameter  $s \in [0, L]$  to  $\mathbb{R}^3$ . The orientation of the cross sections is given by a material frame varying along the curve  $r$  and represented here as a rotation matrix  $R(s) \in SO(3)$ . Much like in the planar case we will formulate the problem of finding the stable static equilibria of rods as an optimal control problem where the states are the frame and the centerline  $(R(s), r(s)) \in SE(3)$ . A spatial Kirchhoff rod, as depicted in Figure 5.3, can bend in two directions and can furthermore twist around its centerline. We collect the bending strains  $\kappa_1, \kappa_2$ , and the twisting strain  $\tau$  in a vector  $\kappa := (\kappa_1, \kappa_2, \tau)^T$  which will be interpreted as the control input in our OCP.

The kinematics of the material frame reads  $R' = R\hat{\kappa}$ , where  $\hat{\kappa}$  is the skew symmetric matrix  $\hat{\kappa} = \begin{bmatrix} 0 & -\tau & \kappa_2 \\ \tau & 0 & -\kappa_1 \\ -\kappa_2 & \kappa_1 & 0 \end{bmatrix}$ . The Euler-Bernoulli constraint in the spatial case reads  $r' = Re_3$ , where  $e_3 = (0, 0, 1)^T$ ; it encodes the incompressibility and no shear conditions, and couples the frame to the centerline (the frame is said to be adapted to the centerline). Again, the  $SE(3)$  structure of the problem becomes apparent by rewriting the frame kinematics and Euler-Bernoulli constraint as the  $SE(3)$  reconstruction equation,

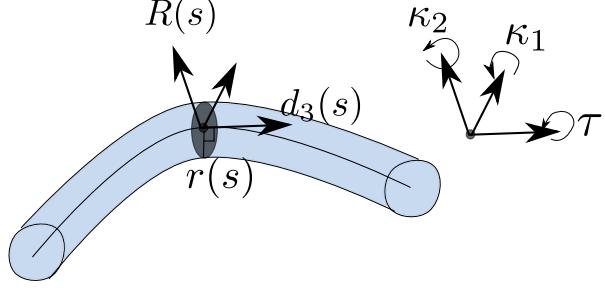


Figure 5.3: Centerline, material frame and curvatures of a Kirchhoff rod.

$$\frac{d}{ds} \begin{bmatrix} R(s) & r(s) \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R(s) & r(s) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\kappa} & e_3 \\ 0 & 0 \end{bmatrix}. \quad (5.6)$$

We will denote the orthonormal column vectors of the material frame  $R$  as  $(d_1, d_2, d_3)$ , the vector tangent to the centerline is  $d_3$ , the other two span the cross section.

**Energy of a spatial Kirchhoff rod** The elastic bending and twisting energy density of a spatial Kirchhoff rod is

$$U_{bt} = \frac{EI_1}{2}(\kappa_1 - \bar{\kappa}_1)^2 + \frac{EI_2}{2}(\kappa_2 - \bar{\kappa}_2)^2 + \frac{\mu J}{2}(\tau - \bar{\tau})^2,$$

where  $E$  is Young's modulus,  $\mu$  is the shear modulus,  $I_1$  and  $I_2$  are the cross section inertias and  $\bar{\kappa} := (\bar{\kappa}_1, \bar{\kappa}_2, \bar{\tau})^T$  are the natural curvatures and twist of the rod. One can also take into account the potential energy due to gravity. Denoting by  $\rho$  the mass density of the rod,  $S$  the area of the cross section and  $g$  the gravitational constant, the expression of the gravitational potential energy density  $U_g$  of the rod is

$$U_g = \rho S g r_3.$$

**Optimal control formulation.** Finding the stable static equilibria of a spatial Kirchhoff rod subject to boundary conditions and frictionless contact constraints under gravity can then be formulated as an Optimal control problem (OCP) where the states are  $(R, r)$  and the controls  $\kappa$ , it reads

$$\begin{aligned} \min_{R, r, \kappa} \quad & \int_0^L \mathcal{L}(R(s), r(s), \kappa(s)) ds \\ \text{s. t.} \quad & R' = R\hat{\kappa} \\ & r' = Re_3 \\ & g_I(R, r, \kappa) \leq 0 \\ & g_{bd}(R(0), r(0), R(L), r(L)) = 0, \end{aligned} \quad (5.7)$$

where the 'Lagrange cost' is the sum of bending, twisting and gravitational potential energy densities  $\mathcal{L} = U_{bt} + U_g$ ,  $g_I$  models inequality constraints and  $g_{bd}$  encodes the boundary conditions.

### Alternative formulations and their discrete counterparts.

**The Frenet, Material and Bishop frames.** A curve alone cannot represent by itself a physical rod as it would be unable to account for the twisting of the rod. Rods are therefore represented mathematically as *framed curves*, that is curves carrying a director frame. The

angular velocity of the frame along the centerline then accounts for bending and twisting strains. The traditional way of framing curves is to make use of the Frenet frame from the classical differential geometry of curves (see for instance Chapter 1 of [86]). However the Frenet frame requires the curve to have non vanishing second derivatives to be defined and can therefore not account for a rod with a perfectly straight centerline and a twisted configuration. Fortunately, there is more than one way to frame a curve [87]. The *Natural* (or *Bishop*) frame, denoted  $R_B = (U, V, T)$  is defined as an adapted frame to the curve  $r$ , that is  $r' = T$  which is twistless (its kinematics read  $R'_B = R_B \hat{u}$  with  $u = (u_1, u_2, 0)^T$ ). Introducing the angle between the Bishop and material frames leads to the curve angle representation [88].

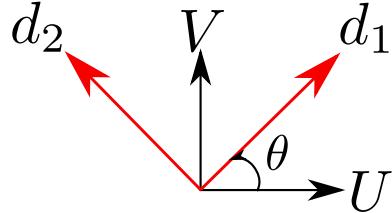


Figure 5.4: The Bishop and material frames coincide in the tangent vector  $d_3 = T = r'$  but differ by an angle  $\theta(s)$  which varies along the rod.

**Absolute.** In [88] the *curve-angle* representation is introduced to treat the Kirchhoff rod problem from a Lagrangian point of view. It consists in representing the rod using the centerline curve  $r(s)$ , the angle  $\theta(s)$  between the Bishop and the material frame as well as the initial director vector  $d_1(0)$  of the material frame. Twist is then measured as  $\tau(s) = \theta'(s)$ . In [89] a discrete version of the curve angle formulation for the dynamics of Kirchhoff rods is derived. A similar staggered grid approach is employed in [85] but with rotations represented as unit quaternions. However, both methods require imposing quadratic constraints on the position variables to enforce the inextensibility of the Kirchhoff rod.

In [90] the statics of Kirchhoff rods is discretized leveraging the Kirchhoff kinetic analogy and the principles of discrete time Lagrangian mechanics on Lie groups. Continuing on the same path as [90], Moser-Veselov integrators are developed in [91] for the statics of Kirchhoff rods. These method however don't account for the gravitational potential energy of the rod as then the classical Kirchhoff kinetic analogy breaks down.

**Strain based.** In [1, 54] the strain based formulation is discretized using respectively piecewise constant and piecewise linear shape functions leading to helical and clothoidal elements for Kirchhoff rods. These methods preserve the inextensibility and no shear constraints intrinsically. Our proposed discretizations in Chapter 6 can be seen as a generalization of these strain based methods. In the Cosserat case strain based discretizations are considered in [59, 92, 93, 94] although the use of

### 5.2.2 Spatial Cosserat rods

**Kinematics** Spatial Cosserat rods [95, 81, 82, 85] are like spatial Kirchhoff rods which are allowed to stretch and shear. A Cosserat rod is described as a framed curve in  $SE(3)$ , having a centerline  $r$  and a material frame  $R$ . However, the tangent vector  $r'$  needs not be unit (thus accounting for stretch), nor aligned with the director  $d_1$  (thus accounting for shear). The extra

degrees of freedom are then the stretching strain  $v_3 := r' \cdot d_3$  and the two shearing strains  $v_1 := r' \cdot d_1$  and  $v_2 := r' \cdot d_2$ .

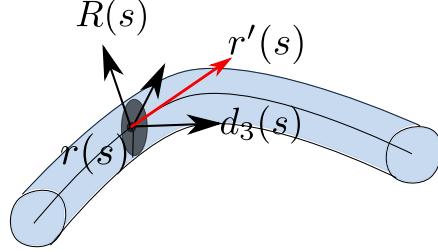


Figure 5.5: Centerline and material frame of a spatial Cosserat rod. The tangent vector  $r'(s)$  isn't unit (accounting for stretch). The cross sections don't remain orthogonal to the centerline (accounting for shear).

Note however that, although the Euler-Bernoulli constraint has been relaxed, the frame and the centerline of a spatial Cosserat rod are still coupled. The kinematics of the frame and the centerline follow the  $SE(3)$  reconstruction equation

$$\frac{d}{ds} \begin{bmatrix} R(s) & r(s) \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R(s) & r(s) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\kappa} & v \\ 0 & 0 \end{bmatrix}. \quad (5.8)$$

**Energy.** The spatial Cosserat rod model has the same energy as the spatial Kirchhoff rod, with the addition of an extra term in the elastic energy which accounts for extension and shearing  $U_{es}$ . Denoting  $S$  the area of the cross section,  $k_1, k_2$  the Timoshenko shear correction factors and  $\mu$  the shear modulus, the extension and shearing term in the elastic energy reads

$$U_{es} = \frac{1}{2}(v - e_3)^T C(v - e_3),$$

where  $C = \text{diag}(\mu k_1 S, \mu k_2 S, E S)$ .

**Optimal Control formulation.** The problem of finding stable static equilibria of spatial Cosserat rods can then be formulated as for Kirchhoff rods:

$$\begin{aligned} \min_{R, r, \kappa, v} \quad & \int_0^L \mathcal{L}(R(s), r(s), \kappa(s), v(s)) ds \\ \text{s. t.} \quad & R' = R\hat{\kappa} \\ & r' = Rv \\ & g_I(R, r, \kappa, v) \leq 0 \\ & g_{bd}(R(0), r(0), R(L), r(L)) = 0, \end{aligned} \quad (5.9)$$

where now  $\mathcal{L} = U_b + U_{es} + U_g$  and where the controls are  $\kappa$  and  $v$ .

**Objectivity of Strain measures.** As alluded to in the introduction of this chapter, in the  $SE(d)$  formalism the definition of strain measures is very natural and enjoys the very important property that they are invariant with respect to superimposed euclidean transformations. The importance of this property for rods was identified in [96] and coined 'Objectivity of strain measures'. Obviously the stored internal elastic energy of the rod should not depend on where the rod is in space or on how it is oriented, but only on the relative position and orientation of

the cross sections. However when designing finite elements for beams in the classical settings, strains need to be expressed in terms of the nodal positions and orientations and it is easy for the resulting expression to violate objectivity. A configuration of a rod, in both the Cosserat and Kirchhoff models, is given by a frame and a centerline  $(R(s), r(s)) \in SE(d)$ . The strains associated to that configuration are the  $\mathfrak{se}(d)$  elements  $(\kappa(s), v(s))$  such that

$$\frac{d}{ds} \begin{bmatrix} R(s) & r(s) \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R(s) & r(s) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\kappa} & v \\ 0 & 0 \end{bmatrix}.$$

If we superimpose an euclidean transformation  $(A, b) \in SE(d)$  to the configuration  $(R(s), r(s))$ , the resulting configuration of the rod is given by  $(AR(s), Ar(s) + b)$ . The strains of the resulting configuration are seen to remain unchanged as the following simple computation shows:

$$\frac{d}{ds} \begin{bmatrix} AR & Ar + b \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} AR\hat{\kappa} & ARv \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} AR & Ar + b \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\kappa} & v \\ 0 & 0 \end{bmatrix}.$$

### 5.3 The Kirchhoff Kinetic analogy

Now that we have seen how the problems of rod statics can be formulated as optimal control problems we take a more detailed look at the classical Kirchhoff kinetic analogy and at what its optimal control formulation suggests.

**The classical Kirchhoff kinetic analogy.** As pointed out by Kehrbaum and Maddocks in [67], many authors are quick to overlook or be discrete about the quirks in the Kirchhoff kinetic analogy. We therefore present in the following table a summary of the similarities and differences between Kirchhoff rod statics and heavy top dynamics. The most notable difference is that rod statics is a BVP instead of an IVP, there is no counterpart to the position of the rod centerline in the dynamics of the top and finally the role of gravity (a known parameter) in the the top dynamics is played by the internal force (an unknown) in the Kirchhoff rod.

Kirchhoff Rod statics	Heavy Top dynamics
arclength $s$ Material frame $(d_1, d_2, d_3)$ Bending and twisting strains $\kappa := (\kappa_1, \kappa_2, \tau)$ Bending and twisting stiffnesses $EI_1, EI_2, \mu J$ Elastic Energy $\frac{1}{2}(EI_1\kappa_1^2 + EI_2\kappa_2^2 + \mu J\tau^2)$ Bending and twisting Moments $m = K\kappa$	time $t$ Frame aligned with axes of inertia $(d_1, d_2, d_3)$ Body angular velocities $\omega := (\omega_1, \omega_2, \omega_3)$ Principal moments of inertia $I_1, I_2, I_3$ Kinetic energy $\frac{1}{2}(I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2)$ Angular momenta $\Pi = I\omega$
Internal Force $F$ (unknwon) Total Energy density $\frac{1}{2}\kappa^T K\kappa + F^T d_3$ Centerline $r(s)$ Constraint $r' = d_3$ Initial Bending and twisting $\kappa_{s=0}$ (unknown) Boundary conditions at $s = L$	Gravity Force $F_g := -mge_3$ (known) Lagrangian $\frac{1}{2}\omega^T I\omega + lF_g^T d_3$ None None Initial Angular velocity $\omega_{t=0}$ (known) None

Where  $l$  is the distance from the fixed point of the top to its center of mass.

**A Kirchhoff Kinetic analogy from the control viewpoint.** The Optimal Control formulation of Kirchhoff rod statics provides an alternative, and more explicit, form of the Kirchhoff kinetic analogy. What it suggests is that the statics of a Kirchhoff rod is equivalent to the non-holonomic mechanics of a rigid body in  $d$ -dimensional space, where  $d = 2$  for planar rods

and  $d = 3$  for spatial rods. The configuration space of the analogous rigid body is  $SE(d)$ , its angular velocity is analogous the the bending strains  $\kappa$  of the rod. The analogous rigid body can assume any angular velocity, reflecting the fact that for a Kirchhoff rod the bending (and twisting) strains are unconstrained. However it is subject to the non-holonomic constraint that the velocity  $v$  of its center of mass expressed in the body frame is always constant,  $v = e_3$  in the spatial case and  $v = e_1$  in the planar case. Thus reflecting the fact that for a Kirchhoff rod the stretching and shearing strains are imposed. The analogous rigid body is thus like a rocket whose back thrusters constrain it to move at constant unit speed in the direction of one of its axes of inertia, while it is impossible for it to instantaneously move orthogonally to that axis.

## 5.4 Conclusion

We have seen that the problem of computing the stable static equilibria of a Kirchhoff or of a Cosserat rod can be formulated as an Optimal Control problem on  $SE(d)$ , where  $d = 2$  for planar rods and  $d = 3$  for spatial rods. We have seen that the optimal control formulation is simply a mixed variational formulation where the strains are interpreted as 'controls' and the position and orientation as 'states'. We have discussed its connection with other more standard formulations of rod statics and exposed the alternative form of the Kirchhoff kinetic analogy which it suggests. In the next chapter we propose to discretize our Optimal Control formulation by using direct methods of numerical optimal control. We will see how discretizing the strain field and using Runge-Kutta-Munthe-Kaas methods to discretize the kinematics leads to the formulation of strain based and mixed finite elements for rod statics.

# Chapter 6

## A Hamilton-Pontryagin Approach for the statics of Kirchhoff rods

**Overview.** We saw in Chapter 5 that it is possible to formulate the problem of finding stable static equilibria of Kirchhoff rods as an optimal control problem which, we recall, reads

$$\begin{aligned} \min_{R, r, \kappa} \quad & \int_0^L \mathcal{L}(R(s), r(s), \kappa(s)) ds \\ \text{s. t.} \quad & R' = R\hat{\kappa} \\ & r' = Re_3 \\ & g_I(R, r, \kappa) \leq 0 \\ & g_{bd}(R(0), r(0), R(L), r(L)) = 0, \end{aligned} \tag{6.1}$$

where the position and frame fields  $(R, r) : [0, L] \rightarrow SE(3)$  are interpreted as 'states' and the bending and twisting strain fields  $\kappa : [0, L] \rightarrow \mathbb{R}^3$  are interpreted as 'controls'.

From the point of view of numerical optimal control there are two major categories of ways to find local solutions of the problem (6.1), namely *indirect methods* and *direct methods*. Although there are also methods to find global minima of optimal control problems, in this thesis we are only concerned with local methods. Global methods are typically prohibitively expensive for problems having more than a few states due to the 'curse of dimensionality'. Furthermore we're more interested in applications like finding local minima and tracking them as some parameters of the system evolve, hence there is no real need for global methods. That being said, we will see that the direct methods we propose do tend to find local optima of rather low energy levels, in contrast to the indirect methods. For the reader unfamiliar with the distinction between direct and indirect methods, we now summarize it shortly before delving into their precise application to problem (6.1).

**The indirect approach** can be summarized as '*first optimize then discretize*'. First one writes the (infinite dimensional) first order optimality conditions of the OCP (6.1). The result is a boundary value problem (BVP). Then using a suitable numerical integrator one discretizes the BVP, which results in a (finite dimensional) set of non linear algebraic equations. One then solves these algebraic equations using a Newton type method. Actually when inequality constraints are present the result is a set of non linear generalized equations (or inclusions) and one would then have to employ a non smooth Newton type method.

**The direct approach** can be summarized as '*first discretize then optimize*'. First one discretizes the states and controls, a suitable numerical integration scheme is employed to discretize

the constraints of problem (6.1) which then becomes a (finite dimensional) non linear programming problem (NLP). One then solves (optimizes) the NLP using standard NLP software like Ipopt [97], which implements an interior point method<sup>1</sup>.

**When the two commute.** Instead of using a minimization routine to solve the NLP of the direct approach, we can write the first order optimality conditions of the NLP, resulting in a set of non linear algebraic equations (or inclusions). The question is then whether these equations are the same as those obtained via the indirect approach. The answer is that it is not always the case. We refer the reader interested in these questions to the very good review article of Sanz-Serna [98], where in particular it is explained that the answer is positive in the following very common scenario. In the absence of inequality constraints the BVP of the indirect approach is governed by a smooth Hamiltonian dynamical system. If the ODE constraints in the direct approach are discretized using a Runge-Kutta scheme, then the resulting optimality conditions describe a symplectic partitioned Runge-Kutta method to solve Hamiltonian BVP of the indirect approach.

**Remark.** In our opinion the issue of whether or not to use direct or indirect methods is a moot point. One rather needs to understand the advantages and disadvantages of each method and use them accordingly, or even use both if the problem requires it. For example a typical disadvantage of indirect methods is their sensitivity to the initial guess. One can in such cases use a the solution of a direct method with a coarse discretization to initialize the indirect method. Also, it is easier to obtain local minima (so *stable* static equilibria) with direct methods since there is often a line search procedure in the NLP solver which guarantees that the local solution found is not a maximum or a saddle point, whereas with indirect methods one is not guaranteed to find a local minimum but rather any kind of inflexion point of the objective. Another reason why we are interested in developing direct methods is that they yield finite dimensional expressions of the energy and kinematics which can then be reused in the context of dynamics. A very common procedure to simulate the dynamics of beams is to use the so called 'method of lines', where one first discretizes in space and let the spatial degrees of freedom depend on time. The result is a reduction of beam dynamics to a problem of multibody dynamics. But we defer further discussion of dynamics to a future chapter. For now let us leave it a motivational point to study the spatial discretizations of energy and kinematics suggested by the direct approach.

We will see in the remainder of this chapter how the Optimal Control point of view provides a very useful theoretical and numerical framework to conceive, analyse and implement strain based and mixed finite element discretizations of rod statics under constraints.

## 6.1 Indirect Methods

This section deals with indirect methods for planar and spatial Kirchhoff rods in the absence of inequality constraints. We show how to recover the classical Kirchhoff balance equations from the optimal control formulation and show how to solve them numerically using shooting techniques. Some of the numerical solutions obtained will be later used as a reference when evaluating the direct methods.

---

<sup>1</sup>Sequential quadratic programming (SQP) methods can also be used although freely available implementations are scarce.

### 6.1.1 Derivation of the Kirchhoff Equations for rod statics

We consider a Kirchhoff rod in the absence of external forces and contact constraints. In the following we show how to recover the Kirchhoff equations of rod statics from the optimal control formulation in left reduced form (meaning in the material frame) and right reduced form (meaning in the spatial frame). We will then point out how the equations change when the rod's weight and natural curvature are taken into account, as well as the difference and similarities with the analogous derivation for Cosserat rods.

**Proposition 6.1.1.** *Let  $H$  be the  $SE(3)$  frame collecting the material frame  $R$  and the centreline position  $r$  as  $H = \begin{bmatrix} R & r \\ 0 & 1 \end{bmatrix}$ . Let  $\xi$  be the  $\mathfrak{se}(3)$  element collecting the bending-twisting strains  $\kappa$  and the stretching-shearing strains  $v$  as  $\xi = \begin{bmatrix} \hat{\kappa} & v \\ 0 & 0 \end{bmatrix}$ . Then the following points hold.*

- (i) **Constrained  $SE(3)$  formulation.** *The problem of finding stable static equilibria of spatial Kirchhoff rods under fixed-fixed boundary conditions and in the absence of contact constraints and external forces can be reformulated as*

$$\begin{aligned} \min_{H, \kappa, v} \quad & \int_0^L l(\xi(s)) ds \\ \text{subject to} \quad & H(s)^{-1} H'(s) = \xi(s) \\ & v(s) = e_3 \\ & H(0) = H_0, H(L) = H_L, \end{aligned} \tag{6.2}$$

where the reduced Lagrangian  $l$  is defined as  $l(\xi) := \frac{1}{2} \kappa^T K \kappa$  and the boundary conditions as  $H_0 = (R_0, r_0)$  and  $H_L = (R_L, r_L)$ .

- (ii) **Left reduced Pontryagin optimality conditions.** *The necessary conditions for optimality of (6.2) in left trivialized and left reduced form read: there exists a frame field  $H : [0, L] \rightarrow SE(3)$ , a strain field  $\xi = (\kappa, v) : [0, L] \rightarrow \mathfrak{se}(3)$ , a stress field  $\mu = (m, f) : [0, L] \rightarrow \mathfrak{se}(3)^*$ <sup>2</sup> and a multiplier field  $\lambda : [0, L] \rightarrow \mathbb{R}^3$  such that the following equations hold*

$$\left\{ \begin{array}{ll} v &= e_3 \\ r' &= Rv \\ R' &= R\hat{\kappa} \\ m' &= -\kappa \times m - v \times f \\ f' &= -\kappa \times f \\ m &= K\kappa \\ f &= \lambda. \end{array} \right. \begin{array}{l} \text{Euler -- Bernoulli} \\ \text{frame reconstruction} \\ \text{momentum balance} \\ \text{force balance} \\ \text{Legendre transform} \end{array} \tag{6.3}$$

- (iii) **The Kirchhoff equations follow from left reduced optimality conditions (6.3).** *The momentum  $M := Rm$  and the force  $F := Rf$  expressed in the fixed reference frame, that is in spatial coordinates, verify the Kirchhoff equations*

$$\begin{aligned} F' &= 0 \\ M' &= F \times d_3. \end{aligned} \tag{6.4}$$

*Proof.* Item (i) just expresses the constraints of the problem (6.1) in the absence of external forces and contact constraints in a different manner. The Euler-Bernoulli constraint combined

---

<sup>2</sup>While the strains  $\xi = (\kappa, v)$  are elements of  $\mathfrak{se}(3)$ , the stresses  $\mu = (m, f)$  are elements of the dual  $\mathfrak{se}(3)^*$ . The element  $m$  represents the moment and  $f$  the force on the rod, both expressed in the frame  $H$ .

with the frame reconstruction equation on  $SO(3)$  can be seen as a reconstruction equation on  $SE(3)$  with bending-twisting strains  $\hat{\kappa}$  and imposed stretching-shearing strains  $v = e_3$ . Leading to the constraint reformulation as  $\frac{d}{ds} \begin{bmatrix} R & r \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R & r \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\kappa} & v \\ 0 & 0 \end{bmatrix}$  and  $v = e_3$ . The reason to express the reconstruction constraint as  $H^{-1}H' = \xi$  instead of  $H' = H\xi$  is that the former equality takes place in the Lie algebra  $\mathfrak{se}(3)$ , instead of in the tangent space to  $SE(3)$  at  $H$  ( $T_H SE(3)$ ), and thus requires only 6 multipliers to be enforced instead of 12.

Item (ii) is then a minor adaptation of the derivation of Euler-Poincaré and Lie-Poisson equations of  $SE(3)$  [99]. If  $(H, \kappa, v)^*$  is an optimal point of problem (6.2) then there exist multipliers  $(m, f, \lambda)^*$  such that the functional

$$J(H, \kappa, v, m, f, \lambda) := \int_0^L l(\kappa) + \langle (m, f), H^{-1}H' - \xi \rangle + \langle \lambda, v - e_3 \rangle ds$$

is stationary at  $(H, \kappa, v, m, f, \lambda)^*$ . We now compute the first variation of this functional. All quantities are varied independently and variations of  $H$  are such that  $\delta H(0) = \delta H(L) = 0$  (because of the boundary conditions). Variations with respect to the multipliers  $\lambda$  yield back the constraint  $v = e_3$ . Variations with respect to  $\mu$  yield back the constraint  $H^{-1}H' = \xi$ . Variations with respect to  $H$  yield the momentum and force balance equations, we shall take a closer look at them. Variations  $\delta H$  are such that there exists an  $\mathfrak{se}(3)$  field  $\eta$  which vanishes at the endpoints and such that  $\delta H = H\eta$ . The following computation uses the property that  $\delta(H^{-1}) = -H^{-1}\delta HH^{-1}$ , it uses the definitions of the left reduced variations  $\delta H = H\eta$  and of the kinematics  $H' = H\xi$ . It also uses the definition of the adjoint action  $\text{ad}_\xi\eta = [\xi, \eta] = \xi\eta - \eta\xi$  and of its dual (see Appendix C.1).

$$\begin{aligned} \left\langle \frac{\delta J}{\delta H}, \delta H \right\rangle &= \int_0^L \langle \mu, -H^{-1}\delta HH^{-1}H' + H^{-1}(\delta H)' \rangle ds \\ &= \int_0^L \langle \mu, -\eta\xi + H^{-1}(H'\eta + H\eta') \rangle ds \\ &= \int_0^L \langle \mu, -\eta\xi + H^{-1}(H\xi\eta + H\eta') \rangle ds \\ &= \int_0^L \langle \mu, -\eta\xi + \xi\eta + \eta' \rangle ds \\ &= \int_0^L \langle \mu, \text{ad}_\xi\eta + \eta' \rangle ds \\ &= \int_0^L \langle \text{ad}_\xi^*\mu, \eta \rangle + \langle \mu, \eta' \rangle ds \\ &= \int_0^L \langle -\mu' + \text{ad}_\xi^*\mu, \eta \rangle ds + [\langle \mu, \eta \rangle]_0^L. \end{aligned}$$

Thus, for the functional  $J$  to be stationary it is necessary that the following ODE holds:  $\mu' = \text{ad}_\xi^*\mu$ . Expressing the latter in terms of the moments  $m$  and the forces  $f$  and using the definition of the coadjoint action we can establish the following balance equations

$$\begin{aligned} m' &= -\kappa \times m - v \times f \\ f' &= -\kappa \times f. \end{aligned}$$

It remains to take variations of the functional  $J$  with respect to the strain field  $\xi$ . Doing so and

decomposing again in angular ( $\delta\kappa$ ) and linear ( $\delta v$ ) parts we find that

$$\begin{aligned}\left\langle \frac{\delta J}{\delta \xi}, \delta \xi \right\rangle &= \int_0^L \left\langle \frac{\partial l}{\partial \xi}, \delta \xi \right\rangle - \langle \mu, \delta \xi \rangle + \langle (0, \lambda), \delta \xi \rangle ds \\ &= \int_0^L \left\langle \frac{\partial l}{\partial \kappa} - m, \delta \kappa \right\rangle + \left\langle \frac{\partial l}{\partial v} - f + \lambda, \delta v \right\rangle ds.\end{aligned}$$

The latter expression must vanish for all variations  $\delta\kappa$  and  $\delta v$ . Thus we find that  $f = \lambda$  (since  $l$  is independent of  $v$ )<sup>3</sup>, so that the force  $\lambda$  maintaining the imposed longitudinal strains  $v = e_3$  is the same as the multiplier  $f$  maintaining the centreline reconstruction constraint. We also find that

$$m = \frac{\partial l}{\partial \kappa} = K\kappa,$$

which in terms of mechanics is the Legendre transform, allowing to go from 'angular velocities'  $\kappa$  to 'angular momenta'  $m$ . In terms of Optimal Control this last equation allows to express the 'controls'  $\kappa$  in terms of the 'adjoint states'  $m$ . In terms of elasticity these are constitutive equations for the elastic material.

Item (iii) of the proof then easily follows from item (ii). Item (ii) expresses the optimality conditions in the material frame. To recover the optimality conditions in a fixed reference frame one establishes an ODE for the spatial momentum  $M = Rm$  and for the spatial force  $F = Rf$ . This is easily done using the equilibrium equations of item (ii)

$$M' = R'm + Rm' = R(\kappa \times m) + R(-\kappa \times m - e_3 \times f) = -Re_3 \times Rf = -d_3 \times F,$$

so that the momentum balance reads  $M' + d_3 \times F = 0$ . It remains to establish the spatial force balance, we do it similarly as

$$F' = R'f + Rf' = R(\kappa \times f) + R(-\kappa \times f) = 0.$$

We thus recover the well known fact that the internal force on a Kirchhoff rod, expressed in the reference frame, is constant.  $\square$

**Remark.** In the case where the Lagrangian accounts for the elastic energy of a pre-bent and pre-twisted Kirchhoff rod under the influence of gravity, the reduced Lagrangian still depends on  $H$  and can be defined as  $l(H, \xi) = \mathcal{L}(H, H\xi) = \frac{1}{2}(\kappa - \bar{\kappa})^T K(\kappa - \bar{\kappa}) + \rho S g e_3^T r$ . In that case there is an extra term in the variation with respect to  $H$  so that

$$\left\langle \frac{\delta J}{\delta H}, \delta H \right\rangle = \int_0^L \left\langle -\mu' + \text{ad}_\xi^* \mu + H^T \frac{\partial l}{\partial H}, \eta \right\rangle ds + [\langle \mu, \eta \rangle]_0^L.$$

This leads to a force balance equation which reads  $f' = -\kappa \times f + \rho S g R^T e_3$  in the body frame and  $F' = \rho S g e_3$  in the spatial frame as expected [77]. Also, the constitutive equation becomes  $m = K(\kappa - \bar{\kappa})$  instead of  $m = K\kappa$ .

**On boundary conditions.** Note that in the case of fixed-fixed boundary conditions one knows the initial state  $H(0)$  but the initial stress resultants  $\mu(0)$  ('costates' in optimal control terminology) are unknown. Hence we are dealing indeed with a boundary value problem and not with an initial value problem. One is given the final state  $H(L)$  and the objective is to

---

<sup>3</sup>In the case of Cosserat rods  $l$  depends quadratically on  $v$  and one recovers a constitutive law for the stretching-shearing strains and stresses  $f = \frac{\partial l}{\partial v}$ .

find the initial costate  $\mu(0)$  such that after forward integration of the Kirchhoff equations (6.3) the target final state is reached. The problem of fixed-free boundary conditions is similar. One can conclude from the boundary term  $[\langle \mu, \eta \rangle]_0^L$  in the derivation of Kirchhoff's equations that if  $\eta(L)$  is free then  $\mu(L)$  must vanish. We are then left with a two point boundary value problem where the initial state is known as well as the final costate  $\mu(L) = 0$ .

### 6.1.2 Kirchhoff Indirect single shooting experiments

The indirect single shooting approach to solve the OCP (6.2) consists in first obtaining the necessary conditions for optimality in (6.3), which is a boundary value problem (BVP), and then discretizing that BVP numerically. First we shall write the BVP in dimensionless form, then we explain how to discretize it numerically.

**Dimensionless Kirchhoff equations.** One can rewrite the Kirchhoff equations (6.3) in the material frame as

$$\begin{cases} r' &= d_3 \\ R' &= R\hat{\kappa} \\ \kappa' &= K\kappa \times \kappa + f \times e_3 \\ f' &= f \times \kappa. \end{cases}$$

Let us consider a rod with a symmetric cross section, then the cross section inertias are equal:  $I := I_1 = I_2$ . Introducing the dimensionless quantities  $\bar{s} := s/L$ ,  $\bar{\kappa}_i(\bar{s}) := L\kappa_i(\bar{s}L)$ ,  $\bar{r}(\bar{s}) := r(\bar{s}L)/L$ ,  $\bar{R}(\bar{s}) := R(\bar{s}L)$ ,  $\gamma := \frac{\mu J}{EI}$  and  $\bar{f}(\bar{s}) := \frac{L^2}{EI}f(\bar{s}L)$  one can write the Kirchhoff equations in dimensionless form. Omitting the bars to avoid cluttering, they read

$$\begin{cases} r' &= d_3 \\ R' &= R\hat{\kappa} \\ \kappa'_1 &= (1 - \gamma)\kappa_2\tau + f_2 \\ \kappa'_2 &= (\gamma - 1)\kappa_1\tau - f_1 \\ \tau' &= 0 \\ f' &= f \times \kappa, \end{cases} \quad (6.5)$$

together with the boundary conditions. The well-known fact that an equilibrium position of a symmetric rod has a constant material twist  $\tau$  is seen immediately in equations (6.5).

**Indirect Single shooting discretization.** The initial state  $(R(0), r(0)) = (R_0, r_0)$  is fixed, as well as the final state  $(R(1), r(1)) = (R_1, r_1)$ . In indirect single shooting the unknown is the initial co-state  $\mu_0 = (\kappa_0, f_0)$ . One wishes to find the initial co-state which yields the correct final state. Solve numerically the ODE (6.5) with initial conditions  $(R_0, r_0)$  and  $\mu_0$  up to the final arclength  $s = 1$  and denote the output final state  $(R_{\mu_0}, r_{\mu_0})$ . The non linear system of equations to solve for  $\mu_0$  is then

$$r_{\mu_0} = r_1, \quad R_{\mu_0} = R_1.$$

These are however 12 equations for 6 unknowns, the problem is over-constrained. To avoid this difficulty it is necessary to impose the equality of the rotation matrices as 3 equations rather than 9, after all  $SO(3)$  is a three dimensional manifold and this should be possible. The apparent conundrum is resolved by noticing that if the initial frame  $R_0$  is in  $SO(3)$  then solving the forward kinematics with sufficient accuracy guarantees that the final frame  $R_{\mu_0}$  is also in  $SO(3)$ . Hence, intuitively, out of the 9 degrees of freedom in the 3 by 3 rotation matrix  $R_{\mu_0}$ , 6 are blocked since  $R_{\mu_0}$  is orthonormal and satisfies  $R_{\mu_0}^T R_{\mu_0} = I_{3 \times 3}$ . Only 3 degrees of freedom of  $R_{\mu_0}$  remain to be imposed. It is sufficient to impose that  $\text{diag}(R_1^T R_{\mu_0} - I_{3 \times 3}) = 0$  so that column  $i$  of  $R_{\mu_0}$  is aligned with column  $i$  of  $R_1$  and of the same norm.

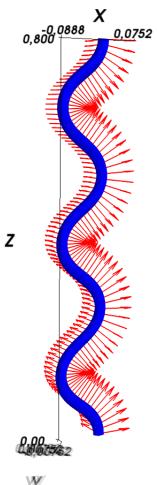
**Comments:** The problem has six equations and six unknowns, it is a very small problem, however evaluating  $(R_{\mu_0}, r_{\mu_0})$  is expensive and the result tends to be very sensitive to the initial conditions  $\mu_0$ . Also note that Newton's method will require the derivatives of  $(R_{\mu_0}, r_{\mu_0})$  with respect to  $\mu_0$ , in our implementation we use Casadi [68] to obtain these derivatives via automatic differentiation.

**Numerical Experiment:** We solved the single shooting problem with boundary conditions  $r_0 = 0, R_0 = I_{3 \times 3}$  and  $R_1 = I_{3 \times 3}, r_1 = (0, 0, 0.8)$ . As a numerical integrator we used a Runge-Kutta order 4 scheme with a fixed step size of  $10^{-3}$ . The non linear algebraic equations were solved using Ipopt, which in this case performs a version of the stabilized Newton-Raphson algorithm. A sample of 100 initial guesses of the form  $\mu_0^0 = (0, \kappa_2^0, 0, f_1^0, 0, f_3^0)$  was generated, where  $(\kappa_2^0, f_1^0, f_3^0)$  were taken randomly in  $[-2, 2]$ . For 72% of the problems attempted the solver converged. The solutions obtained are planar stable or unstable equilibria, some of them are depicted in Figure 6.1. We performed another test, where we generated 100 samples for the whole initial costate  $\mu_0$  in  $[-2, 2]$ , only 12 percent of the problems were solved. Clearly the indirect single shooting problem is quite difficult to solve numerically, even with a very robust non linear solver like Ipopt. One classical remedy is to use indirect multiple shooting where several single shooting problems are solved on subintervals of  $[0, 1]$ , thus trading off a larger problem size for a better structure (reduced non linearities). Other remedies include using a better initial guess for the shooting algorithm, which could come from using the result of a direct method, or from a better knowledge of the solutions. For example the choice of initial guesses of the form  $\mu_0^0 = (0, \kappa_2^0, 0, f_1^0, 0, f_3^0)$  exploits the fact that we know that there are planar solutions and that the strains of a planar configuration must have that form.

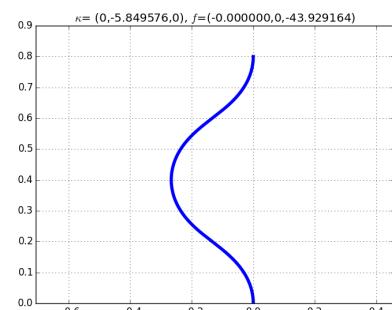
### 6.1.3 Indirect methods for Euler's Elastica with fixed-fixed boundary conditions.

We consider in the following the scenario of a planar Kirchhoff rod with both ends clamped, without gravity and without any contact constraints. We first reproduce the construction of analytical solutions in the small compression regime and compare them to the numerical solutions obtained by indirect single shooting. These solutions will be also used later to evaluate the accuracy of the proposed direct methods.

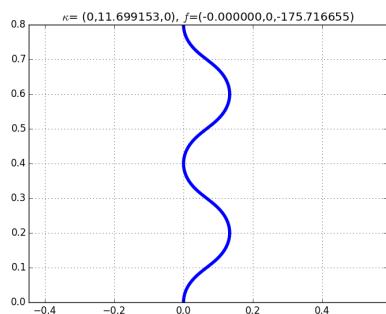
**Buckling of a fixed-fixed Euler elastica, small compression regime.** Here we follow the construction by Audoly and Pomeau [77, Section 8.5.3]. The centerline  $r(s)$  of the planar elastica has two components  $r(s) = (x(s), y(s))^T$ . The tangent to the centerline  $t(s) = r'(s)$  makes an angle  $\theta(s)$  with the horizontal axis so that the tangent is parametrized as  $t(s) = (\cos \theta(s), \sin \theta(s))$ . Let the rod have length  $2L$  and the arclength vary in the interval  $[-L, L]$ . We are looking for solutions with even symmetry in position  $r(-s) = r(s)$  and odd symmetry in orientation  $\theta(-s) = -\theta(s)$ . The objective function reads  $J = \frac{EI}{2} \int_{-L}^L \theta'(s)^2 ds$ . The boundary conditions in orientation are  $\theta(-L) = \theta(L) = 0$ . The boundary conditions in position are  $x(L) - x(-L) = 2L - \delta$  and  $y(L) - y(-L) = 0$ , so that the end-to-end distance of the rod is the total length  $2L$  minus a compression factor  $\delta \in [0, 2L]$  and the relative height between both ends vanishes. The equation  $r' = (x, y)' = (\cos \theta, \sin \theta)$  allows us to get rid of the two variables  $x, y$  (which only appear in the boundary conditions) by expressing the boundary conditions in isoperimetric form:  $\int_{-L}^L \cos \theta(s) ds = 2L - \delta$  and  $\int_{-L}^L \sin \theta(s) ds = 0$ . The second constraint always holds for an odd function  $\theta$  (since the sine function is odd and the integration interval is centered around 0) and one may take the corresponding Lagrange multiplier to be 0. Let  $F$  denote the Lagrange multiplier enforcing the  $x$ -constraint. The equations of equilibrium for the



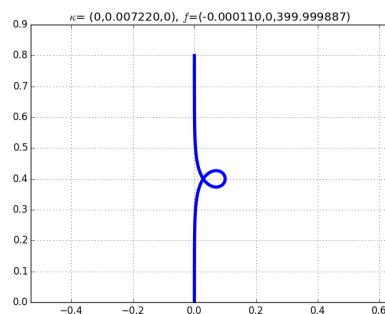
(a)



(b)



(c)



(d)

Figure 6.1: (a) 3d view of a wave-like solution and its material frame. (b) Planar wave-like solution. (c) Planar wave-like solution. (d) Planar orbit-like solution.

rod read

$$\begin{cases} EI\theta''(s) + F \sin \theta(s) = 0 \\ \int_{-L}^L \cos \theta(s) ds = 2L - \delta \\ \theta(L) = \theta(-L) = 0. \end{cases}$$

Now let us rescale this equation by introducing  $\bar{s} = s/L$ ,  $\bar{\theta}(\bar{s}) = \theta(s/L)$ ,  $\bar{F} = \frac{FL^2}{EI}$  and  $\epsilon = \sqrt{\frac{\delta}{2L}}$ . The rescaled equilibrium equations read (omitting the bars)

$$\begin{cases} \theta''(s) + F \sin \theta(s) = 0 \\ \frac{1}{2} \int_{-1}^1 \cos \theta(s) ds = 1 - \epsilon^2 \\ \theta(1) = \theta(-1) = 0. \end{cases} \quad (6.6)$$

Now one proceeds to construct an approximate analytical solution of (6.6) by using perturbation theory with the rescaled compression  $\epsilon$  as a perturbative parameter. Introduce the Ansatz

$$\theta_\epsilon(s) = \epsilon \theta_{[1]}(s) + \epsilon^3 \theta_{[3]}(s) + O(\epsilon^5)$$

and

$$F_\epsilon = F_{[0]} + \epsilon^2 F_{[2]} + O(\epsilon^4),$$

then plug it into (6.6). Regrouping in powers of  $\epsilon$  we find that

$$\epsilon \left( \theta_{[1]}'' + F_{[0]} \theta_{[1]} \right) + \epsilon^3 \left( \theta_{[3]}'' + F_{[0]} \theta_{[3]} + F_{[2]} \theta_{[1]} - \frac{F_{[0]}}{6} \theta_{[1]}^3 \right) + O(\epsilon^5) = 0$$

holds for the ODE part. Whereas the equation

$$\epsilon^2 \left( \frac{1}{4} \int_{-1}^1 \theta_{[1]}(s)^2 ds - 1 \right) + \epsilon^4 \left( \frac{1}{2} \int_{-1}^1 \frac{\theta_{[1]}(s)^4}{4!} - \theta_{[1]}(s) \theta_{[3]}(s) ds \right) + O(\epsilon^6) = 0$$

holds for the constraint part, the boundary conditions being  $\theta_{[i]}(-1) = \theta_{[i]}(1) = 0$ . The order 1 problem reads

$$\begin{cases} \theta_{[1]}''(s) + F_{[0]} \theta_{[1]}(s) = 0 \\ \frac{1}{4} \int_{-1}^1 \theta_{[1]}(s)^2 ds = 1 \\ \theta_{[1]}(1) = \theta_{[1]}(-1) = 0 \end{cases} \quad (6.7)$$

and is a linear second order homogeneous ODE, the two boundary conditions fix the integration constants and the integral constraint fixes  $F_{[0]}$ . The first upward buckling solution of (6.7) is simply

$$\begin{aligned} \theta_{[1]}(s) &= -2 \sin(\pi s) \\ F_{[0]} &= \pi^2 \\ x(s) &= (1 - \epsilon^2)(s + 1) + \frac{\epsilon^2}{2\pi} \sin 2\pi s \\ y(s) &= \frac{2\epsilon}{\pi} (1 + \cos \pi s) \\ \kappa(s) &= -2\epsilon \pi \cos(\pi s), \end{aligned} \quad (6.8)$$

where we remind that, all quantities are unscaled (the bars have been omitted). Knowing  $\theta_{[1]}$  and  $F_{[0]}$  we can proceed to formulate the order 3 problem, which is again linear but is non homogeneous. It reads

$$\begin{cases} \theta_{[3]}'' + F_{[0]} \theta_{[3]} = -F_{[2]} \theta_{[1]} + \frac{F_{[0]}}{6} \theta_{[1]}^3 \\ \int_{-1}^1 \theta_{[1]}(s) \theta_{[3]}(s) ds = \int_{-1}^1 \frac{\theta_{[1]}(s)^4}{24} ds \\ \theta_{[3]}(1) = \theta_{[3]}(-1) = 0 \end{cases} \quad (6.9)$$

and its solution is given by

$$\begin{aligned} \theta_{[3]}(s) &= -\frac{1}{24}(7 + 2 \cos(2\pi s)) \sin 2\pi s \\ F_{[2]} &= \frac{\pi^2}{4}. \end{aligned} \quad (6.10)$$

**From small to large compression: Numerical solutions.** The perturbative solutions for the clamped elastica in Equations (6.8) and (6.10) are valid in the small compression regime. For higher values of the compression rate, the analytical expressions obtained are not that accurate, but still pretty good: at least they verify the boundary conditions. We can hence use these expressions to obtain initial guesses for  $(\kappa(-1), F_x, F_y)$  and use them as input in a shooting scheme. This is a good way to use the indirect approach. One of the weaknesses of the indirect approach is the extreme sensitivity to the initial guess (as will be illustrated shortly in Figures 6.6). Using the solution of a linearized version of the problem as input to the indirect shooting problem helps tremendously.

**Indirect single shooting for the clamped Elastica with arbitrary compression.** The (rescaled) problem to solve is to find  $(\kappa(-1), F_x, F_y)$  which solve the BVP

$$\left\{ \begin{array}{lcl} \dot{x} & = & \cos \theta \\ \dot{y} & = & \sin \theta \\ \dot{\theta} & = & \kappa \\ \dot{\kappa} & = & -F_x \sin \theta + F_y \cos \theta \\ 0 & = & x(-1) = y(-1) = \theta(-1) \\ x(1) & = & 2 - 2\epsilon^2 \\ 0 & = & y(1) = \theta(1). \end{array} \right. \quad (6.11)$$

To solve the BVP (6.11) we setup a residual function which takes as input  $(\kappa, F_x, F_y)_{guess}$ , solves the forward problem using Scipy's odeint<sup>4</sup> and outputs  $(x_1, y_1, \theta_1)_{computed} - (x_1, y_1, \theta_1)_{target}$ . We then solve the system of three equations and three unknowns using a Newton method (Scipy's fsolve). The Jacobian of the residual is computed through finite differences internally by fsolve. In Figure 6.2 we plot for several values of  $\epsilon$  the position of the rod cenetrline and the values of  $\theta$  and  $\kappa$ . The lower row of Figure 6.2 corresponds to the order 1 solutions in Equation (6.8), the upper row corresponds to the results obtained using the shooting algorithm and taking as initial guess the values of  $\kappa, F_x, F_y$  suggested by the order 1 solutions (6.8). The corresponding approximate analytical solutions including the next terms (6.10) in the series expansion are depicted in Figure 6.3. When using the approximate analytical solution (6.8) as an initial guess we find that the shooting method converges in around 20ms on a PC with an Intel core i7 2.6GHz precessor and that furthermore it yields the upward buckling solution with lowest energy. In Figure 6.4 we plot the difference between the computed solution and the order 1 and 3 approximations. Clearly for small values of  $\epsilon$  the difference is very small, and even for a large compression of 80% the relative difference is in a neighbourhood the 10% mark.

**Shooting without a good initial guess.** In Figure 6.5 we plot computed solutions for the fixed-fixed elastica problem for different values of the compression rate, where the initial guesses for  $\kappa(-1)$  and  $F_x$  were taken randomly. This makes clear the sensitivity of the shooting procedure with respect to the initial guess. Furthermore it is observed that the resulting solutions are of much higher energy than those obtained in Figure 6.2 where a first order guess was used. The computation time of the shooting algorithm is also affected, using random initial guesses it might not converge at all or take significantly longer (1s instead of 20 ms) than when initializing smartly, hardly a surprise. Clearly the problem has many local solutions and the Newton scheme, when successful, converges to the closest one from the initial guess. To have a better understanding of the set of possible solutions we decided to plot the residual function for a fixed

---

<sup>4</sup>Scipy's odeint interfaces lsoda from odepak, it is a variable step size method which can switch between Adams and BDF.

### Clamped Euler Elastica

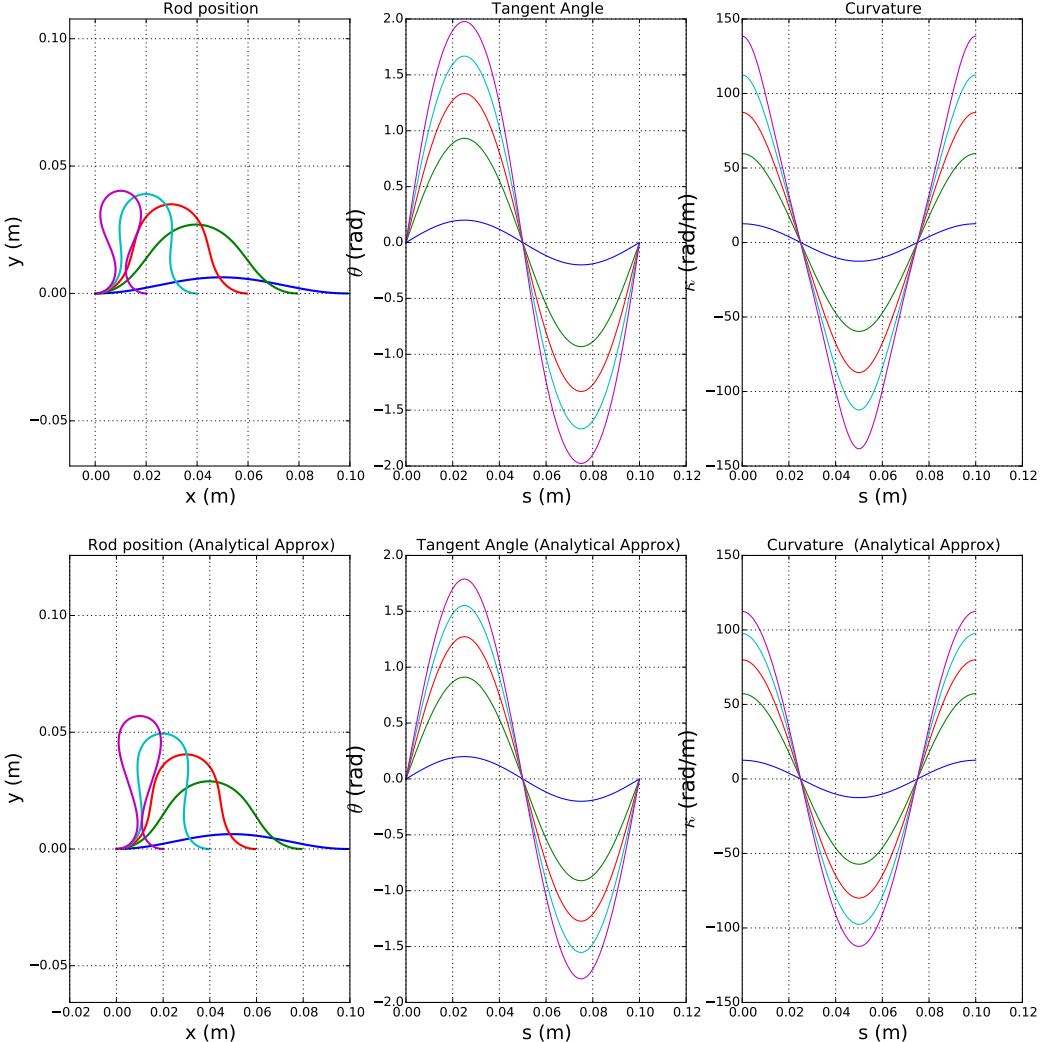


Figure 6.2: Sequence of order 1 approximations from equation (6.8) for different values of  $\epsilon$  taken as input(lower row) and output of the shooting algorithm (upper row). These simulations are of a rod with the characteristics of a hair strand: length 10cm, radius  $50\mu m$  and Young's modulus  $10MPa$ .

value of the compression and for varying initial conditions. In Figure 6.6 we take  $F_y = 0$  and we vary  $F_x$  and  $\kappa(-1)$ , for each one of these initial conditions we solve the forward problem and measure the distance to the target. Blue regions then correspond to initial conditions which take us close to the target, red regions take us far from the target. Deep blue sets of valleys in stripe like patterns appear for increasing values of the internal force  $F_x$ , in each one of these valleys is a solution. The further to the right, the more inflection points (or waves) the centerline has. For negative values of  $\kappa$  the rod buckles down, for positive ones it buckles up. The approximate analytical expression makes us land near the first valley, from left to right, in the upper part. The Figure 6.6 then shows that shooting without a good initial guess can easily yield almost

### Clamped Euler Elastica

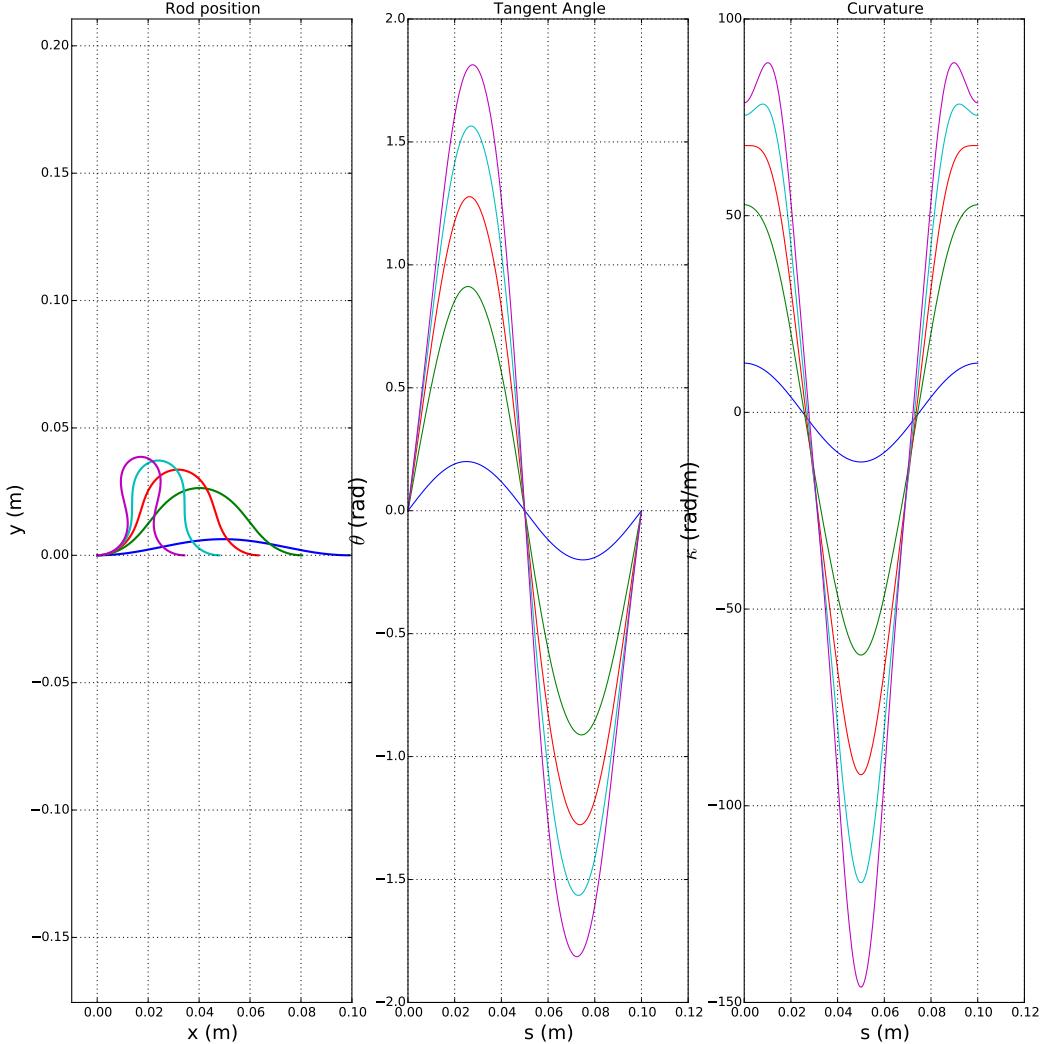


Figure 6.3: Sequence of approximations from equation (6.10) for different values of  $\epsilon$ . Same parameter values as for Figure 6.2

anything, there is no good control over where the shooting algorithm will converge to, or if it will converge at all.

## 6.2 Discrete Kinematics

In this section we take a closer look at the kinematics of rods and their discretization. We present the Runge-Kutta-Munthe-Kaas methods that will be used in the implementation of direct methods later on.

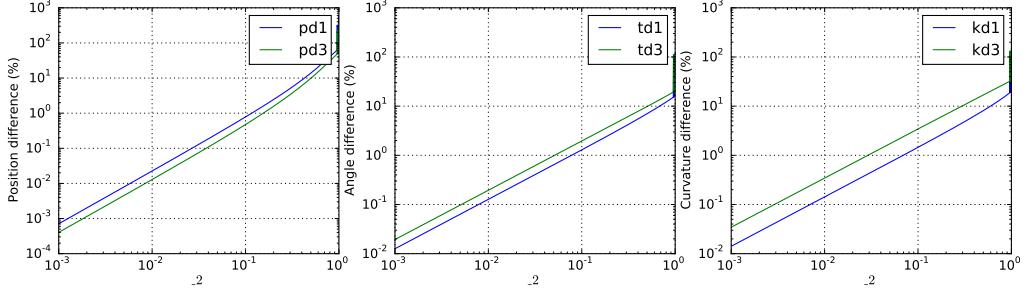


Figure 6.4: Difference between the shooting computed solutions and the order 1 and 3 approximations as a function of the compression  $\epsilon^2$ . The blue curves correspond to the first order solution and green to the third order solution

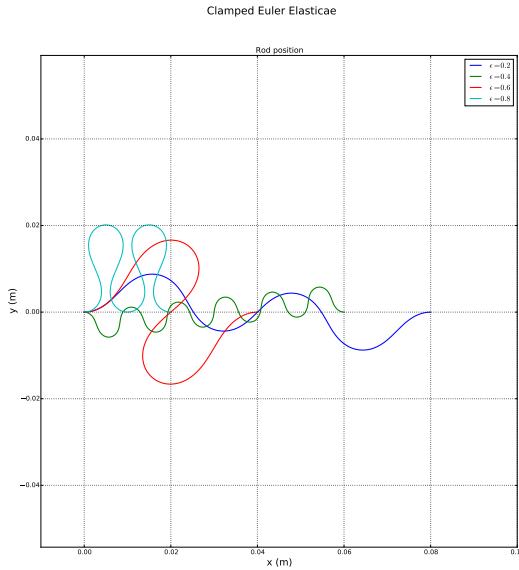


Figure 6.5: Computed solutions for the clamped-clamped elastica problem for different values of the compression rate. Initial guesses for  $\kappa(-1)$  and  $F_x$  were taken randomly.

### 6.2.1 On the choice of the integrator to discretize the Kinematics.

In theory, to discretize the Kinematic constraints of problem (6.1) one could use any numerical method for ODEs (for example any RK method, or even any multistep method, implicit or explicit). However such methods, which are extensively employed in numerical optimal control [100] don't preserve the orthogonality of the frame  $R$ . To employ them one would have to use one of the three following options.

The first option is to use them as such and to take an extremely fine discretization step so that the drift of  $R$  from  $SO(3)$  remains acceptably small. It is also possible to use Baumgarte stabilization to damp the drift from the constraint manifold [30]. Yet for our purposes, such an approach discards the possibility of having qualitatively correct yet coarse and computationally cheap discretizations. It would also add an un-physical parameter to the system.

The second approach would be to use a projection method, so enforcing for instance  $R^T R = I$  as a further constraint using Lagrange multipliers. Similarly if we use unit quaternions  $q$  instead of rotation matrices then we would have to impose  $\|q\| = 1$  as a constraint. The resulting

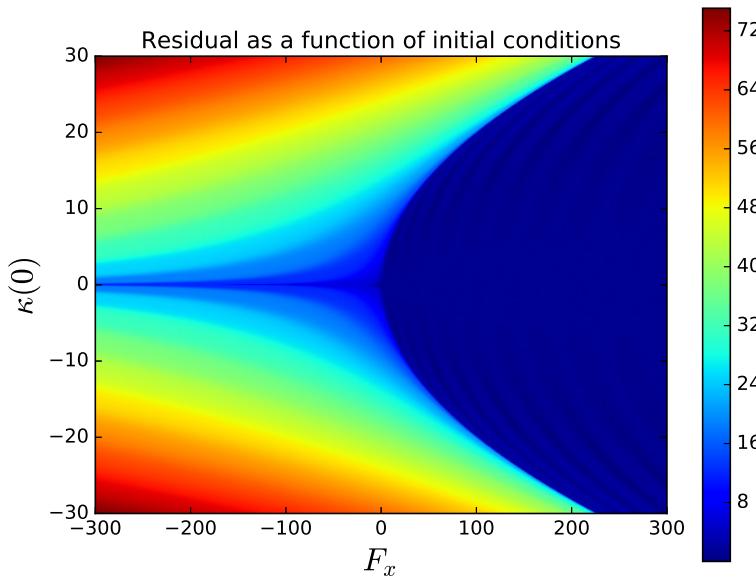


Figure 6.6: Map of the shooting residuals for  $\epsilon = 0.8$  taking  $F_y = 0$  and varying  $\kappa(-1)$  and  $F_x$ . The reader is invited to consult the PDF version of this document if the print quality isn't good enough.

optimal control problem formulation would be one with differential algebraic equations (DAE) as constraints. In such a case what is typically done in the Optimal Control community is to employ Radau Collocation methods which are a subfamily of implicit RK methods well suited for the discretization of DAEs [30]. It is perhaps the right time to mention that there are optimal control software packages like BOCOP [101] which implement the aforementioned direct methods for OCP's with DAE constraints. The reason why we don't use these methods is because we're interested in developing strain based methods for rod statics and strains actually offer a natural parametrization of our configuration manifold. So using a parametrization method seems more fitting in our context.

The third option is to discretize the kinematics using numerical methods based on a local parametrization. The approach we use falls into this category. Because  $SO(3)$  is a 3 dimensional manifold it is possible to parametrize it using only 3 variables, although not globally [102] without introducing singularities. A typical parametrization is given by Euler angles, which, however, have the disadvantage of being subject to gimbal lock. The parametrization variables we employ, as mentioned above, are the strains of the rod. The parametrization maps we use are the exponential map and the Cayley map. We will describe shortly these maps in more detail and explain how to discretize the kinematics of Kirchhoff rods using Runge-Kutta- Munthe-Kaas (RKMK) methods [103]. RKMK methods have the advantage that they preserve the Lie group structure of the problem by construction.

**Using implicit integrators for the rotation reconstruction  $R' = R\hat{\kappa}$  is unnecessary.** The reason is that the reconstruction equation is not stiff. Writing explicitly the ODE in terms

of the column vectors  $d_i$  we find that

$$\begin{aligned} d'_1 &= +\tau d_2 - \kappa_2 d_3 \\ d'_2 &= -\tau d_1 + \kappa_1 d_3 \\ d'_3 &= \kappa_2 d_1 - \kappa_1 d_2. \end{aligned}$$

The Jacobian of the right-hand side reads  $\begin{bmatrix} 0_{3 \times 3} & \tau I & -\kappa_2 I \\ -\tau I & 0_{3 \times 3} & \kappa_1 I \\ \kappa_2 I & -\kappa_1 I & 0_{3 \times 3} \end{bmatrix}$  and has zero and purely imaginary eigenvalues  $(0, +i\|\kappa\|, -i\|\kappa\|)$  each with multiplicity 3. There are no destabilizing eigenvalues with large negative real parts, the equation is non stiff and one can use explicit integrators instead of implicit ones.

### 6.2.2 Runge-Kutta-Munthe-Kaas methods

In the following we shall present the construction of Runge-Kutta-Munthe-Kaas methods. After some preliminary definitions and results we expose the change of variables at the heart of RKMK methods and then proceed to give the explicit expressions for some explicit RKMK integrators.

#### Description

As explained in [104] Munthe-Kaas methods are nothing but numerical integrators for ODE's on manifolds (specifically Lie groups) which are based on a parametrization. In our case the configuration manifold on which the reconstruction ODE takes place is the Lie group  $SE(3)$ . The key is to introduce a map between the Lie algebra  $\mathfrak{se}(3)$ , which is a linear space, and the Lie group  $SE(3)$ , which is a non linear space. In the original construction by Munthe-Kaas [103] the map considered was the exponential map, which can be defined on any Lie group.

Take for instance  $SO(3)$ . Any rotation matrix  $R \in SO(3)$  can be expressed as the exponential of some skew symmetric matrix  $\hat{\kappa} \in \mathfrak{so}(3)$ :

$$\forall R \in SO(3) \exists \hat{\kappa} \in \mathfrak{so}(3) : R = \exp(\hat{\kappa}).$$

The exponential is the usual matrix exponential  $\exp(A) = \sum_{i=0}^{\infty} \frac{A^i}{i!}$ . In general one would truncate the sum, although for  $SO(3)$  it is not necessary, as thanks to Rodrigues' formula it has a closed form expression. For any vector  $\kappa \in \mathbb{R}^3$  the exponential of the skew symmetric matrix  $\hat{\kappa} \in \mathfrak{so}(3)$  reads

$$\exp_{SO(3)} \hat{\kappa} = \begin{cases} I + \frac{\sin \|\kappa\|}{\|\kappa\|} \hat{\kappa} + \frac{1-\cos \|\kappa\|}{\|\kappa\|^2} \hat{\kappa}^2 & \text{if } \kappa \neq 0 \\ I & \text{if } \kappa = 0. \end{cases} \quad (6.12)$$

Similarly, for the special euclidean group  $SE(3)$ , the exponential  $\exp_{SE(3)}$  maps any element  $(\hat{\kappa}, v)$  of the Lie algebra  $\mathfrak{se}(3)$  to an element of the Lie group  $SE(3)$ . The exponential on  $SE(3)$  also has a closed form expression, given by

$$\begin{aligned} \exp_{SE(3)} \begin{bmatrix} \hat{\kappa} & v \\ 0 & 0 \end{bmatrix} &= \begin{bmatrix} \exp_{SO(3)} \hat{\kappa} & B(\kappa)v \\ 0 & 1 \end{bmatrix}, \\ \text{where } B(\kappa) &= \begin{cases} I + \frac{1-\cos \|\kappa\|}{\|\kappa\|^2} \hat{\kappa} + \frac{\|\kappa\| - \sin \|\kappa\|}{\|\kappa\|^3} \hat{\kappa}^2 & \text{if } \kappa \neq 0 \\ I & \text{otherwise.} \end{cases} \end{aligned} \quad (6.13)$$

For quadratic Lie groups the Cayley map can also be employed (see [104] section IV.8.3). For  $SO(3)$  the Cayley map  $\text{cay} : \mathfrak{so}(3) \rightarrow SO(3)$  reads

$$\text{cay}(\hat{\kappa}) = I + \frac{4}{4 + \|\kappa\|^2} (\hat{\kappa} + \frac{\hat{\kappa}^2}{2}). \quad (6.14)$$

For the special Euclidean group the Cayley map  $\text{cay}_{SE(3)} : \mathfrak{se}(3) \rightarrow SE(3)$  reads

$$\text{cay}_{SE(3)}(\hat{\kappa}, v) = \begin{bmatrix} \text{cay}_{SO(3)}(\hat{\kappa}) & d\text{cay}_{\hat{\kappa}}v \\ 0 & 1 \end{bmatrix}, \quad (6.15)$$

where

$$d\text{cay}_{\hat{\kappa}} = \frac{2}{4 + \|\kappa\|^2} (2I + \hat{\kappa}).$$

Having defined the exponential and Cayley maps for  $SE(3)$  we can proceed to the construction of RKMK methods for the Kirchhoff reconstruction equation. Let us denote  $\xi(s) \in \mathfrak{se}(3)$  the strain field, which for the Kirchhoff reconstruction equation is of the form  $\xi(s) := \begin{bmatrix} \hat{\kappa}(s) & e_3 \\ 0 & 0 \end{bmatrix}$ . Let  $H(s) \in SE(3)$  gather the frame and position of the rod in homogeneous coordinates. The frame reconstruction problem is then to solve the following ODE for given initial data  $H_0 \in SE(3)$  and a given strain field:

$$\begin{cases} H'(s) = H(s)\xi(s) \\ H(0) = H_0. \end{cases} \quad (6.16)$$

We introduce as in [62] a map  $\tau : \mathfrak{se}(3) \rightarrow SE(3)$  which parametrizes the special Euclidean group, which is analytic and verifies  $\tau(-\theta)\tau(\theta) = I$  for all  $\theta \in \mathfrak{se}(3)$ . In practice we will use either the exponential map  $\tau = \exp_{SE(3)}$ , or the Cayley map  $\tau = \text{cay}_{SE(3)}$ . Let us denote  $D\tau_\theta : T_\theta \mathfrak{se}(3) \rightarrow T_{\tau(\theta)} SE(3)$  the usual derivative of  $\tau$  at  $\theta$ . Since  $\mathfrak{se}(3)$  is a linear space one has the identification  $T_\theta \mathfrak{se}(3) \approx \mathfrak{se}(3)$ . Let us introduce the right trivialized tangent map  $d\tau_\theta : \mathfrak{se}(3) \rightarrow \mathfrak{se}(3)$  which is defined as the only linear map such that  $D\tau_\theta(\delta) = d\tau_\theta(\delta) \cdot \tau(\theta)$  for all  $\delta \in \mathfrak{se}(3)$ . The inverse of  $d\tau$  will be frequently used. We will denote it  $d\tau_\theta^{-1} := (d\tau_\theta)^{-1}$ .

**Lemma 6.2.1.** *For all  $\theta \in \mathfrak{se}(3)$  and all  $\delta \in \mathfrak{se}(3)$  the identity*

$$d\tau_\theta(\delta) = \text{Ad}_{\tau(\theta)}(d\tau_{-\theta}(\delta)) = \tau(\theta)d\tau_{-\theta}(\delta)\tau(\theta)^{-1} \quad (6.17)$$

holds.

*Proof.* Let  $\delta$  and  $\theta$  be arbitrary elements of  $\mathfrak{se}(3)$ . Consider a curve  $\theta(s)$  such that  $\theta(0) = 0$  and  $\theta'(0) = \delta$ . One has  $\tau(-\theta(s)) = \tau(\theta(s))^{-1}$  for all  $s$ . Differentiating with respect to time and evaluating as  $s = 0$  one has  $D\tau_{-\theta}(-\delta) = -\tau(\theta)^{-1}d\tau_\theta(\delta)\tau(\theta)\tau(\theta)^{-1}$ . Simplifying the expression at the right-hand side and applying the definition of the right trivialized tangent map, one obtains  $d\tau_{-\theta}(-\delta)\tau(-\theta) = -\tau(\theta)^{-1}d\tau_\theta(\delta)$ . Hence  $\tau(\theta)d\tau_{-\theta}(\delta)\tau(\theta)^{-1} = d\tau_\theta(\delta)$  and the identity  $d\tau_\theta(\delta) = \text{Ad}_{\tau(\theta)}(d\tau_{-\theta}(\delta))$  holds.  $\square$

**Lemma 6.2.2.** *For all  $\theta \in \mathfrak{se}(3)$  and all  $\delta \in \mathfrak{se}(3)$  the identity*

$$d\tau_\theta^{-1}(\delta) = (d\tau_{-\theta})^{-1}(\text{Ad}_{\tau(-\theta)}(\delta)) \quad (6.18)$$

holds.

*Proof.* Let  $\delta$  and  $\theta$  be arbitrary elements of  $\mathfrak{se}(3)$ . Depart from the identity  $d\tau_\theta\eta = \delta$  where we denote  $\eta := d\tau_\theta^{-1}(\delta)$ . Left multiply the previous equality by  $\tau(\theta)^{-1}$  and right multiply by  $\tau(\theta)$ , then  $\tau(\theta)^{-1}d\tau_\theta\eta\tau(\theta) = \tau(\theta)^{-1}\delta\tau(\theta)$ , hence  $\text{Ad}_{\tau(-\theta)}(d\tau_\theta\eta) = \text{Ad}_{\tau(-\theta)}(\delta)$ . Using Lemma 6.2.1 on the left-hand side of the last equality one obtains  $d\tau_{-\theta}(\eta) = \text{Ad}_{\tau(-\theta)}(\delta)$ . Input the definition of  $\eta$  to obtain  $d\tau_{-\theta}(d\tau_\theta^{-1}(\delta)) = \text{Ad}_{\tau(-\theta)}(\delta)$ . Thus we can conclude that  $d\tau_\theta^{-1}(\delta) = (d\tau_{-\theta})^{-1}(\text{Ad}_{\tau(-\theta)}(\delta))$  as announced in the lemma.  $\square$

Munthe-Kaas methods are based on a change of variables. Instead of solving for  $H(s)$  which lies on the nonlinear manifold  $SE(3)$ , one introduces an unknown  $\theta(s) \in \mathfrak{se}(3)$  which lies in a linear space.

**Proposition 6.2.1.** *By introducing the ansatz  $H(s) = H_0\tau(\theta(s))$ , the reconstruction equation (6.16) becomes*

$$\begin{cases} \theta' &= d\tau_{-\theta}^{-1}\xi \\ \theta(0) &= 0_{\mathfrak{se}(3)} \\ H(s) &= H_0\tau(\theta(s)). \end{cases} \quad (6.19)$$

*Proof.* • First, let us assume that  $H$  satisfies (6.16) and introduce  $\theta$  such that  $H(s) = H_0\tau(\theta(s))$ . Then one has  $H' = H\xi = H_0\tau(\theta)\xi$  and on the other hand  $H' = H_0D\tau_\theta(\theta') = H_0d\tau_\theta(\theta')\tau(\theta)$ . Equating both expressions for  $H'$  one obtains  $\xi = \tau(\theta)^{-1}d\tau_\theta(\theta')\tau(\theta)$ . By application of Lemma 6.2.1 the expression for  $\xi$  simplifies to  $\xi = d\tau_{-\theta}(\theta')$  and one can conclude that  $\theta' = d\tau_{-\theta}^{-1}\xi$  holds as announced. And if  $H(0) = H_0$  then  $\tau(\theta(0)) = I$  so that the initial condition for  $\theta$  is indeed  $\theta(0) = 0_{\mathfrak{se}(3)}$ .

- Conversely, let us assume that  $\theta' = d\tau_{-\theta}^{-1}\xi$  and that  $\theta(0) = 0_{\mathfrak{se}(3)}$ . Then the function  $H(s) := H_0\tau(\theta(s))$  verifies the initial condition  $H(0) = H_0$  as well as the ODE  $H' = H\xi$ . Indeed, by differentiation  $H' = H_0d\tau_\theta(\theta')\tau(\theta)$ . Inputting the expression for  $\theta'$  one has  $H' = H_0d\tau_\theta(d\tau_{-\theta}^{-1}\xi)\tau(\theta)$ . Applying Lemma 6.2.2 one has  $H' = H_0d\tau_\theta(d\tau_\theta^{-1}\text{Ad}_{\tau(\theta)}(\xi))\tau(\theta)$ , which simplifies to  $H' = H_0\tau(\theta)\xi$  once the definition of  $\text{Ad}$  is applied. Hence  $H' = H\xi$ .  $\square$

Hence to solve the reconstruction equation (6.16) a RKMK integrator proceeds as follows. First use any Runge-Kutta method to solve the differential equation (6.19) which takes place in the linear space  $\mathfrak{se}(3)$ . Then map the solution back to the group using the map  $\tau$ .

We remind the reader the definition of a Runge-Kutta integrator.

**Definition 6.2.1** (Runge-Kutta integrator [104]). To solve an ode  $y' = f(t, y)$ , an  $s$  stage Runge-Kutta method with real coefficients  $a_{ij}$ ,  $b_i$  ( $i, j$  ranging from 1 to  $s$ ) and  $c_i = \sum_{j=1}^s a_{ij}$  is given by

$$\begin{aligned} k_i &= f(t_0 + c_i h, y_0 + h \sum_{j=1}^s a_{ij} k_j), \quad i = 1 \dots s \\ y_1 &= y_0 + h \sum_{i=1}^s b_i k_i. \end{aligned} \quad (6.20)$$

The  $k_i$  are called 'slopes', and the  $y_{0,i}$  defined as  $y_{0,i} := y_0 + h \sum_{j=1}^s a_{ij} k_j$  are called the 'internal stages'. The coefficients defining an RK method are often summarized in a 'Butcher tableau' of the form

$$\begin{array}{c|c} c_i & a_{ij} \\ \hline & b_j \end{array}$$

We can now proceed to give the expressions for several RKMK one step maps.

**RKMK1** Using the coefficients of the explicit Euler method, which is an explicit one-stage RK method, we can construct the RKMK1 integrator. The Butcher tableau of the explicit Euler method reads

$$\begin{array}{c|c} 0 & 0 \\ \hline & 1 \end{array}$$

and this yields the RKMK1 integrator, more commonly known as the 'Lie-Euler method' [104]. The Lie-Euler update formula reads

$$\begin{bmatrix} R_{i+1} & r_{i+1} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R_i & r_i \\ 0 & 1 \end{bmatrix} \exp_{SE(3)} \left( h \begin{bmatrix} \hat{\kappa}_i & e_3 \\ 0 & 0 \end{bmatrix} \right). \quad (6.21)$$

One can also use the Cayley map, leading to a variation of the Lie-Euler method which reads

$$\begin{bmatrix} R_{i+1} & r_{i+1} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R_i & r_i \\ 0 & 1 \end{bmatrix} \text{cay}_{SE(3)} \left( h \begin{bmatrix} \hat{\kappa}_i & e_3 \\ 0 & 0 \end{bmatrix} \right). \quad (6.22)$$

**RKMK2** One can use the coefficients of an explicit trapezoidal rule to obtain a 2 stage RKMK method which reads

$$\begin{array}{c|cc} 0 & 0 & 0 \\ \hline 1 & 1 & 0 \\ \hline & 1/2 & 1/2 \end{array}$$

$$\begin{aligned} K_1 &= \xi(t_i) \\ K_2 &= d\tau^{-1}(-h\xi(t_i), \xi(t_i + h)) \\ \begin{bmatrix} R_{i+1} & r_{i+1} \\ 0 & 1 \end{bmatrix} &= \begin{bmatrix} R_i & r_i \\ 0 & 1 \end{bmatrix} \tau \left( \frac{h}{2} K_1 + \frac{h}{2} K_2 \right), \end{aligned} \quad (6.23)$$

where  $\xi$  is the strain field  $\xi = (\kappa, e_3)$ .

**RKMK4** One can use the coefficients of a classical explicit RK4 to obtain a 4 stage RKMK method which reads

$$\begin{array}{c|ccc} 0 & & & \\ \hline 1/2 & 1/2 & & \\ 1/2 & 0 & 1/2 & \\ \hline 1 & 0 & 0 & 1 \\ \hline & 1/6 & 2/6 & 2/6 & 1/6 \end{array}$$

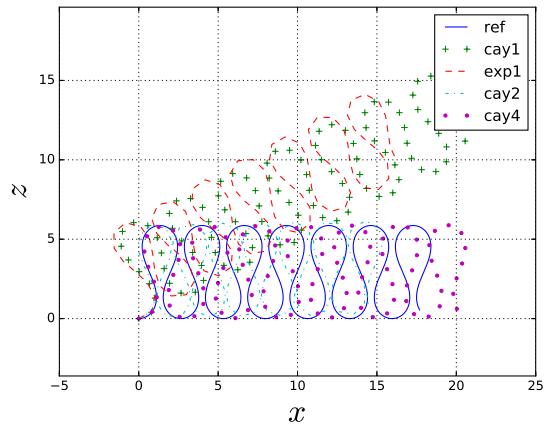
$$\begin{aligned} K_1 &= \xi(t_i) \\ K_2 &= d\tau^{-1}(-\frac{h}{2}\xi(t_i), \xi(t_i + \frac{h}{2})) \\ K_3 &= d\tau^{-1}(-\frac{h}{2}\xi(t_i + \frac{h}{2}), \xi(t_i + \frac{h}{2})) \\ K_4 &= d\tau^{-1}(-h\xi(t_i + \frac{h}{2}), \xi(t_i + h)) \\ \begin{bmatrix} R_{i+1} & r_{i+1} \\ 0 & 1 \end{bmatrix} &= \begin{bmatrix} R_i & r_i \\ 0 & 1 \end{bmatrix} \tau \left( \frac{h}{6} K_1 + \frac{2h}{6} K_2 + \frac{2h}{6} K_3 + \frac{h}{6} K_4 \right), \end{aligned} \quad (6.24)$$

where  $\xi$  is the strain field  $\xi = (\kappa, e_3)$ .

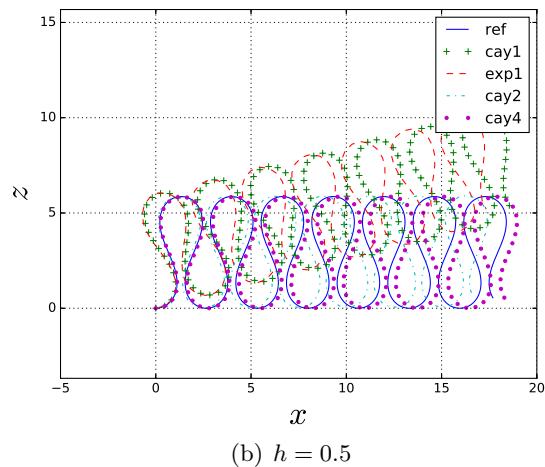
## Evaluation

We will now illustrate the accuracy and computational effort of the different Runge-Kutta Munthe-Kaas methods explicated above. We shall do this by solving the 3D Kirchhoff reconstruction equation with a prescribed strain field of the form  $\kappa_1(s) = k_0 cn(\frac{k_0}{2p}s, p)$ ,  $\kappa_2 = \tau = 0$ , where  $k_0$  and  $p$  are parameters and  $cn$  is the Jacobi elliptic cosine. Such a strain field is not arbitrary, it is the form of the solution of the planar wave-like Elastica as explained in [79]. We take as parameters  $L = 100$ ,  $k_0 = 1.$ ,  $p = 0.8$ . In Figure 6.7 we plot the centerline of the reference solution and compare it with the numerical solutions computed by the different RKMK methods for 3 different stepsizes. In Figure 6.8 we plot the error of each numerical

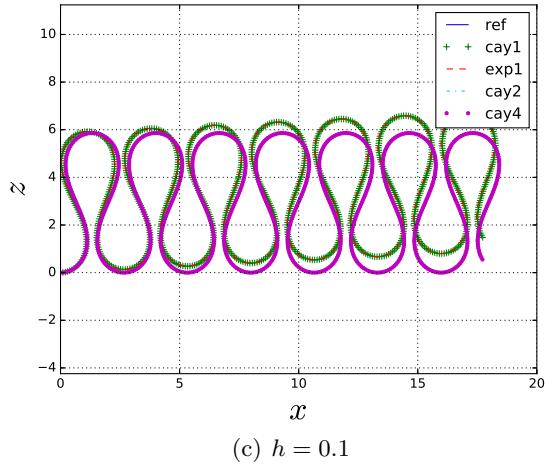
solution as well as the computational load against the number of discretization intervals used. We can see that the one-stage RKMK methods are of order 1 and that for a small enough step size there is virtually no difference between the exponential and the Cayley versions of RKMK1. The RKMK2 method with the Cayley map is seen to be of order 2. There doesn't seem to be much of an advantage in using RKMK4 instead of RKMK2 for this particular example as the error plot shows. As for the computational load, there isn't much of a difference between any of the methods in particular for a low number of elements. These remarks motivate in particular the use of RKMK1 and RKMK2 for the discretization of the reconstruction equations in the Kirchhoff statics OCP.



(a) Step size  $h = 1$ .

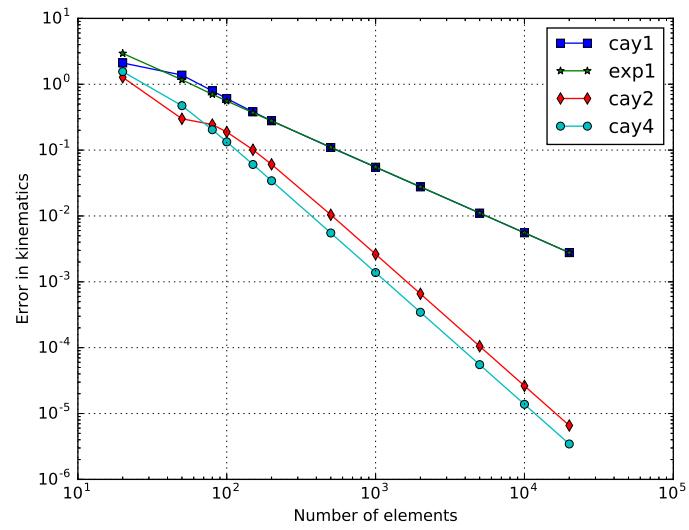


(b)  $h = 0.5$

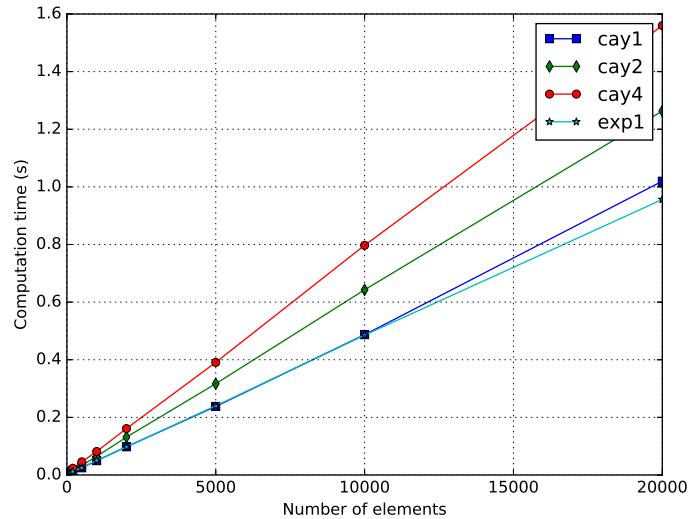


(c)  $h = 0.1$

Figure 6.7: Comparison of the different methods to solve the kinematics in terms of centerline position for different step sizes. Reference solution in solid blue. RKM<sub>K</sub>1 with exp (red dashes) and Cayley (green crosses) maps. RKM<sub>K</sub>2 (turquoise dashes) and RKM<sub>K</sub>4 (violet dots) with Cayley map.



(a)



(b)

Figure 6.8: (a) Accuracy of the different methods to solve the kinematics. (b) Computational effort of the different methods to solve the kinematics.

## 6.3 Direct Methods and the derivation of Strain based and mixed discretizations

### 6.3.1 Direct Single Shooting

In numerical optimal control, 'direct single shooting' is a discretization strategy which consists in projecting the control field onto a finite dimensional subspace and expressing the states as functions of the controls using a numerical integrator. The discretized OCP then has as unknowns only the discrete controls.

In our context this will lead to strain based methods for the statics of Kirchhoff rods. In direct single shooting, we choose a set of basis functions  $\varphi_i : [0, L] \rightarrow \mathbb{R}^3$  so that our approximation of the strain field  $\kappa$  reads  $\kappa(s) = \sum_{i=0}^n \kappa_i \varphi_i(s)$ . The basis functions are typical finite element functions associated with a grid  $(s_0, \dots, s_N)$  subdividing the segment  $[0, L]$ . In practice we will use piecewise constant step functions (0 order discontinuous Galerkin) or continuous piecewise affine hat functions (1st order continuous Galerkin) on a regular grid. The unknowns of the discrete problem will then be the discretized strains. We discretize the reconstruction equations (6.16) using an s-stage RKMK method with coefficients  $a_{ij}, b_j, c_j$ . The objective function of the Kirchhoff OCP (6.1) is discretized using the same weights than for the RKMK method.

#### Direct single shooting with piecewise constant strains.

The most common thing to do in direct single shooting for optimal control is to restrict the controls to the space of piecewise constant functions. For that purpose, let us introduce a regular grid  $(s_0, \dots, s_N)$  on  $[0, L]$  so that  $s_i = ih$  where the step size  $h$  is  $h = \frac{L}{N}$  for some integer  $N$ . The basis functions are then unit on a subinterval and zero elsewhere, which we denote  $\varphi_i(s) = \chi_{[s_i, s_{i+1}[}(s)$ . Furthermore let us introduce the finitely many degrees of freedom  $q = (u_0, \dots, u_{N-1}) \in \mathbb{R}^{3N}$ , where each  $u_i$  is a vector in  $\mathbb{R}^3$  collecting the curvatures and twist of the rod in the interval  $[s_i, s_{i+1}[$ . The ODE constraints of the OCP (6.1) are discretized using an RKMK1 method (6.21) (6.22). Let us denote  $\Psi_h$  the one step map to solve the reconstruction ODE  $r' = Re_3$ ,  $R' = R\hat{k}$  for a given function  $\kappa$  and given initial conditions. The initial state  $H_0 := (R_0, r_0)$  is given, the state  $H(s_{i+1}) = (R(s_{i+1}), r(s_{i+1}))$  is approximated by  $H_{i+1}$  as

$$H_{i+1} = \Psi_h(H_i, u_i) = H_i \tau(h(\hat{u}_i, e_3)).$$

The final state  $H(L) = (R(L), r(L))$  is thus approximated by a function  $H_N(H_0, q)$  of the initial state and of all the strains  $q$ , it is defined by sequentially applying the integrator as

$$H_N = \Psi_h(H_{N-1}, u_{N-1}) = \Psi_h(\Psi_h(H_{N-2}, u_{N-2}), u_{N-1}) = \dots = H_N(H_0, q).$$

The resulting finite dimensional optimization problem is the following

$$\begin{aligned} \min_{q \in \mathbb{R}^{3N}} \quad & \frac{1}{2}(q - \bar{q})^T \mathbf{K}(q - \bar{q}) + U(q) \\ \text{s. t.} \quad & g_I(R_i(q), r_i(q)) \leq 0 \\ & g_{bd}(R_0, r_0, R_N(q), r_N(q)) = 0, \end{aligned} \tag{6.25}$$

where  $U(q) = \sum_{i=0}^{N-1} h U_g(r_i(q))$  is the discrete gravitational potential energy,  $\mathbf{K} := h \text{diag}(EI_1, EI_2, \mu J \dots)$  is the stiffness matrix and  $\bar{q} := (\bar{u}_0, \dots, \bar{u}_{N-1})^T$  collects the natural curvatures and twists of the rod (also discretized as piecewise constants).

Let us comment further on the structure of this problem. There are many more variables than in indirect single shooting and fewer variables than in direct multiple shooting as we will see. The kinematics ODE constraints have disappeared, instead the centerline and the material

frame are given explicit (closed form) expressions of the strains  $q$ . The internal elastic energy remains a quadratic function of the degrees of freedom, whereas the gravitational potential energy is no longer a simple linear function but a complicated nonlinear one. If fixed-fixed boundary conditions are considered then the boundary constraint function depends on all the strains. This leads to **small but dense** linear algebra problems at the Newton iterations level when solving problem (6.25) even in the absence of gravity.

**Mechanical Interpretation.** The single shooting approach with piecewise constant strains and a Lie-Euler integrator coincides with the Super-Helix [1] strain based finite element approach for the statics of Kirchhoff rods.

**Remark.** For the case of piecewise constant strains, on each subinterval  $[s_i, s_{i+1}]$  the constraint ODE is autonomous and linear in the states. It reads  $H' = H\xi$  with  $\xi$  constant. Its exact solution on each subinterval is given by  $H_{i+1} = H_i \exp(h\xi)$ , which is exactly the Lie-Euler method on  $SE(3)$  (6.21). Hence for the case of piecewise constant strains no other method is more appropriate than the Lie-Euler method since it coincides with the exact solution on each subinterval and is cheap to compute thanks to the Rodrigues formula.

**Numerical example.** In Figure 6.9 we plot the obtained numerical solutions for the stable static equilibrium of a Kirchhoff rod in the absence of gravity, without natural curvature and with fixed-fixed boundary conditions. We assumed a symmetric cross section and rescaled the problem as in 6.5 so that the total length is 1 and the relevant parameters are the boundary conditions and the twisting to bending stiffness ratio  $\gamma := \frac{\mu J}{EI}$ . The top row shows a solution for 20 elements, the bottom row a solution for 100 elements. Even though the initial guess was the same for both (all curvatures equal to 1) we obtain different solutions, this not a surprise as the solutions aren't unique. The computed numerical solutions have very low energy even though the initialization was quite arbitrary, a phenomenon which we also observe when initializing randomly. This is in sharp contrast with the indirect methods of the previous section where the algorithm was much more sensitive to the initial guess and tended to yield high energy solutions when initialized randomly. As expected the internal twist of the rod is constant in the absence of gravity. We can also see that the discrete model shows a good behavior even for a coarse discretization.

**Experimental and cross validation.** In this experiment we simulated a fixed-free static equilibrium of a rod under gravity using single shooting with piecewise constant strains and exp-RMK1 with 50 elements. The rod parameters were length  $20\text{cm}$ , radius  $2.23\text{mm}$ , Young's modulus  $E = 1\text{MPa}$ , Poisson's ratio  $\nu = 0.5$ , natural Frenet curvature  $\kappa_F = 66.67\text{m}^{-1}$ , natural Frenet torsion  $\tau_F = -24.79\text{m}^{-1}$  and density  $\rho = 1200\text{kg m}^{-3}$ . These silicone rubber rods are fabricated by Victor Romero, a post doc at our team at the time of writing of this thesis. Using computer vision techniques he is also able to recover the centerline position of the rod at equilibrium. In Figures 6.10 and 6.11 we compare the computed solution with the experimental observation. We also include in Figure 6.10 a numerical solution computed using the implementation of super helix dynamics described in [1]. To obtain a static equilibrium using a dynamics code we include a lot of damping and we wait for the generalized velocities to approximately vanish. We can see that our proposed method, the super helix method [1] and the experimental observation are in good agreement. With respect to computation time, using the dynamics approach takes of course a lot longer (2 minutes in this case) than our statics approach (80 ms), not a fair comparison but it does justify the need to have a code tailored for Kirchhoff rod statics.

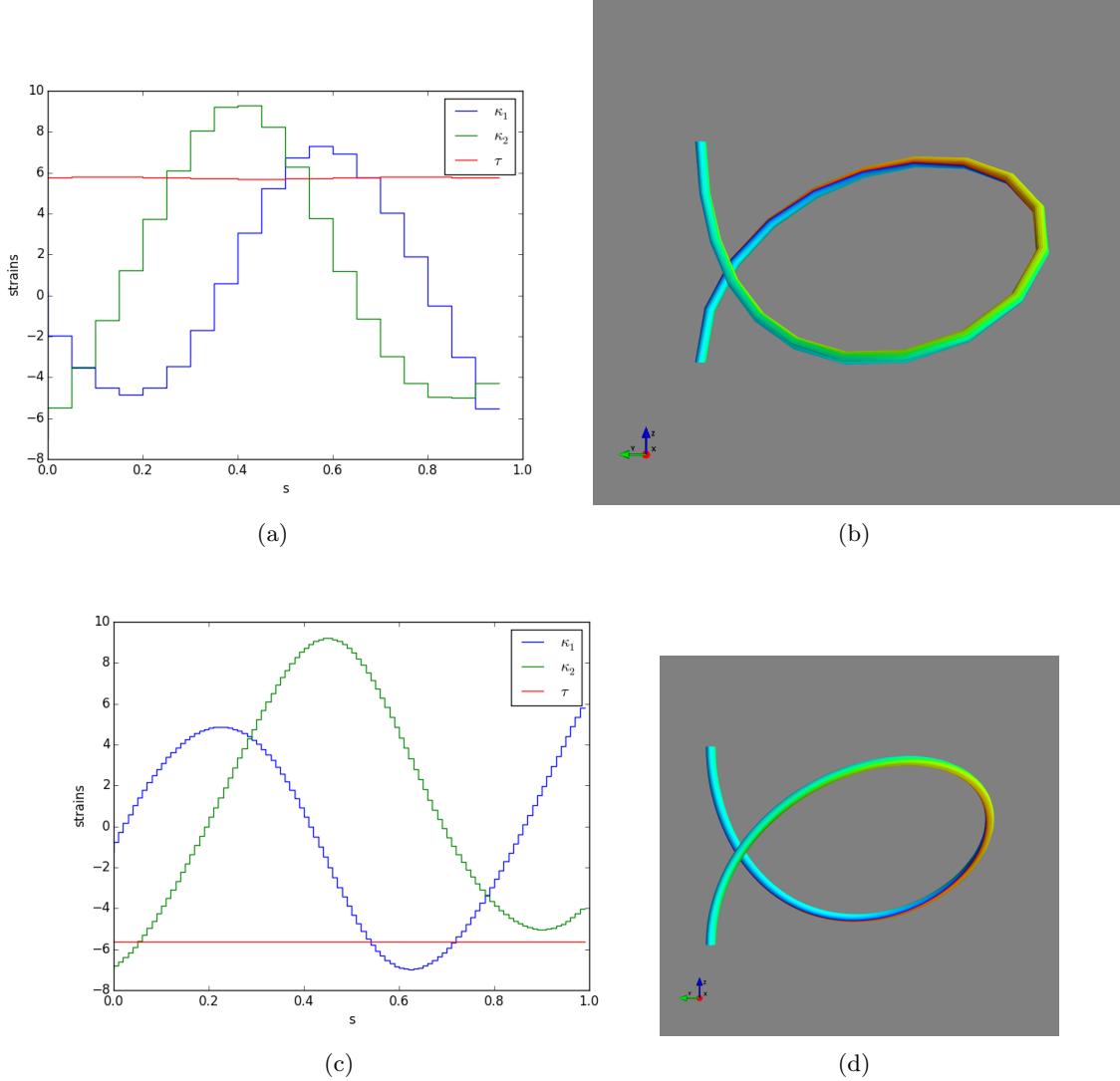


Figure 6.9: Solving a fixed-fixed statics problem with a piecewise constant strain based method and RKMK1. No natural curvature and no gravity,  $\gamma = \frac{\mu J}{EI} 0.7$ , end to end distance 20 percent of rod length. (a) Strains, 20 elements. (b) Rod position, 20 elements. (c) Strains, 100 elements. (d) Rod position, 100 elements

### Direct single shooting with Piecewise linear strains.

Let us again consider a regular grid  $(s_0, \dots, s_N)$  on  $[0, L]$  with step size  $h = L/N$ . Let us discretize the strain field by introducing for every node  $i \in [0, \dots, N]$  the strains  $u_i \in \mathbb{R}^3$  at that node and letting the strain field be piecewise linear. For all  $s$  in  $[s_i, s_{i+1}]$  we have  $\kappa(s) := \frac{s-s_i}{h} u_{i+1} + \frac{s_{i+1}-s}{h} u_i$ . We collect all the nodal strains in a vector  $q := (u_0, \dots, u_N) \in \mathbb{R}^{3(N+1)}$ . Let the natural strain field  $\bar{\kappa}$  be discretized in the same fashion and denote  $\bar{u}_i$  the natural strains at node  $i$ . We denote by  $\Delta u_i$  the difference between actual and natural curvature at node  $i$ . To discretize the kinematics we employ a RKMK2 method with the Cayley map defined in equation (6.23). We use the weights in the Butcher tableau of (6.23) to discretize the objective, that is we employ trapezoidal quadrature, so that the discrete total elastic and gravitational energy

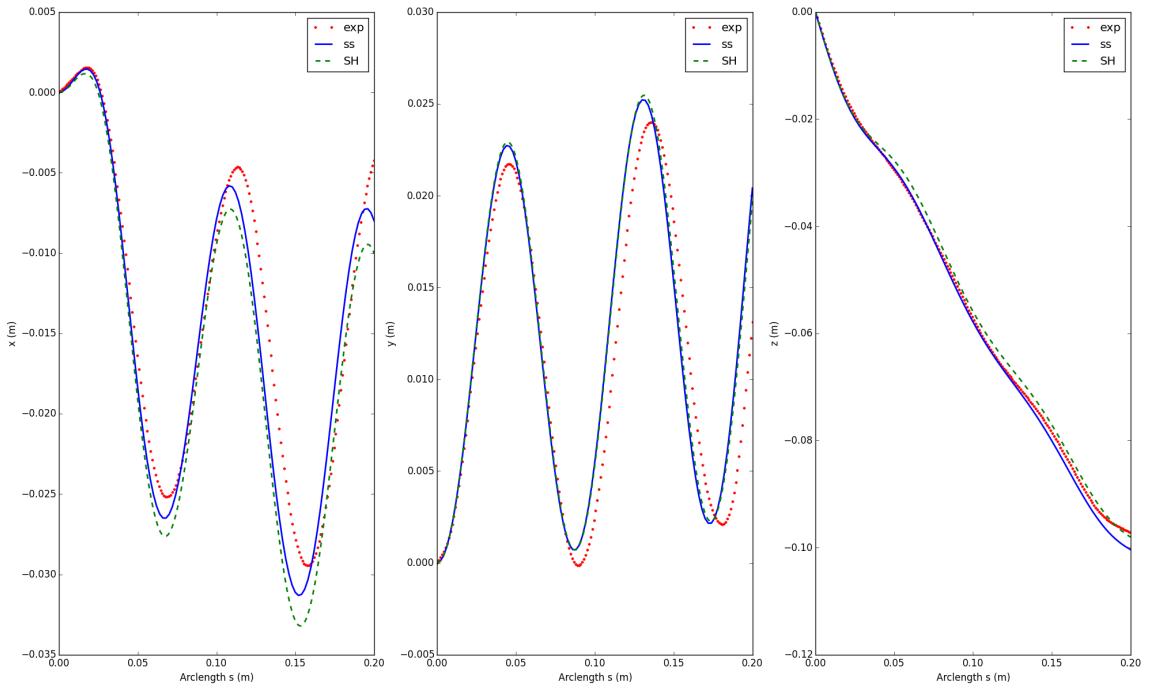


Figure 6.10: Comparison of the centerline positions at equilibrium. In red dots (exp) is the experimental curve. In solid blue (ss) is our single shooting with piecewise constant strains and exp-rkmk1. In green dashes is the equilibrium obtained using the super helix (SH) implementation by the authors of [1].

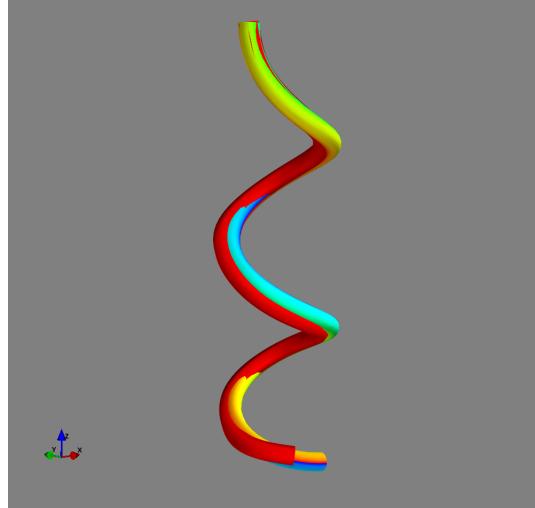


Figure 6.11: 3D view of centerline positions at equilibrium. In red is the experimental curve. In colors is our computed solution with single shooting, piecewise constant strains and exp-rkmk1.

reads

$$U_{tot} = \sum_{i=0}^{N-1} \frac{h}{4} \Delta u_i^T K \Delta u_i + \frac{h}{4} \Delta u_{i+1}^T K \Delta u_{i+1} + h \rho g S \frac{z_i(q) + z_{i+1}(q)}{2}.$$

As in direct single shooting with piecewise constant strains, the height  $z_i$  of the rod at node  $i$  is a nonlinear function of the strains  $q$  so that the gravitational potential energy is non linear. The internal elastic energy is again quadratic. The total number of degrees of freedom is  $3(N + 1)$  so just 3 more degrees of freedom than in the piecewise constant strain version. However the discrete strain field is now continuous. A closer look at the relative performance and precision of these methods will be done shortly.

**Numerical application.** We solve a fixed-fixed boundary value problem in the absence of gravity with the method of single shooting with piecewise linear strains and plot the results in Figure 6.12. We again use a rod with a symmetric cross section and solve the rescaled problem with a twisting to bending stiffness ratio of  $\gamma = 0.7$  and an end to end compression of 50%. We again initialize the strains to be all equal to 1 and solve the problem for 20 elements and 100 elements. Both obtained numerical solutions were in this instance the same. We again observe good behavior even for a coarse discretization. The material twist is constant as expected. The obtained numerical solution is actually the global minimum although we used an arbitrary initial guess.

### 6.3.2 Direct Multiple Shooting

The second branch of direct methods is 'direct multiple shooting' where both the strains  $\kappa$  and the states  $(R, r)$  are discretized, leading to a *mixed formulation*. The NLP in the mixed formulation is of higher dimensionality but with more sparsity and simpler non linearities. Using Lie group methods to discretize the kinematics allows us to avoid formulating supplementary orthogonality constraints for the frame  $R_i$  at each node  $s_i$ . The direct multiple shooting approach may be seen as a 'lifting' of the direct single shooting approach. The idea is to trade a small but highly nonlinear problem for one that is larger but less nonlinear, so trading off small size for better structure.

The implementation of multiple shooting that we will describe employs piecewise constant strains. The variables in direct multiple shooting are not only the discrete strains  $q \in \mathbb{R}^{3N}$  but also the states at each discretization node  $X := (H_0, \dots, H_N)$ , where  $H_i = (R_i, r_i)$ . The total number of variables is thus  $12(N + 1) + 3N$ , so that the problem is indeed larger than in single shooting. The discretization of the objective is the same as in single shooting, except that now the gravitational potential energy keeps its linear structure since the centerline nodal positions are part of the degrees of freedom. The position, frame and strains are coupled via the introduction of constraints which are nothing but the discretized kinematics.

**Direct Multiple Shooting NLP.** The discretized Kirchhoff statics OCP for fixed-fixed boundary conditions in the absence of inequality constraints reads

$$\begin{aligned} & \min_{H_0, u_0, \dots, H_{N-1}, u_{N-1}, H_N} && \sum_{i=0}^{N-1} \frac{h}{2} u_i^T K u_i + h \rho g S e_3^T r_i \\ & \text{subject to} && H_0 = (R_I, r_I) \\ & && H_{i+1} = \Psi_h(H_i, u_i) \quad i = 0 \dots N - 1 \\ & && H_N = (R_L, r_L), \end{aligned} \tag{6.26}$$

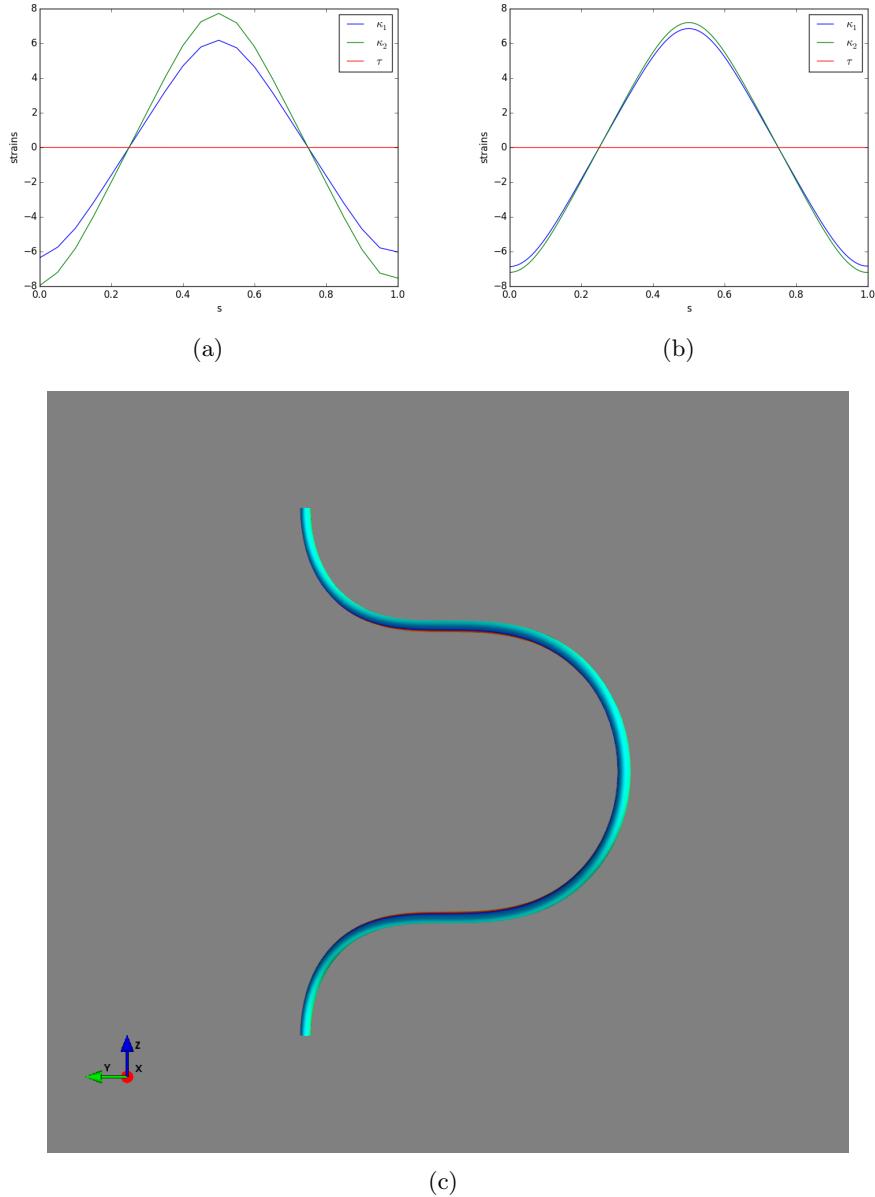


Figure 6.12: Solving a fixed-fixed statics problem with a piecewise linear strain based method and RKMK2.  $\gamma = 0.7$ , end to end distance = 0.5(a) 20 elements strains (b) 100 elements strains (c) 100 elements rod position.

where  $(R_I, r_I)$  and  $(R_L, r_L)$  are the given boundary conditions. The map  $\Psi_h$  is the same as in direct single shooting with piecewise constant strains, that is one step of a RKMK1 integrator as defined in equations (6.22) and (6.21). Notice that there are many more constraints than in single shooting, however each one depends only on a few of the variables :  $H_i, H_{i+1}, u_i$ . This sort of local dependence leads to sparse linear algebra problems at the level of the Newton iterations (see Figure 6.13). The block banded structure of the Hessian makes it much more economical to evaluate than for single shooting as the number of non zeros grows linearly with the number of elements instead of quadratically.

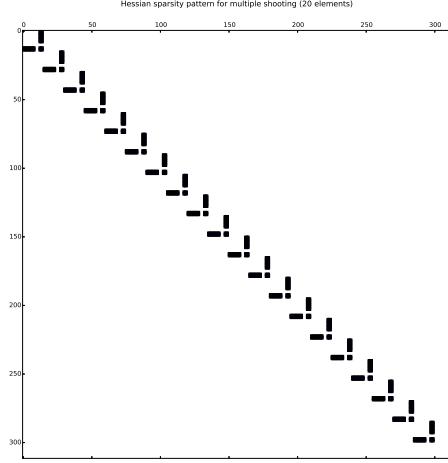


Figure 6.13: Direct Multiple Shooting Hessian sparsity pattern

**Mechanical interpretation.** The direct multiple shooting approach with piecewise constant controls and a Lie-Euler integrator also yields a piecewise helical rod geometry as for super helices. However it is not a purely strain based model but a mixed one, the variables are strains, positions and frames. Such mixed discretizations have already been considered in [64] for Cosserat rods and in [65] for planar Kirchhoff rods.

**Numerical application.** In Figure 6.14 we present the numerical solution obtained for the stable static equilibrium of a symmetric rod under periodic boundary conditions and in the absence of gravity. The obtained solution as seen in Figure 6.14 (d) has a circular centerline with a constant non zero twist. The twist is constant along the centerline as expected. The internal forces are plotted in Figure 6.14(b), they are computed as the Lagrange multipliers enforcing the discretized position constraints. The computed internal forces are seen to be constant, as they should be since according to the Kirchhoff balance equations (6.4) the internal force  $F$  satisfies  $F' = 0$  in the absence of gravity. In Figure 6.15 we solve the same problem but with gravity. The solution is no longer of a circular nature but gets deformed to a carabiner-looking curve. The computed internal forces  $F_1, F_2$  are constant and  $F_3$  is linear, in agreement with the Kirchhoff balance equations under gravity which read  $F' = \rho S g e_3$ .

**Invariants.** For a Kirchhoff rod in the absence of gravity the following quantities are conserved: the momentum  $m^T f$ , which is the analogue for the projection of angular momentum onto the axis of gravity for the heavy top, the Hamiltonian  $\frac{1}{2} \kappa^T K \kappa + f^T e_3$ , which is the analogue of the total energy of the heavy top. For a symmetric rod the quantity  $m^T e_3$  is also constant, meaning that the twist  $\tau$  is constant: it is the analogue of the rotation speed of a symmetric top around its axis of symmetry. In Figure 6.16 we plot the evolution of the momentum and of the Hamiltonian along arclength for different values of the number of elements. The typical behaviour of symplectic integrators is observed, where the computed value of the Hamiltonian oscillates less and less around the true constant value as the step size becomes smaller and

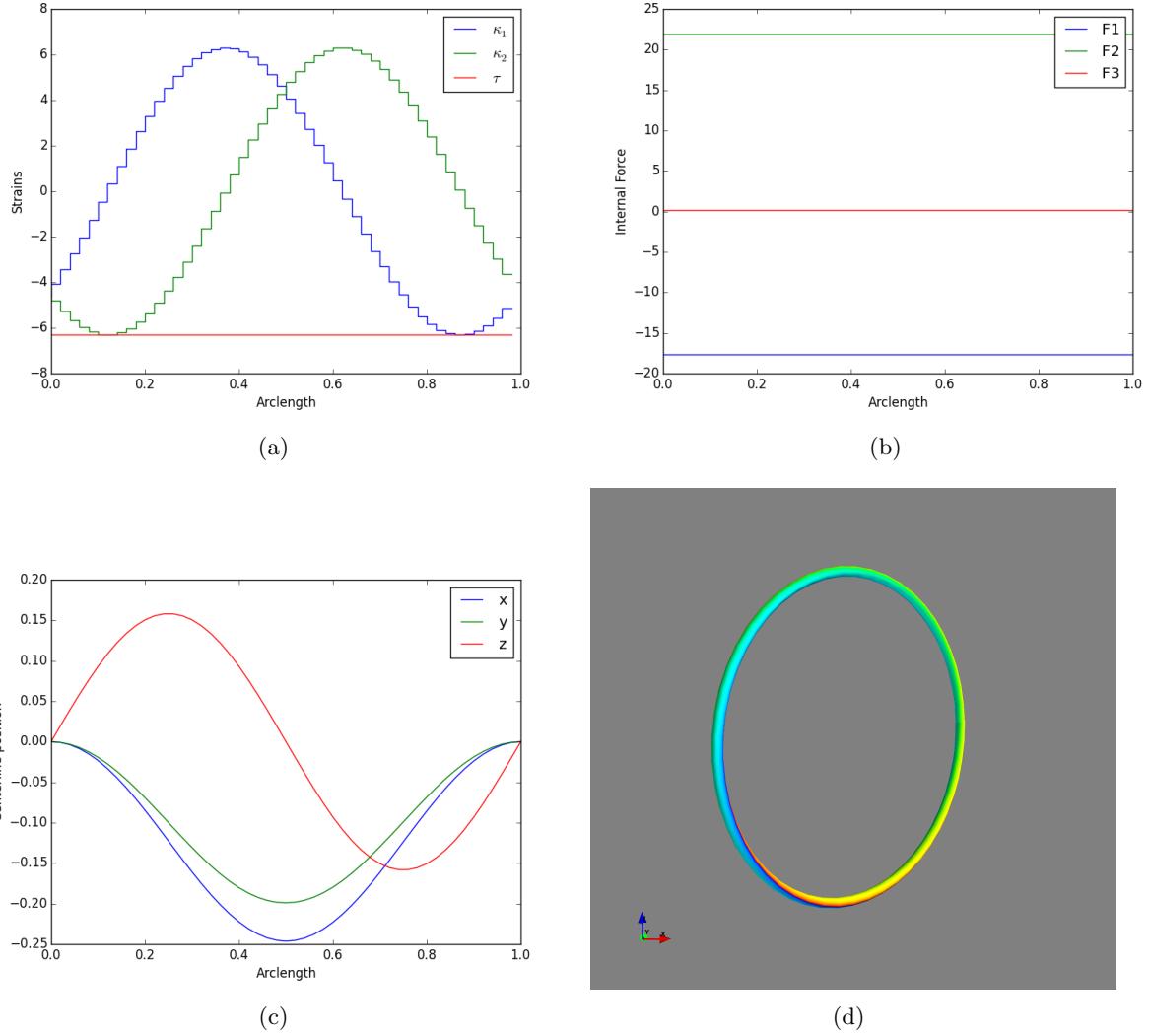


Figure 6.14: Static equilibrium configuration of a rod under periodic boundary conditions without gravity,  $\gamma = 0.7$ . 50 elements.(a) Strains. (b) Internal force. (c) Position of the centerline. (d) 3d view of the rod's equilibrium.

smaller. Symplecticity is one the reasons why even for a coarse discretization the qualitative behaviour of the discrete model is correct. The twist  $\tau$  is found to be constant as in the example of Figure 6.14.

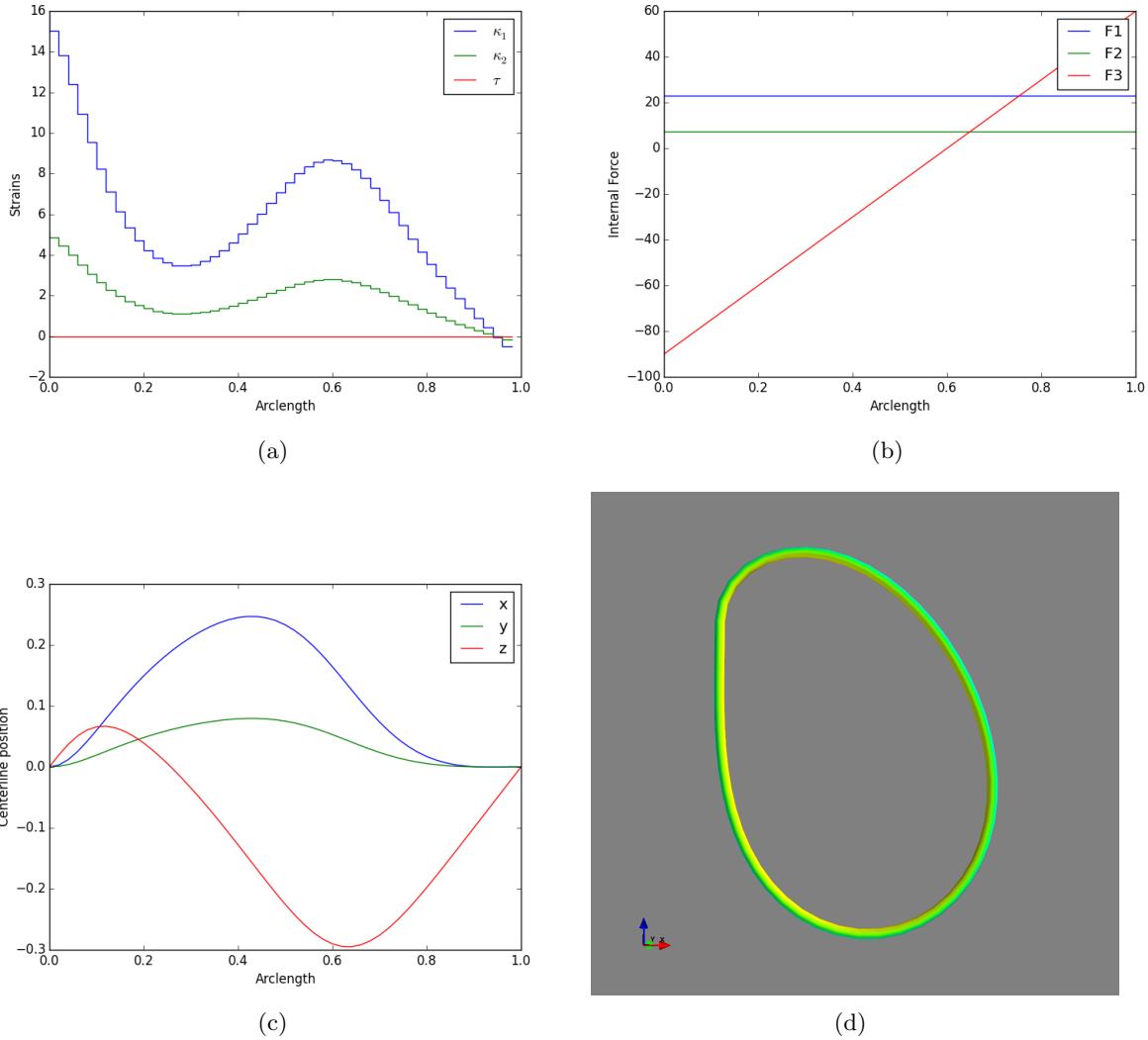


Figure 6.15: Static equilibrium configuration of a rod under periodic boundary conditions with strong gravity  $\gamma = 0.7$ ,  $\beta := \frac{\rho S g L^3}{EI} = 150$ . 50 elements.(a) Strains. (b) Internal force. (c) Position of the centerline. (d) 3d view of the rod's equilibrium.

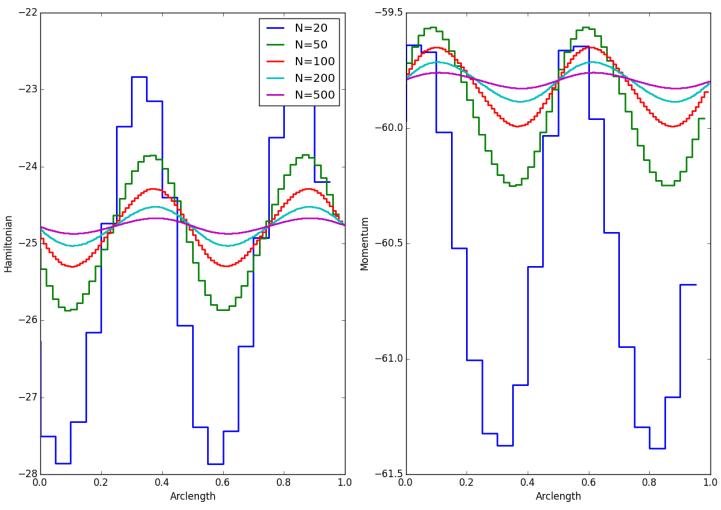


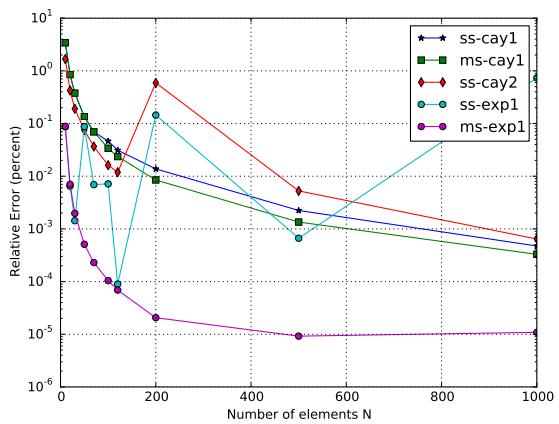
Figure 6.16: Computed Hamiltonian and momentum evolution of the rod along arclength for an increasing number of elements. The problem instance solved was a rescaled fixed-fixed problem with  $\gamma = 0.7$ , no gravity, an end-to-end distance of 0.8 and an end-to-end rotation of  $\pi/2$ .

### 6.3.3 Numerical evaluation

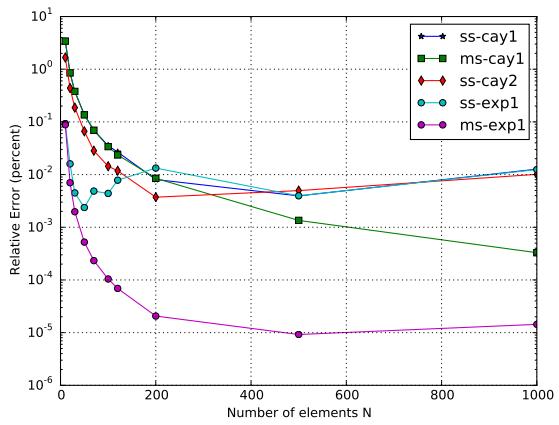
To give an idea of the accuracy and relative performance of the direct methods presented in the previous sections we solve with all methods the same problem: a fixed-fixed statics problem for a symmetric rod without gravity. We fix the end-to-end distance to 90% of the total length and use the same boundary conditions for the material frame at both ends of the rod. The obtained numerical solutions are all planar and can be compared to the solution of the Elastica via indirect methods exposed in Section 1 of this chapter. In Figure 6.17 we plot the relative error as a function of number of elements used. The methods tested are the following: direct single shooting with piecewise constant strains with both the Cayley (ss-cay1) and exponential (ss-exp1) maps, direct single shooting with piecewise affine strains with a Cayley RKM2 integrator (ss-cay2) and finally direct multiple shooting with piecewise constant strains and both the Cayley (ms-cay1) and exponential (ms-exp1) maps.

For almost all problems solved the error is beneath 1 percent, something that goes unnoticed by the bare eye. However the direct single shooting methods seemed to be yielding the unexpected results seen in Figures 6.17(a,b), where their error, although it remains small, doesn't decrease with the number of elements. The observed behavior appears to be influenced by two things: the use of the Hessian approximation instead of a full Hessian and the use of reduced boundary conditions instead of natural boundary conditions. We call natural boundary conditions the 12 conditions  $R_N = R^{target}$ ,  $r_N = r^{target}$ , in contrast to reduced boundary conditions where for the rotation part we only impose the three conditions  $\text{diag}(R_N^T R^{target} - I) = 0$ . The intuition behind this is, that natural boundary conditions are redundant since the orthonormality of the frame is already guaranteed by construction since we use an RKM method: this blocks already 6 degrees of freedom among the 9 of  $R_N$ . The boundary condition for the frame should be imposed via three constraints. Imposing 9 constraints is redundant. A fact which, however, seems to affect in practice only the direct single shooting implementation and not the direct multiple shooting.

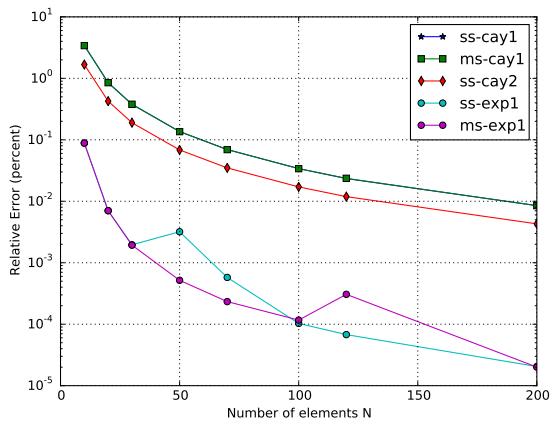
In Figure 6.17(c) we can see that for the piecewise constant strain discretization the single shooting and the multiple shooting versions are equally precise, as expected, since they are theoretically equivalent. The Cayley versions ss-cay1 and ms-cay1 are less precise than the exponential map versions ss-exp1 and ms-exp1, although they seem to be of the same order. The single shooting solver with piecewise linear strains is more precise than ss-cay1 although surprisingly less precise than ss-exp1, at least in this problem instance.



(a)



(b)



(c)

Figure 6.17: Convergence plots using (a) Natural boundary conditions and BFGS(b) Reduced boundary conditions and BFGS (c) Reduced boundary conditions and full hessian.

## 6.4 Quasi static experiments

In this section we present various numerical experiments of quasistatics of Kirchhoff rods carried out using our proposed direct methods.

### 6.4.1 Instability when varying $\bar{\kappa}_1$ .

In this experiment we reproduce numerically the dynamic instability that occurs to the statics equilibrium of a rod under gravity as its natural curvature is varied. We compute the static equilibrium configurations of a rod using direct single shooting with piecewise constant strains for a circularly symmetric rod of length 30 cm, Poisson's ratio 0.5, Young's modulus 1.3 MPa, density  $1200 \text{ kg/m}^3$  and radius 1.5 mm. We vary natural curvature  $\bar{\kappa}_1$  and let  $\bar{\kappa}_2 = \bar{\tau} = 0$ . The initial guess for the first problem is random. Then each computed equilibrium configuration is used as the initial guess for the next problem where the natural curvatures have changed. In Figure 6.18 we visualize the different stages of this process. At first (a), when  $\bar{\kappa}_1 = 0$ , the solution is planar and straight except at a boundary layer near the clamping end. As  $\bar{\kappa}_1$  augments (b), the rod begins to bend near the free end as well, releasing bending energy but accumulating gravitational energy, the solution is still planar. In (c) an instability occurs and the solution loses its planar symmetry, the vertical loop that was forming becomes too heavy and it becomes cheaper for the rod to use some twist, switching the loop to a horizontal configuration and thus releasing gravitational energy. In (d) the process continues and the rod assumes the configuration of a coiled spring under gravity.

In Figure 6.19 we plot the 'load displacement' diagram of the process by monitoring the evolution of the total twist  $\frac{1}{2\pi} \int_0^L \tau(s)ds$  as a function of the natural curvature  $\bar{\kappa}_1$ . The transition through the critical value of the load parameter  $\bar{\kappa}_1$  appears to be non smooth at our level of sampling, it occurs extremely fast.

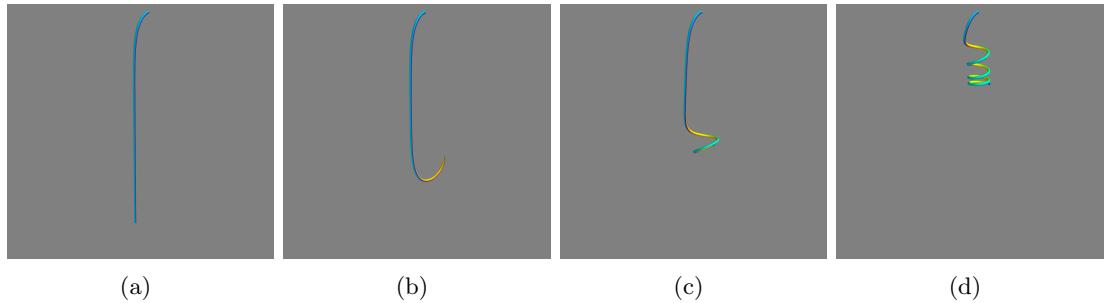


Figure 6.18: Static equilibrium configurations of a rod for varying natural curvature  $\bar{\kappa}_1$ , the rod is circularly symmetric and has length 30 cm, poissons ratio 0.5, Youngs modulus 1.3 MPa, density =  $1200 \text{ kg/m}^3$  and radius = 1.5 mm. (a)  $\bar{\kappa}_1 = 0m^{-1}$ . (b)  $\bar{\kappa}_1 = 50m^{-1}$  (c)  $\bar{\kappa}_1 = 60m^{-1}$  (d)  $\bar{\kappa}_1 = 90m^{-1}$

### 6.4.2 Circle Packing

In this numerical experiment we would like to find the stable static equilibria of an elastic ring packed inside a smaller ring. For this we consider the optimal control formulation of planar Kirchhoff statics and discretize it using direct multiple shooting with piecewise constant strains. The continuous problem reads

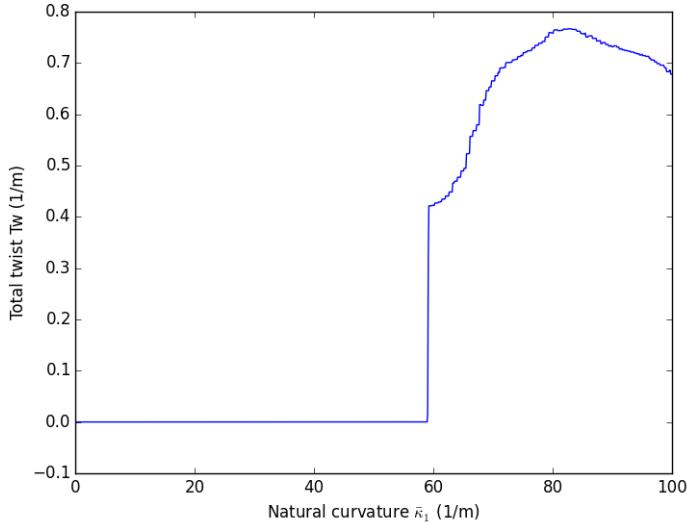


Figure 6.19: Evolution of total twist of the equilibrium configuration as a function of the natural curvature of the rod.

$$\begin{aligned}
& \min_{R, r, \kappa} \frac{EI}{2} \int_0^L \kappa(s)^2 ds \\
& \text{s. t. } R' = R \begin{pmatrix} 0 & -\kappa \\ \kappa & 0 \end{pmatrix} \\
& \quad r' = Re_1 \\
& \quad r(0) = r(L) \\
& \quad R(0) = R(L) \\
& \quad R(0)^T R(0) = I_{2 \times 2} \\
& \quad \|r\|^2 \leq c,
\end{aligned}$$

where  $c > 0$  is the radius of the packing circle. The boundary conditions are periodic and since we don't impose explicitly the initial frame  $R(0)$  it is necessary to impose that it be orthonormal. The kinematics will ensure that the rest of the frames are also orthonormal provided that the first one is. We plot in Figure 6.20 two obtained numerical solutions starting from random initial guesses. They are deformed versions of two classical solutions of the unconstrained periodic elastica: the circle and the figure eight. The curious reader may obtain experimentally the curve in Figure 6.20(a) by packing a sheet of paper with two ends glued together (so that it is cylindrical) inside a smaller cylinder (maybe a coffee mug?): the profile of the packed sheet of paper is a solution of our packed ring problem. For more on the topic of confined developpable sheets and its relation to the Elastica the reader may consult [105].

One advantage of multiple shooting here is that since we use the states  $R_i, r_i$  and strains  $\kappa_i$  as variables in the NLP, the inequality constraints  $\|r_i\|^2 < c$  remain convex quadratic. Those inequality constraints would have been non convex for a direct single shooting (strain-based) discretization.

Also note that in this mixed formulation where the strains are also variables we can also impose the topology of the solution by imposing a further linear constraint. Indeed, discretizing

the curve index theorem

$$\int_0^L \kappa(s) ds = 2\pi \text{Index}$$

leads to the linear constraint in  $\kappa$

$$h \sum_{i=0}^{N-1} \kappa_i = 2\pi \text{Index}.$$

The simulation used 100 elements, the total simulation time was 1.3s on a 2.6GHz Intel Core i7 processor parting from a random initial configuration using Ipopt. Our proposed method is quite robust to the choice of the initial configuration, as shown in Figure 6.21 where we display the iterations performed by Ipopt.

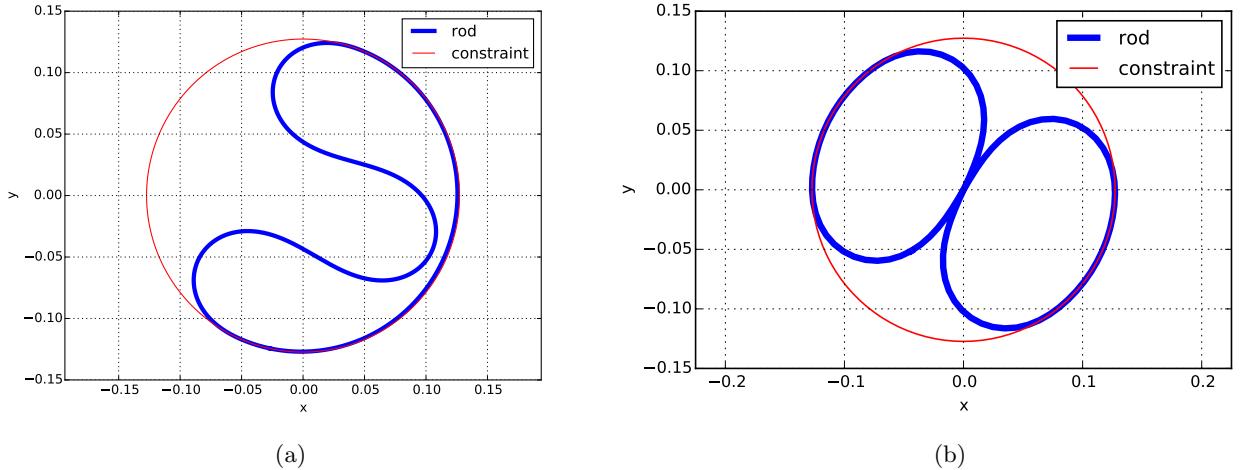


Figure 6.20: Static equilibria of planar Kirchhoff rods with periodic boundary conditions and contact constraints. (a) Index 1 solution. (b) Index 0 solution.

#### 6.4.3 Plectoneme Formation

Encouraged by the ease with which we solved statics problems for planar Kirchhoff rods subject to inequality constraints in the previous section, we set out to use our 3D statics solvers to simulate numerically the formation of plectonemes. Plectonemes are the self coiling formations which appear very frequently in long slender rods. In Figure 6.22 we show a numerical simulation of plectoneme formation. The experiment we perform is to keep fixed the end to end distance of the rod with both tangents aligned and iteratively rotate one the material frames at the end around its tangent, thus twisting the rod. In Figure 6.23 we plot the evolution of total twist as a function of the clamping angle. Twist accumulates and after about 3 full turns an instability occurs and the twist drops instantly. The configurations before and after this event are plotted in Figure 6.22 (a) and (b). We can see that the twist is released and the first contact is formed. Then as the clamping angle grows, twist accumulates further and after another instability it

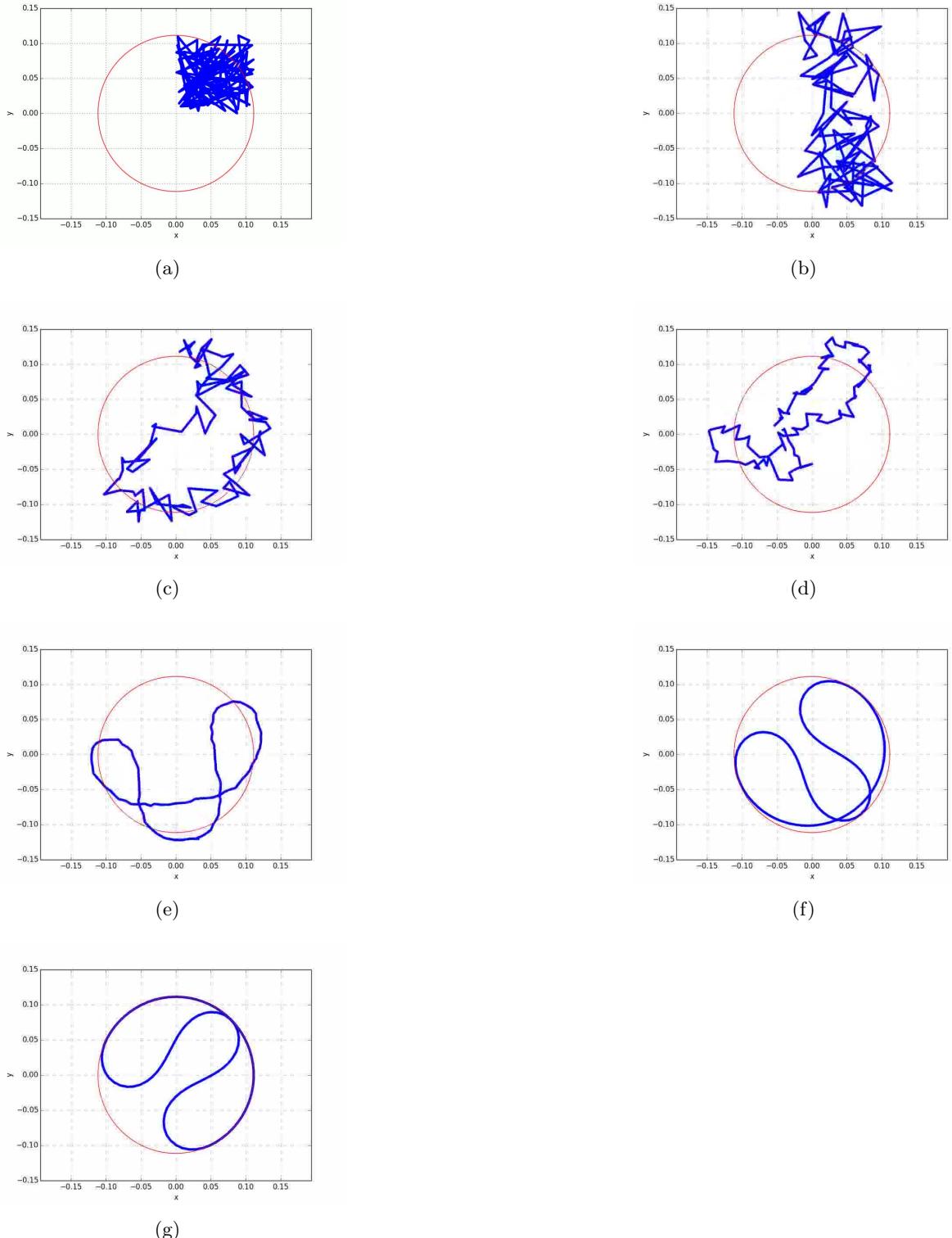


Figure 6.21: Ipopt Iterations. (a) Initial guess: random. (b) Iteration 20. (c) Iteration 30. (d) Iteration 60. (e) Iteration 90. (f) Iteration 100. (g) Optimal configuration : Iteration 200.

drops and a second contact is formed (Figures 6.22 (c) and (d)). In Figures 6.22 (e) and (f) the rod displays 3 and 4 loops respectively and the number of self contact points is larger.

The constraints involved in plectoneme formation are inequality constraints that forbid self collision. They exhibit two difficulties. The first difficulty is that they are non local. In the discrete setting this translates into the fact that if there are  $N$  elements then there are  $O(N^2)$  inequality constraints since we need to take into account the possible collision of one element with all others. The second difficulty is that the constraints are non convex. We require that the distance between any two points on different elements be larger than the diameter of the rod so that  $\|r(s) - r(t)\|^2 \geq (2\text{radius})^2$  for all  $s \in [s_i, s_{i+1}]$  and  $t \in [s_j, s_{j+1}]$  and  $i \neq j$ , where  $r$  denotes the centerline position.

In our first attempt at simulating plectoneme formation we used direct multiple shooting with Ipopt. After each problem solved, the solution was used as an input for the next problem where the clamping angle was augmented. What we observed was that no contact was formed at the moment of the instability (Figures 6.22 (a) to (b)). Instead, during the internal iterations of Ipopt, interpenetration of the rod occurred and the solver found another local minimum of the problem without contact and with less twist. Thus no plectonemes were formed.

To perform the simulation shown in Figure 6.22 our approach was the following. We use a direct single shooting discretization with piecewise constant strains, the degrees of freedom are the strains on each element collected in the vector  $q \in \mathbb{R}^{3N}$ . The spatially discretized energy of the system is  $U(q) := \frac{1}{2}q^T \mathbf{K}q$ , with  $\mathbf{K}$  diagonal and positive definite as shown in Section 6.3.1. The fixed-fixed boundary conditions are modelled as the equality constraints  $g_{E,\theta}(q) = 0$ , they depend on the clamping angle  $\theta$ . The non penetration constraints are modelled as the inequality constraints  $g_I(q) \geq 0$ . Assume we are given a solution  $q^{\theta-\Delta\theta}$  of the problem with clamping angle  $\theta - \Delta\theta$ . We find the next solution  $q^\theta$  by forward solving the first order complementarity system

$$\begin{cases} \dot{q} + \nabla U(q) &= \nabla g_{E,\theta}(q)\lambda + \nabla g_I(q)\mu \\ g_{E,\theta}(q) &= 0 \\ 0 \leq g_I(q) &\perp \mu \geq 0 \end{cases} \quad (6.27)$$

with  $q_0 := q^{\theta-\Delta\theta}$  as an initial condition until  $\|\dot{q}\| < \epsilon$ . If  $\|\dot{q}\| < \epsilon$  for  $\epsilon$  small enough then a solution of the first order optimality conditions of the problem minimizing  $U(q)$  subject to  $g_{E,\theta}(q) = 0$  and  $g_I(q) \geq 0$  has been found.

To numerically solve the complementarity system (6.27), at each time step  $t_k$  compute active set at  $q_k$  and solve for  $v, \lambda, \mu, q_{k+1}$  the Mixed Linear Complementarity Problem (MLCP)

$$\begin{cases} v &= \frac{q_{k+1}-q_k}{\Delta t} \\ v + \nabla U(q_k) &= \nabla g_{E,\theta}(q_k)\lambda + \nabla g_I(q_k)\mu \\ \nabla g_{E,\theta}(q_k)v &= -\gamma g_{E,\theta}(q_k) \\ 0 \leq \nabla g_I(q_k)v &\perp \mu \geq 0. \end{cases} \quad (6.28)$$

In practice we reformulate (6.28) as a Quadratic Program (QP) and solve it using qpOASES [106].

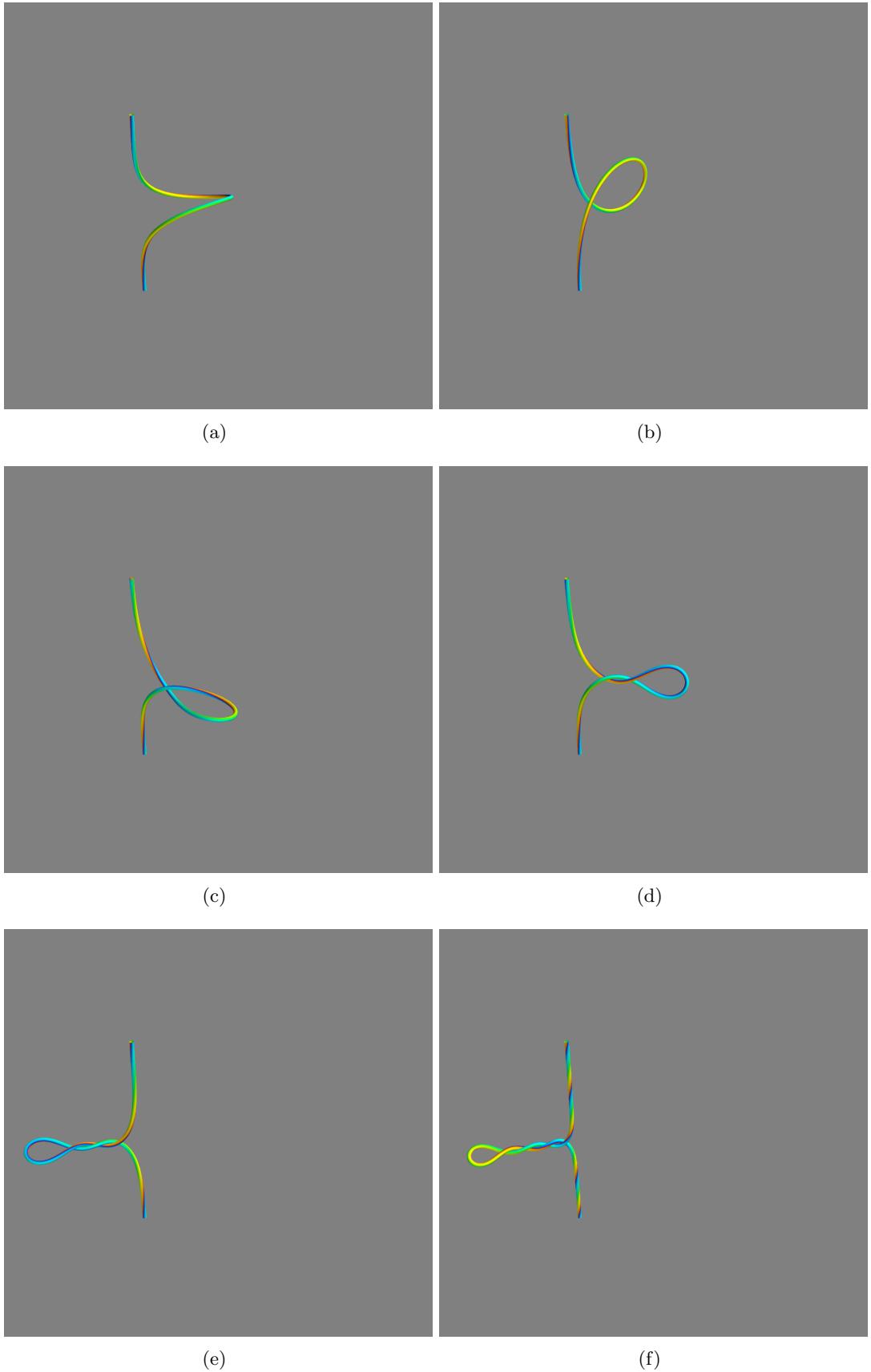


Figure 6.22: Visualization of plectoneme formation without gravity.

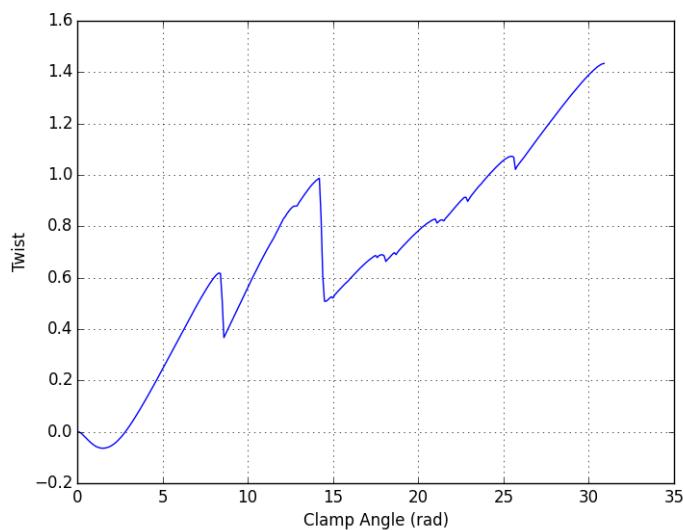


Figure 6.23: Total twist as a function of clamping angle during plectoneme formation without gravity.

# Conclusion

**Recapitulation.** We have seen in this part how the problems of computing the stable static equilibria of Kirchhoff and Cosserat rods can be formulated as Optimal Control Problems. The strains of the rod are interpreted as control variables and the frame and position as state variables. This interpretation has lead us to the exploration of numerical methods which hadn't been considered before in the context of rod statics. We propose to discretize the rod statics OCPs using direct single shooting and direct multiple shooting methods. Both methods rely on a discretization of the strain field, our implementations covered the case of piecewise-constant and piecewise-linear strain fields. The orientation and position of the rod are recovered from the strains by numerically solving the rod kinematics with a Runge-Kutta-Munthe-Kaas method on  $SE(3)$ . The discretized OCPs take the form of Non-linear programs (NLPs).

Our proposed methods are easy to implement. The objective function and constraints of the NLPs are easy to assemble using Casadi [68], a computer algebra system designed for Optimal Control. Casadi can compute the Jacobians and Hessians of the objective function and constraints thanks to automatic differentiation and can furthermore automatically detect the sparsity patterns in these matrices, making the implementation extremely easy. If needed, C code can be generated and compiled for the evaluation of all the terms involved in the NLP.

The direct single shooting discretizations lead to purely strain based methods, where the no extension and no shear constraints of the Kirchhoff rod are handled intrinsically. In the case of a piecewise constant strain field, our discretization coincides with that of [1]. Hence our construction generalizes the super-helix element of [1].

The direct multiple shooting discretizations lead to mixed methods, where the discrete strains, as well as the nodal frames and positions are degrees of freedom in the problem. It may at first seem unnecessary to introduce so many degrees of freedom, but while the problem size increases the non linearities are reduced. Furthermore, the Hessian matrix which needs to be inverted by the NLP solver is sparse and block banded in the direct multiple shooting formulation, whereas it is dense in the single shooting formulation. These mixed methods are advantageous when the number of elements is high. In problems where the contact of the rod with an external object or with itself needs to be taken into account, a mixed discretization yields simpler-to-handle constraints.

**Limitations.** If the forces acting on the rod are non conservative then the Optimal Control formulation is no longer possible as the balance equations would not have a variational structure. Our same spatial discretization procedures could nevertheless be used to tackle the weak form of the balance equations. We would however obtain equilibrium configurations which are not necessarily stable.

**Future work.** Some interesting theoretical questions could be pursued. For instance, in the presence of inequality constraints, the balance equations describe a non-smooth boundary value problem. Is there any known time stepping scheme for the solution of the non-smooth

BVP which would yield the same algebraic system of equations as the first order optimality conditions of our discretized OCPs? In other words, when do indirect methods commute with direct methods if inequality constraints are involved?

From a numerical point of view, a comparison between our strain based and mixed methods against absolute coordinate methods remains to be done. As well the application of our methods to some benchmark problems.

The application of the optimal control framework to compute the statics of developable ribbons [107] is part of ongoing work.

# **Part III**

## **Strain based dynamics of developable shells**

# Chapter 7

## Introduction

**From Kirchhoff rods to developable shells.** In the previous part we saw how to model and simulate inextensible and unshearable thin elastic rods. In particular we explored a class of strain based methods for the simulation of Kirchhoff rods. The no shear and no extension constraints of the rod can be handled intrinsically, by construction, using the strains of the rod as the main unknowns in the simulation instead of the positions and orientations of the cross sections. In the strain based approach we avoid the use of excessively stiff elastic forces to handle the inextensibility and no shear constraints. In this part our goal is to extend the strain based approach so as to simulate the dynamics of shells with an inextensible and unshearable mid-surface.

**Definition of a Shell.** A *shell*, as depicted in Figure 7, is a thin elastic solid with two dimensions of comparable magnitude (length and width), both much larger than the third (height or *thickness*). Given appropriate kinematic and physical assumptions, the behaviour of the shell may be entirely described by the behaviour of its mid surface. The mid surface of a shell can take on a non planar rest configuration. If the mid surface of the shell is planar in its rest configuration it is called a *plate*.

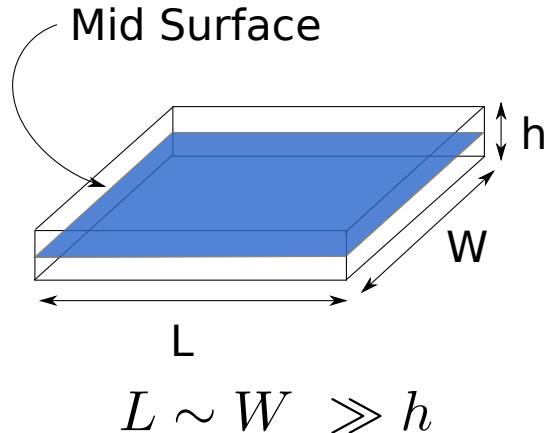


Figure 7.1: A shell and its mid surface.

**Motivation.** Many surface-like objects around us such as paper, leaves or boat sails, may easily bend but hardly stretch or shear. One is thus faced with the need for numerical models

able to handle these constraints properly not only in the case of rods but also in the case of shells.

**Developable shells.** Contrary to the case of rods, for shells in general it is very difficult to undergo pure bending deformations. The obstruction to pure bending deformations is related to the geometry of the mid surface. For instance closed convex surfaces are *rigid*, it is impossible to deform them in a smooth fashion without stretching or shearing<sup>1</sup>. Hence shells with a closed convex mid surface cannot undergo pure bending deformations. The implication for us is that we have to restrict our study to a class of shells for which pure bending deformations are possible. In the present part we thus restrict ourselves to the modeling of thin elastic shells with a *developable* mid-surface. A developable surface is one which always remain isometric to a planar configuration. We shall from now on refer to shells with a developable mid-surface as *developable shells*.

**Contributions.** Our goal is to design a strain based discrete model for simulating the motion of developable shells. The approach we follow is directly inspired by the super-helix model for thin elastic rods [1]. In this model, each rod element is characterized by uniform material curvatures and twist (i.e., uniform bending and twisting strains); a multibody rod dynamical system is then derived with curvatures and twist as degrees of freedom, leading to a perfectly inextensible dynamical rod model. Similarly, we build here an inextensible and unshearable shell patch by taking as degrees of freedom the *curvatures* (or *bending strains*) of its mid-surface, expressed in the local frame; equations of motion are then directly solved for the curvatures. Compared to the 1D (rod) case however, significant difficulties arise in the 2D (plate/shell) case, where compatibility conditions are to be treated carefully.

In the construction proposed in Section 9.2 the curvatures of the shell are taken to be spatially constant (the direct analog of the helical element in [1]). This leads to a shell patch that takes on circularly cylindrical configurations. The compatibility conditions are enforced as quadratic holonomic constraints. It is remarkable that in this case the expression for the position of the mid surface is an explicit function of the curvatures. This work was done in collaboration with Romain Casati who had already proposed the kinematic construction of the cylindrical patch in his PhD thesis [108]. The reinterpretation of his kinematic construction within the  $SE(3)$  formalism already employed in Part II is due to us as well as the formulation of the dynamics of the shell patch.

In Section 9.3 we propose an analysis of the compatibility equations for developable surfaces. It leads to the derivation of spatially affine shape functions for the curvatures which automatically satisfy the compatibility equations. The position of the mid surface is in this case no longer explicit but can be numerically obtained using a Runge-Kutta-Munthe-Kaas method to solve the surface reconstruction equations. The resulting shell patch takes on configurations which are also included in cylinders but of a clothoidal rather than a circular base.

To the best of our knowledge the strain based simulation of developable shells is a problem that had not been treated before. Our contributions represent a small step in this direction and we hope that it spurs further developments in the strain based approach to the numerical mechanics of shells.

**Overview.** In the following we shall first, in Chapter 8, present the strain based approach for Koiter shells (which can also stretch and shear) and review the literature on the subject. Then

---

<sup>1</sup>This fact is the nightmare of cartographers as it explains that it is impossible to make a perfect map of the earth, one without distortion of lengths, angles or areas.

in Chapter 9 we introduce a method for simulating the pure bending dynamics of a perfectly inextensible and unshearable patch of a developable thin elastic shell.

# Chapter 8

## The Strain based approach to the mechanics of shells

In this chapter we present the strain based approach for shells. We begin by giving some important definitions and concepts related to the geometry of surfaces, for which we rely heavily on the references [86] and [109]. We present Koiter's shell energy, which depends on the first and second fundamental forms of the mid-surface of the shell. We state the fundamental theorem of surface theory which is at the heart of the strain based approach for shells. Since we will deal in particular with shells or plates with a developable mid surface, we shall include some background on developable surfaces in this chapter. We will also quickly review the literature on the strain based approach for shells.

### 8.1 The first and second fundamental forms

**Parametrized surfaces.** A *parametrized surface* is a differentiable map  $r$  from an open domain  $D$  of  $\mathbb{R}^2$  into  $\mathbb{R}^3$ . A point of coordinates  $(s^1, s^2)$ <sup>1</sup> in the domain  $D$  gets mapped by  $r$  to a point  $r(s^1, s^2)$  in  $\mathbb{R}^3$ , as depicted in Figure 8.1. Let us denote  $\mathcal{S}$  the set  $r(D)$  which is called the trace of  $r$ . It is frequent to refer to the map  $r$  as the *parametrization* and the trace  $\mathcal{S}$  as the *surface*.

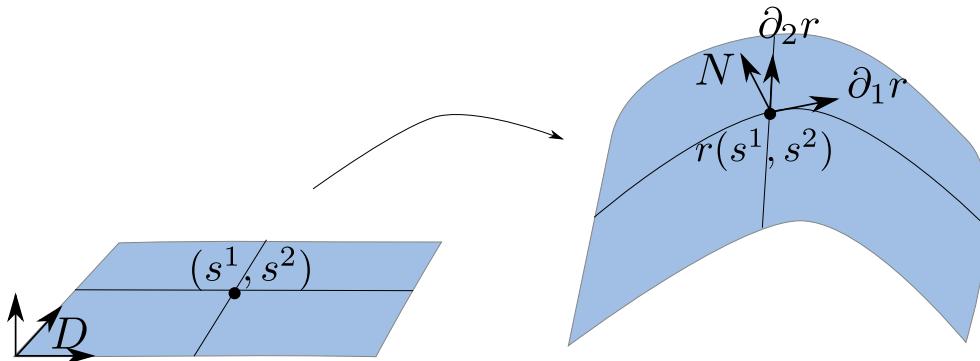


Figure 8.1: A regular parametrized surface.

We will denote  $d_1$  and  $d_2$  the tangent vectors defined as  $d_i := \partial_i r$ , where we have used

---

<sup>1</sup>We will be using classical tensor notation so  $s^i$  denotes the ' $i$ -th component of  $s$ ' and not ' $s$  to the power  $i$ '.

the shorthand notation  $\partial_i := \frac{\partial}{\partial s^i}$ <sup>2</sup>. If the tangent vectors are linearly independent for all  $(s^1, s^2) \in D$  then the parametrized surface  $r$  is said to be *regular* or an *immersion*. When the surface is regular at a point  $s := (s^1, s^2) \in D$  then the tangent vectors span a plane called the tangent plane to the surface  $\mathcal{S}$  at the point  $r(s)$ , which is denoted  $T_{r(s)}\mathcal{S}$ . When the surface is regular we can also define the normal vector  $N(s^1, s^2) := \frac{d_1 \times d_2}{\|d_1 \times d_2\|}(s^1, s^2)$ . We can collect the three columns  $(d_1, d_2, N)$  in a 3 by 3 invertible matrix  $F := (d_1, d_2, d_3) \in GL_3(\mathbb{R})$  called *the moving frame induced by the parametrization*. In general the frame  $F$  is not orthogonal. The parametrization stretches and shears the domain  $D$  in a smooth fashion to convert it into the surface  $\mathcal{S}$ . These changes in lengths and angles (or stretching and shearing) is what the **first fundamental form** of the surface is meant to capture.

**Definition 8.1.1.** First Fundamental Form. The first fundamental form  $a$  of a regular parametrized surface  $r : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$  is a 2 by 2 symmetric positive definite matrix field defined on  $D$  as

$$a_{ij} := \partial_i r \cdot \partial_j r \quad (8.1)$$

for  $(i, j)$  in  $\{1, 2\}$ . The first fundamental form is also called the 'metric' of the surface.

The diagonal components  $a_{ii}$  measure the amount of stretching along the  $s^i$  coordinate line. The off diagonal components  $a_{12} = a_{21} = d_1^T d_2$  measure the amount of shearing. We will see later on that the 'membrane' part of Koiter's shell energy depends on the first fundamental form. The 'flexural' or 'bending' part will depend on the second fundamental form, which we now define.

**Definition 8.1.2.** Second fundamental form. The second fundamental form  $b$  of a regular parametrized surface  $r : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$  is a 2 by 2 symmetric matrix field defined on  $D$  as

$$b_{ij} := -\partial_i r \cdot \partial_j N \quad (8.2)$$

for  $(i, j)$  in  $\{1, 2\}$ .

As one would expect, the curvature information of a surface is contained in the derivatives of the normal field. If the surface is flat the normal field is constant and the curvature is zero. The more the surface bends along one direction, the faster the normal field changes along that direction and the larger the magnitude of the curvature. The second fundamental form captures that information. The components  $b_{ij}$  are also called 'bending strains' in mechanics.

**Remark** We see in the definition of the second fundamental form the projection of the derivatives of the normal onto the tangent plane. There is no loss of information through that projection since the derivative of the normal field along any direction is tangent to the surface. Indeed, the normal field is of unit norm so we have that  $\partial_i(N \cdot N) = 0$  and hence  $\partial_i N \cdot N = 0$ , that is  $\partial_i N$  is tangent to the surface.

**Remark** We can also express  $b_{ij}$  as  $\partial_{ij} r \cdot N$ . Since the normal  $N$  is always orthogonal to the tangent vector  $\partial_i r$ , we have that  $\partial_j(\partial_i r \cdot N) = 0$  hence  $\partial_{ij} r \cdot N = -\partial_i r \cdot \partial_j N = b_{ij}$ .

---

<sup>2</sup>What we really mean is  $d_i(s^1, s^2) := \frac{\partial r}{\partial s^i}(s^1, s^2)$  for  $i \in \{1, 2\}$ , but as this is notationally heavy we will often use the shorthand notation for partial derivatives.

**Mixed components of the second fundamental form.** As we have seen, the derivatives of the normal map  $\partial_i N$  are tangent to the surface. Thus  $\partial_i N$  can be decomposed in the local basis  $(d_1, d_2)$  as a linear combination  $\partial_i N = -b_i^k d_k$ . We are using the Einstein summation convention, where if an upper and lower index is repeated then summation is implied, so  $b_i^k d_k = b_i^1 d_1 + b_i^2 d_2$ . The relationship between  $b_{ij}$  and  $b_i^k$  is easy to work out: on the one hand  $\partial_i N \cdot d_j = -b_{ij}$  by definition of the second fundamental form and on the other hand  $\partial_i N \cdot d_j = -b_i^k d_k \cdot d_j = -b_i^k a_{kj}$ . Hence  $b_i^k a_{kj} = b_{ij}$ . Denoting  $a^{ij}$  the components of the inverse of the metric  $a_{ij}$ , we can rewrite the former equality as

$$b_i^k = b_{ij} a^{jk}.$$

In general to raise the indices of a tensor we multiply by the inverse of the metric as  $v^i = a^{ij} v_j$ . To lower indices of a tensor we multiply by the metric as  $v_i = a_{ij} v^j$ .

The matrix with components  $b_i^k$  is called the *second fundamental form in mixed components* or the *Shape Operator*. In matrix form, if we denote  $S$  the matrix of the shape operator then  $S = ba^{-1}$ . The shape operator is a self-adjoint linear map with respect to the metric  $a$ .

**Definition 8.1.3.** Principal, Mean and Gaussian curvatures.

We call *principal curvatures* of the surface the eigenvalues of the shape operator  $S$  and denote them  $\kappa_{\min}$  and  $\kappa_{\max}$ .

The *Gaussian curvature*, denoted  $K_G$ , is the geometric mean of the principal curvatures:

$$K_G := \kappa_{\min} \kappa_{\max} = \det(b_i^j) = \frac{\det(b_{ij})}{\det(a_{ij})}.$$

The *mean curvature*  $K_M$  is the arithmetic mean of the principal curvatures:

$$K_M := \frac{\kappa_{\min} + \kappa_{\max}}{2} = \frac{\text{tr}(b_i^j)}{2}.$$

We now give the definition of a surface isometry.

**Definition 8.1.4.** Isometry. A surface isometry is an application between surfaces that preserves lengths and angles, that is, which doesn't change the metric. If  $\Phi : (\bar{\mathcal{S}}, \bar{a}) \rightarrow (\mathcal{S}, a)$  is an application between a surface  $\bar{\mathcal{S}}$  with metric  $\bar{a}$  and a surface  $\mathcal{S}$  with a metric  $a$  then  $\Phi$  is an isometry if and only if for all points  $p$  in  $\bar{\mathcal{S}}$  and for all vectors  $X, Y$  in the tangent plane  $T_p \bar{\mathcal{S}}$

$$a_{\Phi(p)}(D\Phi(p)X, D\Phi(p)Y) = \bar{a}_p(X, Y),$$

where we have denoted  $D\Phi(p)$  the differential of  $\Phi$  at  $p$ .

Our definition is more accurately that of a local isometry. A global isometry is a local isometry that is also a diffeomorphism, in particular a global isometry preserves the topology of the surface. In that sense the plane and the cylinder are only locally isometric and not globally isometric. In the future if we write isometric we implicitly mean *locally isometric*.

**Developable Surfaces.** In the next chapter we will focus our attention to shells with a developable mid surface, we are now in a position to give a definition of developability. A surface is said to be *developable* if its Gaussian curvature is identically zero ( $K_G = 0$  for all points on the surface). Clearly  $K_G = 0$  holds if and only if at least one of the principal curvatures vanishes. If both principal curvatures vanish then the surface is planar. If only one principal curvature vanishes then the surface is *parabolic*, it is straight in the direction of the zero principal curvature and curved in the orthogonal direction. The sign of mean curvature determines whether the surface is curved upwards, that is in the same direction as the normal,

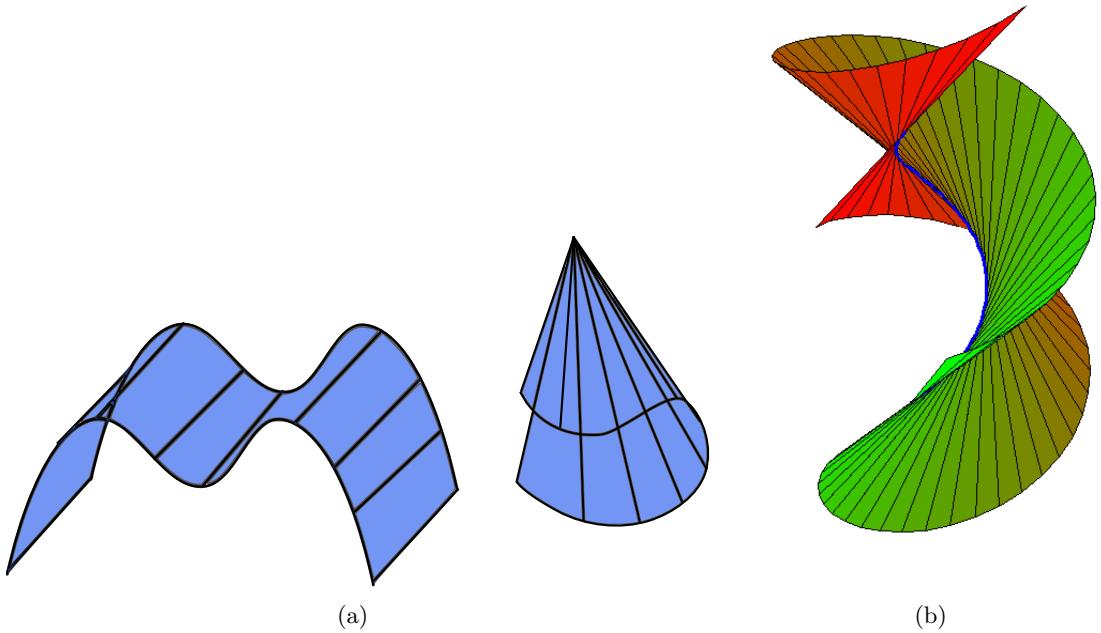


Figure 8.2: (a) Generalized cylinder and cone. (b) Tangent developable of a helix, image from [Wikipedia en.wikipedia.org/wiki/Tangent\\_developable](https://en.wikipedia.org/wiki/Tangent_developable)

or downward, opposite to the normal. Developable surfaces are *ruled*, the curves along which the principal curvature vanishes (called *rulings*) are straight lines, they entirely cover the surface and the normal to the surface is constant along those lines. Developable surfaces are comprised of patches of planes, generalized cylinders (not necessarily with a circular base), generalized cones and 'tangent developables' (see Figure 8.2). Tangent developables are surfaces constructed by taking a curve and extruding it along its tangent field. The term 'developable' comes from the fact that they can be developed onto a flat configuration without any stretching or shearing; they are locally isometric to the plane.

## 8.2 The fundamental theorem of surface theory

We have seen that given a regular parametrized surface we can deduce its first and second fundamental forms. In terms of the mechanics of shells this means that we can deduce stretching, shearing and bending strains from the displacements. In order to have a strain based approach to the mechanics of shells, where the displacements are deduced from the strains we must address the following question: can one recover the surface given its first and second fundamental forms? The answer to this question is positive, provided that certain compatibility conditions are met. The latter statement constitutes what is known in differential geometry as the *fundamental theorem of surface theory*.

**Reconstructing a surface given its fundamental forms.** It is clear that if the tangent vectors  $d_i$  are given on the domain  $D$  then the parametrized surface  $r$  could be recovered by integration of the equations

$$\partial_i r = d_i \quad (8.3)$$

on  $D$ . The situation is very similar to the Kirchhoff kinematics of Chapter 5 but now there are 2 spatial variables  $(s^1, s^2)$  instead of just the arclength of the rod, which means that we are dealing with PDE's instead of ODE's, and also the directors are not (yet) assumed orthonormal. As we know from Chapter 5, to recover the directors from the curvatures we need to look at the kinematics of the directors. By decomposing the derivatives  $\partial_i d_j$  and  $\partial_i N$  in the basis  $(d_1, d_2, N)$  we obtain the frame reconstruction equations

$$\begin{cases} \partial_i d_j = \Gamma_{ij}^k d_k + b_{ij} N \\ \partial_i N = -\bar{b}_i^k d_k. \end{cases} \quad (8.4)$$

The coefficients  $\Gamma_{ij}^k$  are called the Christoffel symbols. The Christoffel symbols are defined as  $\Gamma_{ijk} := \partial_{ij} r \cdot \partial_k r$  and  $\Gamma_{ij}^k := a^{kl} \Gamma_{jkl}$ .

Now, behold the following miracle: the Christoffel symbols  $\Gamma_{ijk}$  can be shown to depend only on the metric  $a_{ij}$  and not on the parametrization. Indeed, by the formula of Koszul we have that

$$\Gamma_{ijk} = \frac{1}{2} (\partial_i a_{jk} + \partial_j a_{ik} - \partial_k a_{ij}). \quad (8.5)$$

The conclusion is that, if the fundamental forms  $a$  and  $b$  are given, then the frame reconstruction equation (8.4) is a linear first order PDE for the frame  $(d_1, d_2, N)$  with coefficients depending on  $a$  and  $b$ . To reconstruct the surface from its fundamental forms in practice, we first integrate (8.4) to obtain the directors  $d_1, d_2$  and then we integrate (8.3) to recover the parametrized surface  $r$ .

**Integrability conditions.** The tensor fields  $a$  and  $b$  cannot be arbitrary. First,  $a$  has to be symmetric positive definite and  $b$  has to be symmetric. Second,  $a$  and  $b$  necessarily satisfy the **Gauss-Codazzi-Mainardi-Peterson** equations (GCMP)

$$\begin{cases} R_{ikj}^p = b_{ij} b_k^p - b_{ik} b_j^p \\ \partial_k b_{ij} - \partial_j b_{ik} = \Gamma_{ik}^l b_{lj} - \Gamma_{ij}^l b_{lk} \end{cases} \quad (8.6a)$$

$$R_{ikj}^p = \partial_k \Gamma_{ji}^p - \partial_j \Gamma_{ik}^p + \Gamma_{kl}^p \Gamma_{ji}^l - \Gamma_{jl}^p \Gamma_{ki}^l \quad (8.6b)$$

where  $R_{ikj}^p := \partial_k \Gamma_{ji}^p - \partial_j \Gamma_{ik}^p + \Gamma_{kl}^p \Gamma_{ji}^l - \Gamma_{jl}^p \Gamma_{ki}^l$  is the **Riemann curvature tensor**.

**Remark** The GCMP equations (also known as *compatibility conditions*) were first derived as necessary conditions verified by the fundamental forms. They result from requiring the commutation of cross derivatives  $\partial_{ijk} r = \partial_{ikj} r$  and using the reconstruction equations (8.3) and (8.4). The tangential projection  $\partial_{ijk} r \cdot d_p = \partial_{ikj} r \cdot d_p$  results in the *Gauss equation* (8.6a). The normal projection  $\partial_{ijk} r \cdot N = \partial_{ikj} r \cdot N$  results in the Codazzi-Mainardi-Peterson equations (8.6b). It was proven independently by Peterson [110] and Bonnet that the GCMP equations are also the sufficient conditions for the symmetric positive definite matrix field  $a_{ij}$  and symmetric matrix field  $b_{ij}$  to determine the surface locally.

The symmetries of  $a$  and  $b$  imply that there are only 3 independent equations in (8.6), which are

$$\begin{cases} R_{1212} &= b_{11} b_{22} - (b_{12})^2 \\ \partial_2 b_{11} - \partial_1 b_{12} &= \Gamma_{12}^1 b_{11} + (\Gamma_{12}^2 - \Gamma_{11}^1) b_{12} - \Gamma_{11}^2 b_{22} \\ \partial_2 b_{12} - \partial_1 b_{22} &= \Gamma_{22}^1 b_{11} + (\Gamma_{22}^2 - \Gamma_{12}^1) b_{12} - \Gamma_{12}^2 b_{22}. \end{cases} \quad (8.7)$$

**Remark** The celebrated *Theorema Egregium* of Gauss follows from the first equation in (8.7). The Riemann curvature tensor depends only on the metric and  $\det(b_{ij}) = K_G \det(a_{ij})$ . Hence Gaussian curvature also depends only on the metric:  $K_G = \frac{R_{1212}}{\det(a_{ij})}$ . In particular, Gaussian

curvature is invariant with respect to surface isometries. This is of interest to us since it means that the Gaussian curvature of a shell's mid surface will be preserved by pure bending deformations.

The following theorem collects Theorem 2.8-1 and Theorem 2.9-2 of [109]. We denote  $\mathbb{S}_2$  the set of real symmetric 2 by 2 matrices and  $\mathbb{S}_2^{++}$  the set of real symmetric positive definite 2 by 2 matrices.

**Theorem 8.2.1.** *Fundamental Theorem of the theory of surfaces. Let  $D$  be a connected and simply connected open subset of  $\mathbb{R}^2$ . Let  $a \in \mathcal{C}^2(D; \mathbb{S}_2^{++})$  be a symmetric positive definite matrix field on  $D$  and  $b \in \mathcal{C}^2(D; \mathbb{S}_2)$  be a symmetric matrix field on  $D$ . If  $a$  and  $b$  satisfy the Gauss-Codazzi-Mainardi-Peterson equations (8.7) then there exists a  $\mathcal{C}^3$  immersion  $r : D \rightarrow \mathbb{R}^3$  having  $a$  and  $b$  as fundamental forms, that is  $a_{ij} = \partial_i r \cdot \partial_j r$  and  $b_{ij} = \partial_{ij} r \cdot \frac{\partial_1 r \times \partial_2 r}{\|\partial_1 r \times \partial_2 r\|}$ . Furthermore the immersion  $r$  is unique up to a rigid body motion.*

The compatibility conditions (8.7) of the surface are just integrability conditions for the linear first order PDE's (8.4). In practice to integrate the frame reconstruction PDE (8.4) we first integrate in one direction (say  $s^1$ ) for a given initial datum and then integrate in the other direction. The compatibility conditions (8.7) guarantee that it is indifferent to integrate in direction  $s^1$  first and then in direction  $s^2$  or vice versa (provided that the domain is simply connected). In the textbook of Palais [111] a very clear proof of Theorem (8.2.1) is given by invoking Frobenius' theorem on the integrability of first order linear PDEs.

The regularity conditions in Theorem 8.2.1 can be significantly relaxed, although that goes beyond the scope of this thesis. The reader may consult [109] and references therein. Another important property proved in [109] is that the immersion is a continuous function of its fundamental forms, that is if we vary continuously the fundamental forms then the reconstructed surface varies continuously as well.

**Remark** The reconstructed surface is guaranteed to be an immersion but not an embedding, in particular the reconstructed surface can self-intersect. If self-penetration of the surface is to be avoided, it has to be enforced by additional constraints.

**Remark** For a developable surface with an identity first fundamental form, the Christoffel symbols vanish and the compatibility equations take the relatively simple form

$$\begin{cases} b_{11}b_{22} - (b_{12})^2 &= 0 \\ \partial_2 b_{11} - \partial_1 b_{12} &= 0 \\ \partial_2 b_{12} - \partial_1 b_{22} &= 0. \end{cases} \quad (8.8)$$

### 8.3 Koiter Shells

In this section we present Koiter's shell energy which is valid for small strains but large displacements and large rotations. The reduction of the full three dimensional elastic energy of the shell to an energy depending only on the mid surface is based on two hypothesis:

- The Kirchhoff-Love kinematic hypothesis which states that material segments initially orthogonal to the mid surface remain orthogonal to the mid surface and their length remains unchanged throughout the deformation.
- The no transverse shear hypothesis which states that shearing in the direction orthogonal to the mid-surface can be neglected.

Similar versions of Koiter's shell theory can be obtained by relaxing these assumptions, see for instance [81, Chapter 17]. If the material segments are restrained to remain straight but neither orthogonal to the mid-surface nor of constant length, one obtains the so-called *one director theory of Cosserat shells*. If the material segments remain straight, of constant length but not normal to the mid-surface, then one obtains the so-called *special theory of Cosserat shells*.

Let  $\bar{a}_{ij}$  and  $\bar{b}_{ij}$  be respectively the first and second fundamental forms of the mid surface of the shell in its natural configuration. The elasticity tensor  $\bar{a}^{ijkl}$  is defined as

$$\bar{a}^{ijkl} = \frac{2E\nu}{1-\nu^2} \bar{a}^{ij}\bar{a}^{kl} + \frac{E}{1+\nu} (\bar{a}^{ik}\bar{a}^{jl} + \bar{a}^{il}\bar{a}^{jk}),$$

where  $E$  and  $\nu$  are Young's modulus and Poisson's ratio of the elastic material. Let  $a_{ij}$  and  $b_{ij}$  denote the first and second fundamental forms of the shell's mid-surface in the actual (deformed) configuration. The stored elastic energy density  $w_K$  of a Koiter shell of thickness  $h$  reads

$$w_K = \frac{h}{4} \bar{a}^{ijkl} (a_{ij} - \bar{a}_{ij})(a_{kl} - \bar{a}_{kl}) + \frac{h^3}{48} \bar{a}^{ijkl} (b_{ij} - \bar{b}_{ij})(b_{kl} - \bar{b}_{kl}). \quad (8.9)$$

**Remark** The internal elastic energy (8.9) of a Koiter shell is a quadratic function of the first and second fundamental forms. The internal elastic energy is the sum of two terms, a *membrane* part and a *bending* part. The membrane part  $\frac{h}{4} \bar{a}^{ijkl} (a_{ij} - \bar{a}_{ij})(a_{kl} - \bar{a}_{kl})$  is seen to penalize the in plane deformations of the mid surface, it penalizes in plane shearing and stretching. The bending part  $\frac{h^3}{48} \bar{a}^{ijkl} (b_{ij} - \bar{b}_{ij})(b_{kl} - \bar{b}_{kl})$  penalizes the out of plane or bending deformations of the mid surface. It is important to note that the membrane term is linear in the thickness parameter  $h$  while the bending term is cubic in the thickness. The thinner the shell the harder it is to stretch and shear and the easier it is to bend. In the limit of zero thickness the shell undergoes pure bending deformations, provided it can do so.

**Remark** Pure bending deformations are those that leave the first fundamental form of the shell's mid surface unchanged. In geometric terms pure bending deformations are surface isometries. The question of whether it is possible to deform a surface isometrically is a non trivial one in general. For an interesting discussion on surface isometries and their relevance to shell theories the reader can consult [77, Chapter 11]. To find pure bending deformation of a surface one would need to find a family of second fundamental forms  $b_t$  depending continuously on a parameter  $t$  which comply with the compatibility conditions (8.7) for a fixed metric  $a$ . For a fixed metric the compatibility conditions (8.7) are a set of linear first order PDEs (the Codazzi-Mainardi-Peterson equations) but coupled with a non linear quadratic constraint (the Gauss equation). Even the linearization of (8.7) is non trivial since it can be recast into a second order linear PDE of mixed type, which is parabolic where the Gaussian curvature vanishes, hyperbolic where the Gaussian curvature is negavite and elliptic where the Gaussian curvature is positive [77]. One class of surfaces for which pure bending deformations exist, is the class of developable surfaces. This is one of the reasons why we restrict our attention to the pure bending deformations of shells with a developable mid surface. The other reason is that developable surfaces often arise in applications.

## 8.4 Strain based methods for shells

Just as for the rod theories of Chapter 5, it is possible to tackle the problem of shells using strain variables instead of position and orientation variables as primary unknowns. The appeal of the strain based method for shells is the same as for rods, the internal elastic energy (8.9) is

quadratic in the strains just as for Kirchhoff or Cosserat rods. The strain based method for rods is relatively well understood and has been tried on numerically for several rod models, using different basis functions and different reconstruction procedures. For shells the literature on strain based methods is much more scarce. The strain based formulation, also called *intrinsic* formulation, dates back to the forties and was originally proposed by Chien [112]. Since then the theory has been refined in [113], [114], [115]. In [116] an intrinsic formulation for the dynamics of composite plates was proposed which does not involve the first and second fundamental form directly but rather instantaneous displacement and rotation vectors as in a Cosserat description. In these references focus is put on theoretical aspects and the subject of numerically solving the obtained intrinsic balance equations is not dealt with. Furthermore it is unclear how to handle forces or moments that involve the position or orientation of the plate/shell structure. The proposed solution strategy in [115] is to first solve the equations for the strains and in a post processing step to recover the displacement and rotation fields. However this decoupling is rarely possible in scenarios of interest. We are unaware of any numerical method in the literature employing only strains as primary unknowns for the simulation of shells.

It should also be noted that the term 'intrinsic formulation' is often used to refer to a formulation for which the balance equations involve only strains/stresses and don't involve the position and orientation variables explicitly. In this sense the formulations of Simo and Fox [117] are intrinsic, although their main unknowns are the position and a director, not the strains.

# Chapter 9

## Strain based dynamics of developable shells

In this chapter we first present the kinematics of developable shells in the same  $SE(3)$  formalism used in Part II and developed in [118]. Background material on  $SE(3)$  is provided in Appendix C.1. We then proceed to the kinematic and dynamic construction of a shell patch with a spatially constant second fundamental form. In Section 9.3 we propose an analysis of the compatibility equations for developable surfaces and derive spatially affine shape functions for the curvatures which automatically satisfy the compatibility equations. A shell patch with a spatially affine second fundamental form is constructed. We present the main implementation features and some simulation results.

### 9.1 Strain based kinematics of developable shells

Let us consider a thin elastic shell with a developable mid-surface. We normally parametrize the mid-surface by a function  $r(s^1, s^2) \in \mathbb{R}^3$  of two spatial parameters  $(s^1, s^2)$  living on a planar domain  $D \subset \mathbb{R}^2$ . Our goal here is to explain how the surface parametrization can be obtained as a function of the bending strains of the shell.

**Enforcing developability.** As we are interested in simulating the pure bending dynamics of developable shells, we need to enforce the developability of the mid-surface throughout its deformation. The way in which we choose to do this is to require the first fundamental form of the surface to be always equal to the identity matrix. This implies that  $r$  is a local isometry between the planar configuration  $D$  (with metric  $\bar{a} = I$ ) and the actual mid-surface. The tangent vectors  $d_1 := \partial_1 r$  and  $d_2 := \partial_2 r$  are then orthonormal throughout the motion of the surface. The moving frame  $R(s^1, s^2) := (d_1, d_2, N)$ , defined by the two tangent vectors and the normal  $N := d_1 \times d_2$ , is thus constrained to remain a rotation matrix  $R(s^1, s^2) \in SO(3)$ .

**Surface and frame reconstruction.** This puts us in a situation very similar to the one encountered for Kirchhoff rods. If the frame field  $R$  is known on the domain then the position reconstruction equation reads

$$\frac{\partial r}{\partial s^i}(s) = R(s)e_i \quad \forall s \in D, \tag{9.1}$$

where  $i \in \{1, 2\}$ ,  $e_1 = (1, 0, 0)^T$  and  $e_2 = (0, 1, 0)^T$ . This is a 2D version of the Euler-Bernoulli constraint of Kirchhoff rods. In mechanical terms, the in-plane strains are prescribed. Among

the strains only the out of plane or "bending strains" remain unknown, geometrically they correspond to the entries of the second fundamental form of  $r$  and they appear naturally when one takes the derivatives of the frame with respect to  $(s^1, s^2)$ . Since  $R(s)$  belongs to  $SO(3)$  there exist two fields  $\Omega_1(s) \in \mathbb{R}^3$  and  $\Omega_2(s) \in \mathbb{R}^3$  such that

$$\frac{\partial R}{\partial s^i}(s) = R(s)\hat{\Omega}_i(s) \quad \forall s \in D, \quad (9.2)$$

where (we recall) the hat operator takes 3D vectors into skew symmetric matrices, that is

$$\hat{\cdot}: x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \rightarrow \hat{x} = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix} \in \mathfrak{so}(3).$$

The entries of  $\Omega_i$  are precisely the bending strains and the entries of the second fundamental form. Although these seem to be 6 parameters, only 3 of them are independent as we will shortly see. Again, as in the case of Kirchhoff rods, the  $SE(3)$  structure of the kinematics becomes apparent by collecting the frame  $R$  and the position  $r$  into the matrix

$$H(s) := \begin{bmatrix} R(s) & r(s) \\ 0 & 1 \end{bmatrix} \in SE(3).$$

We can restate the position reconstruction equation (9.1) and the frame reconstruction equation (9.2) as the  $SE(3)$  reconstruction equation

$$\frac{\partial H}{\partial s^i}(s) = H(s)\tilde{f}_i(s), \quad (9.3)$$

where  $\tilde{f}_i$  is the matrix  $\tilde{f}_i := \begin{bmatrix} \hat{\Omega}_i(s) & e_i \\ 0 & 0 \end{bmatrix} \in \mathfrak{se}(3)$ .

**Compatibility conditions.** The bending strain fields  $\Omega_i$  can't be arbitrary, they satisfy compatibility conditions which are nothing else than the requirement that second cross derivatives should commute,

$$\begin{aligned} \partial_1 \partial_2 H = \partial_2 \partial_1 H &\Leftrightarrow \partial_1(H\tilde{f}_2) = \partial_2(H\tilde{f}_1) \\ &\Leftrightarrow H\tilde{f}_1\tilde{f}_2 + H\partial_1\tilde{f}_2 = H\tilde{f}_2\tilde{f}_1 + H\partial_2\tilde{f}_1 \\ &\Leftrightarrow H\partial_2\tilde{f}_1 - H\partial_1\tilde{f}_2 = H\tilde{f}_1\tilde{f}_2 - H\tilde{f}_2\tilde{f}_1 \\ &\Leftrightarrow \partial_2\tilde{f}_1 - \partial_1\tilde{f}_2 = \tilde{f}_1\tilde{f}_2 - \tilde{f}_2\tilde{f}_1. \end{aligned}$$

In terms of the  $\Omega_i$  these conditions read

$$\begin{bmatrix} \partial_2\hat{\Omega}_1 - \partial_1\hat{\Omega}_2 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \hat{\Omega}_1\hat{\Omega}_2 - \hat{\Omega}_2\hat{\Omega}_1 & \hat{\Omega}_1e_2 - \hat{\Omega}_2e_1 \\ 0 & 0 \end{bmatrix},$$

which we can rewrite using cross products as

$$\begin{cases} \Omega_1 \times e_2 &= \Omega_2 \times e_1 \\ \partial_2\Omega_1 - \partial_1\Omega_2 &= \Omega_1 \times \Omega_2. \end{cases} \quad (9.4)$$

One has thus two sets of conditions, the first reads  $\Omega_1 \times e_2 = \Omega_2 \times e_1$ , yielding three equations which reduce the number of unknown bending strains from six down to three. Explicitly they tell us that  $\Omega_1^3 = \Omega_2^3 = 0$  and that  $\Omega_2^2 = -\Omega_1^1$ , so that one has

$$\Omega_1 = \begin{bmatrix} \Omega_1^1 \\ \Omega_1^2 \\ 0 \end{bmatrix} \quad \text{and} \quad \Omega_2 = \begin{bmatrix} \Omega_2^1 \\ \Omega_2^2 = -\Omega_1^1 \\ 0 \end{bmatrix}.$$

The fact that  $\Omega_1^3 = \Omega_2^3 = 0$  means that there is no drilling degree of freedom for developable shells. The three remaining bending strains are precisely the three components of the second fundamental form. The relationship between the instantaneous rotation vectors and the second fundamental form can be easily seen. The second fundamental form has components defined by  $b_{ij} = -\partial_i N \cdot \partial_j r$ . Plugging in the kinematics  $\partial_j r = R e_j$  and  $\partial_i N = R \hat{\Omega}_i e_3$  one concludes that

$$\begin{aligned} b_{11} &= -\Omega_1^2 \\ b_{12} &= \Omega_1^1 \\ b_{22} &= \Omega_2^1. \end{aligned}$$

Hence the instantaneous rotation vectors can be expressed in terms of the entries of the second fundamental form as follows,

$$\Omega_1 = \begin{bmatrix} b_{12} \\ -b_{11} \\ 0 \end{bmatrix} \quad \text{and} \quad \Omega_2 = \begin{bmatrix} b_{22} \\ -b_{12} \\ 0 \end{bmatrix}. \quad (9.5)$$

The second set of equations in the compatibility equations (9.4) is  $\partial_2 \Omega_1 - \partial_1 \Omega_2 = \Omega_1 \times \Omega_2$ . In terms of the second fundamental form, the latter equations read

$$(GCMP) \left\{ \begin{array}{lcl} \partial_2 b_{12} - \partial_1 b_{22} & = & 0 \\ \partial_2 b_{11} - \partial_1 b_{12} & = & 0 \\ b_{11} b_{22} - (b_{12})^2 & = & 0, \end{array} \right. \quad (9.6)$$

thus they are precisely the Gauss-Codazzi-Mainardi-Peterson equations for a surface with a metric tensor equal to the identity (encountered already in Equation (8.8)). Hence it follows by the Fundamental theorem of surface theory that the reconstruction equations (9.3) have a unique solution given an initial datum  $H_0 \in SE(3)$  if the domain  $D$  is connected and simply connected, if the bending strain vectors  $\Omega_i$  are defined as in Equation (9.5) and if their components satisfy the GCMP equations (9.6).

## 9.2 Dynamics of a developable shell patch with spatially constant curvatures

We wish to simulate the pure bending dynamics of a developable shell patch. We proceed just like in the construction of the super-helix element for the dynamics of Kirchhoff rods. Instead of formulating the problem as a non linear PDE involving functions of time and of the two spatial variables  $s^1, s^2$ , we use a method of lines. By using a method of lines, the problem of the pure bending dynamics of a developable shell patch can be reduced to one of multibody dynamics. We shall consider the entries of the second fundamental form to be spatially constant but time varying and take them to be the degrees of freedom of the shell patch. We collect the curvatures (or bending strains) in a vector  $q := [b_{11}, b_{22}, b_{12}]^T \in \mathbb{R}^3$ . These will be our generalized coordinates in the Lagrangian dynamics of the shell patch. However, these degrees of freedom aren't independent since they have to obey the GCMP equations (9.6): this is a major discrepancy with the rod case, where no such conditions arise. The kinetic and potential energies of the shell are to be expressed as functions of  $q$  and  $\dot{q}$ .

### 9.2.1 Kinematics of a developable shell patch with spatially constant curvatures

Let us consider that the  $\Omega_i$  are constant in space and given on the whole domain  $D$  and that they verify the compatibility conditions. Furthermore let us consider that  $D$  is star-shaped, centered

at the origin, so that any point  $s$  in the domain can be reached by a straight line segment emanating from the origin. The segment can be parametrized as the curve  $c(t) = (ts^1, ts^2)$  and one can consider the value of  $H$  along this segment, let us call it  $H_c(t) := H(c(t))$ . Then one can write a simple differential equation for  $H_c$ :

$$\begin{aligned}\frac{dH_c}{dt}(t) &= \partial_1 H(c(t))s^1 + \partial_2 H(c(t))s^2 \\ &= H_c(t)\tilde{f}_1(c(t))s^1 + H_c(t)\tilde{f}_2(c(t))s^2 \\ &= H_c(t)\tilde{A}(s),\end{aligned}$$

where  $\tilde{A}(s) = \tilde{f}_1(c(t))s^1 + \tilde{f}_2(c(t))s^2 = \tilde{f}_1s^1 + \tilde{f}_2s^2$ . Note that since  $\tilde{f}_i$  is constant on  $D$  then  $\tilde{A}(s)$  is independent of the parameter  $t$ . The ODE  $H'_c = H_c\tilde{A}(s)$  is an autonomous linear differential equation for  $H_c$ , its solution is given by

$$H_c(t) = H_c(0) \exp(t\tilde{A}(s)).$$

Evaluating at  $t = 1$  we recover the solution to (9.3) with initial condition  $H(0) = H_0 \in SE(3)$ , which reads

$$H(s^1, s^2) = H_0 \exp(s^1\tilde{f}_1 + s^2\tilde{f}_2). \quad (9.7)$$

As we have seen, the exponential of a matrix in  $\mathfrak{se}(3)$  has a closed form expression which comes from Rodrigues' formula. Hence the expression (9.7) for the surface and the frame is explicit.

**Explicit expression of the position** Let us collect the curvatures in a vector  $q := [b_{11}, b_{22}, b_{12}]^T$ . These will be our generalized coordinates in the Lagrangian dynamics of the shell patch.

The position of a material particle with coordinates  $(s^1, s^2)$  in the planar configuration is given by

$$r(q, s^1, s^2) = r_0 + R_0(I + \frac{1 - \cos(\theta)}{\theta^2}\hat{\Omega} + \frac{\theta - \sin(\theta)}{\theta^3}\hat{\Omega}^2)(s^1e_1 + s^2e_2), \quad (9.8)$$

where

$$\Omega = \begin{bmatrix} s^1q_2 + s^2q_1 \\ -s^1q_0 - s^2q_2 \\ 0 \end{bmatrix},$$

and  $\theta = \|\Omega\|_2$ . The reconstructed surface takes on the shape of a cylindrical patch as seen in Figure 9.1. Evidently if all entries of the second fundamental form are constant, then the mean curvature of the mid-surface is also constant and the lines of curvature are the straight rulings and circular arcs of constant curvature.

**Remark** Choosing spatially constant  $\Omega_i$  not only has the advantage of leading to an explicit reconstruction for the frame and position, it also simplifies the GCMP equations. The second fundamental form of the mid surface is then constant on  $D$  so the Codazzi-Mainardi-Peterson equations are automatically satisfied, that is

$$\begin{aligned}\partial_2 b_{12} - \partial_1 b_{22} &= 0 \\ \partial_2 b_{11} - \partial_1 b_{12} &= 0.\end{aligned}$$

Only the Gauss equation

$$b_{11}b_{22} - (b_{12})^2 = 0$$

remains to be enforced. It will be considered as a holonomic constraint

$$q_0q_1 - (q_2)^2 = 0$$

on the generalized coordinates  $q$ .

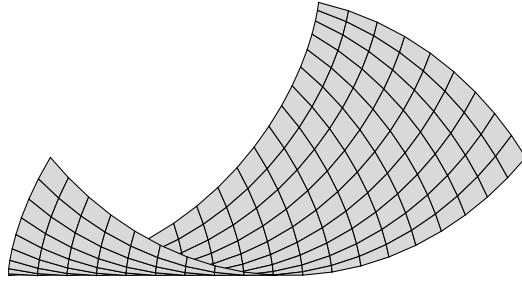


Figure 9.1: Geometry of our shell patch for  $q_0 = q_1 = q_2 = 1$ .

**Remark** The  $\mathfrak{se}(3)$  elements  $\tilde{f}_i$  containing the in-plane stretching/shearing strains and the bending strains are also spatially constant. In terms of the  $\tilde{f}_i$  the compatibility equations read

$$\partial_2 \tilde{f}_1 - \partial_1 \tilde{f}_2 = \tilde{f}_1 \tilde{f}_2 - \tilde{f}_2 \tilde{f}_1$$

which simplifies to

$$[\tilde{f}_1, \tilde{f}_2] = 0$$

for spatially constant  $\tilde{f}_i$ . Meaning that the matrices  $\tilde{f}_1$  and  $\tilde{f}_2$  commute. Under this condition we have that the exponential of the sum is the product of the exponentials and thus  $\exp(s^1 \tilde{f}_1 + s^2 \tilde{f}_2) = \exp(s^1 \tilde{f}_1) \exp(s^2 \tilde{f}_2) = \exp(s^2 \tilde{f}_2) \exp(s^1 \tilde{f}_1)$ . If the compatibility conditions did not hold then our solution for the frame and surface  $H(s) = H_0 \exp(s^1 \tilde{f}_1 + s^2 \tilde{f}_2)$  (obtained by radial integration) would have been different to that obtained by integrating first with respect to  $s^1$  and then with respect to  $s^2$  or vice versa.

**Manual simplification of the position expression using the eigen-basis of the second fundamental form** In the case of formula (9.7) one has  $\Omega = s^1 \Omega_1 + s^2 \Omega_2 = \begin{bmatrix} s^1 b_{12} + s^2 b_{22} \\ -s^1 b_{11} - s^2 b_{12} \\ 0 \end{bmatrix}$

and  $v = s^1 e_1 + s^2 e_2$ . Using the Gauss equation ( $b_{11} b_{22} = b_{12}^2$ ) one finds that the expression for  $\theta$  is given by

$$\theta = \|\Omega\|_2 = \sqrt{(b_{11} + b_{22}) s^T b s}.$$

This expression can be further simplified by introducing an eigen decomposition of the matrix of the second fundamental form  $b$ . Indeed,  $b$  is symmetric and has a zero eigen-value, if the other eigenvalue is non zero then the two eigen-directions are orthogonal and one can introduce a matrix  $P$  such that  $b = P^T \begin{bmatrix} 0 & 0 \\ 0 & b_{11} + b_{22} \end{bmatrix} P$ . The expression for  $\theta$  then reads  $\theta = |(b_{11} + b_{22})(Ps)_2|$ , the problem is that the matrix  $P$  changes depending on which entries of  $b$  are zero.

**Property** (Expression for  $\theta$ ). *Let  $\theta = \|s^1 \Omega_1 + s^2 \Omega_2\|$  where  $\Omega_i$  are the instantaneous rotation vectors defined in (9.2). Then the expression for  $\theta$  is given by*

- If  $b_{11} = 0$  and  $b_{22} = 0$  then  $b = 0$  and  $\theta = 0$ .

- If  $b_{11} \neq 0$  then  $P = \frac{1}{\sqrt{b_{11}^2 + b_{12}^2}} \begin{bmatrix} -b_{12} & b_{11} \\ b_{11} & b_{12} \end{bmatrix}$  and

$$\theta = |(b_{11} + b_{22}) \frac{b_{11}s^1 + b_{12}s^2}{\sqrt{b_{11}^2 + b_{12}^2}}| \quad (9.9)$$

- If  $b_{22} \neq 0$  then  $P = \frac{1}{\sqrt{b_{22}^2 + b_{12}^2}} \begin{bmatrix} -b_{22} & b_{12} \\ b_{12} & b_{22} \end{bmatrix}$  and

$$\theta = |(b_{11} + b_{22}) \frac{b_{12}s^1 + b_{22}s^2}{\sqrt{b_{22}^2 + b_{12}^2}}| \quad (9.10)$$

Furthermore, since in the expressions for  $\exp(\hat{\Omega})$  and  $B$  the function  $\theta$  appears only as an argument in the even functions  $\frac{\sin(\theta)}{\theta}$ ,  $\frac{1-\cos(\theta)}{\theta^2}$  and  $\frac{\theta-\sin(\theta)}{\theta^3}$ , one can omit the absolute values in the expression of  $\theta$ . Finally, with all this information at hand, one can arrive at the following closed form expression for the position of the surface as a function of its curvatures.

**Property** (Expression for  $r$ ).

- If  $b_{11} = 0$  and  $b_{22} = 0$  then  $b = 0$  and

$$r = r_0 + R_0 \begin{bmatrix} s^1 \\ s^2 \\ 0 \end{bmatrix} \quad (9.11)$$

- If  $b_{11} \neq 0$ :

$$r = r_0 + R_0 \left( \begin{bmatrix} s^1 \\ s^2 \\ 0 \end{bmatrix} + \frac{1-\cos\theta}{b_{11}+b_{22}} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + \frac{\theta-\sin\theta}{(b_{11}+b_{22})^2} \sqrt{b_{11}^2 + b_{12}^2} \begin{bmatrix} -1 \\ -b_{12}/b_{11} \\ 0 \end{bmatrix} \right) \quad (9.12)$$

- If  $b_{22} \neq 0$ :

$$r = r_0 + R_0 \left( \begin{bmatrix} s^1 \\ s^2 \\ 0 \end{bmatrix} + \frac{1-\cos\theta}{b_{11}+b_{22}} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + \frac{\theta-\sin\theta}{(b_{11}+b_{22})^2} \sqrt{b_{22}^2 + b_{12}^2} \begin{bmatrix} -b_{12}/b_{22} \\ -1 \\ 0 \end{bmatrix} \right) \quad (9.13)$$

As a test, one can verify that for  $b = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$  the expression (9.12) coincides with the explicit solution of (9.3), that is  $r(s) = \begin{bmatrix} \sin s^1 \\ s^2 \\ 1 - \cos s^1 \end{bmatrix}$ , which is a cylinder of radius 1 with circular base in the xz plane.

### 9.2.2 Dynamics of the developable shell patch with spatially constant curvatures

**Kinetic energy of the shell patch** A material particle which is located at coordinates  $(s^1, s^2)$  in the planar configuration will have coordinates  $r(q, s^1, s^2)$  in space. The velocity of such a particle can thus be computed as

$$\dot{r} = \frac{\partial r}{\partial q}(q, s^1, s^2)\dot{q}.$$

The kinetic energy due to each particle in the shell can be expressed as  $\frac{1}{2}dmr^2$ , where the elementary mass  $dm$  is deduced from the volumic mass density  $\rho$ , the thickness  $h$  and the elementary surface area  $ds^1 ds^2$  as  $dm = \rho h ds^1 ds^2$ .

The total kinetic energy of the shell is obtained by summing the kinetic energies of each elementary particle, and is given by

$$T(q, \dot{q}) = \iint_D \frac{1}{2} \rho h \dot{q}^T \frac{\partial r}{\partial q}(q, s^1, s^2)^T \frac{\partial r}{\partial q}(q, s^1, s^2) \dot{q} ds^1 ds^2.$$

Since the velocity  $\dot{q}$  and the density  $\rho$  do not depend on the coordinates  $(s^1, s^2)$ , they can be taken out of the integral and the kinetic energy thus reads

$$T(q, \dot{q}) = \frac{1}{2} \dot{q}^T M(q) \dot{q},$$

where the mass matrix  $M(q)$  is defined as

$$M(q) = \rho h \iint_D \frac{\partial r}{\partial q}(q, s^1, s^2)^T \frac{\partial r}{\partial q}(q, s^1, s^2) ds^1 ds^2.$$

**Gravitational potential energy** Each material particle contributes  $\rho g h r_3(q, s^1, s^2) ds^1 ds^2$  to the potential energy due to weight. The total gravitational potential energy thus reads

$$U_g(q) = \rho g h \iint_D r_3(q, s^1, s^2) ds^1 ds^2.$$

**Elastic potential energy** The elastic energy is entirely due to bending, the first fundamental form of the mid surface remains constant throughout its evolution. Hence the membrane part in Koiter's shell energy (8.9) vanishes. The total bending energy of the shell can be simplified to the form

$$U_b(q) = \frac{1}{2} (q - \bar{q})^T K (q - \bar{q}),$$

where  $\bar{q} := (\bar{b}_{11}, \bar{b}_{22}, \bar{b}_{12})^T$  is the vector of curvatures of the shell at rest (the so-called *natural* curvatures), and  $K$  the stiffness matrix defined as  $K = \text{area}(D) \frac{Eh^3}{12(1-\nu^2)} \begin{bmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & 2(1-\nu) \end{bmatrix}$ .

**Remark** When the natural curvatures vanish, the bending energy  $U_b(q)$  is just proportional to the mean curvature squared, that is, Willmore's energy functional. As we can see in that case  $U_b$  is proportional to  $(b_{11}^2 + b_{22}^2 + 2\nu b_{11} b_{22} + 2(1-\nu)b_{12}^2)$ . Using the Gauss equation the latter expression becomes simply  $(b_{11} + b_{22})^2$ , which is proportional to the square of the mean curvature of the mid-surface.

**Equations of motion.** To compute the pure bending dynamics of our shell patch, we build a constrained Lagrangian mechanical system with generalized coordinates  $q$  and Lagrangian of the form

$$L(q, \dot{q}) = \frac{1}{2} \dot{q}^T M(q) \dot{q} - U_g(q) - U_b(q),$$

subject to the holonomic constraint

$$g(q) := q_0 q_1 - q_2^2 = 0.$$

The equations of motion reduced to index 1 read

$$\begin{aligned} M(q) \ddot{q} &= F(q, \dot{q}) + G^T(q) \lambda \\ G(q) \ddot{q} &= -\dot{G} \dot{q}, \end{aligned} \tag{9.14}$$

where the Lagrange multiplier  $\lambda$  is just one dimensional here, and  $G(q)$  is the jacobian matrix of the constraints  $G(q) := \frac{\partial g}{\partial q}(q)$ . The vector  $F(q, \dot{q})$  is defined as

$$F(q, \dot{q}) = \nabla_q T(q, \dot{q}) - \dot{M} \dot{q} - \nabla_q U(q).$$

### 9.2.3 Numerical solution of the equations of motion

Since the system is small and we are not interested in its long term dynamics we choose to solve numerically the DAE (9.14) by reducing it to an ODE. We explicitly compute the Lagrange multiplier as

$$\lambda(q, \dot{q}) = -\frac{G(q)M^{-1}(q)F(q, \dot{q}) - \dot{G}\dot{q}}{G(q)M^{-1}(q)G^T(q)},$$

and plug it back into the force balance equation. This procedure leaves us with the ODE

$$\begin{aligned} \dot{q} &= v \\ \dot{v} &= M^{-1}(q) (F(q, v) + G^T(q)\lambda(q, v)), \end{aligned} \tag{9.15}$$

which we then solve simply with an *RK4* scheme. A problem of drift from the constraints arises, though, when using this scheme. Indeed we enforce the constraints in the acceleration level and all information about the constraints in the position and velocity level is lost. Although the formulations are analytically equivalent, numerically one witnesses how small errors in the acceleration level constraints lead to large ones in the position level constraints. To remedy this problem we add a stabilization term in the equations of motion which does not change the physical dynamics of the system and helps keep the curvatures on the constraint set. This is known as Baumgarte stabilization. In figure 9.2 we plot the evolution of the curvatures and the constraint violations. With the stabilization off one sees a quadratic growth of the position level constraint violation, in about 2 seconds it goes from 0 to  $10^{-2}$ . With the stabilization on, one sees that although the constraint is not kept exactly at 0, the error is controlled and it oscillates around the value  $4 \cdot 10^{-5}$ .

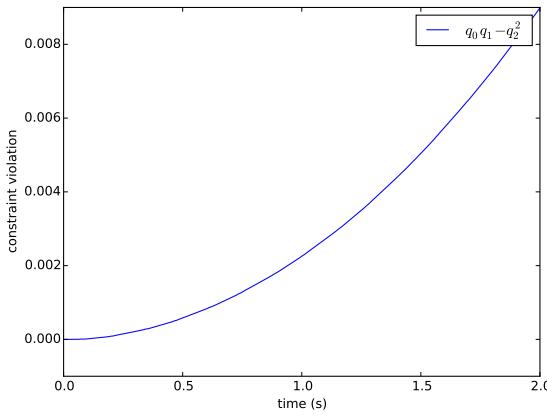
**Performance:** The numerical simulation of the shell dynamics runs in real time. Using a time step of  $dt = 0.001$ , 33 steps are taken to produce one frame of an animation, which is done in under 5ms. The computational cost of 10,000 time steps is of about 1s on a 2.6 GHz processor. Frames of an animation are depicted in Figure 9.3.

### 9.2.4 Implementation

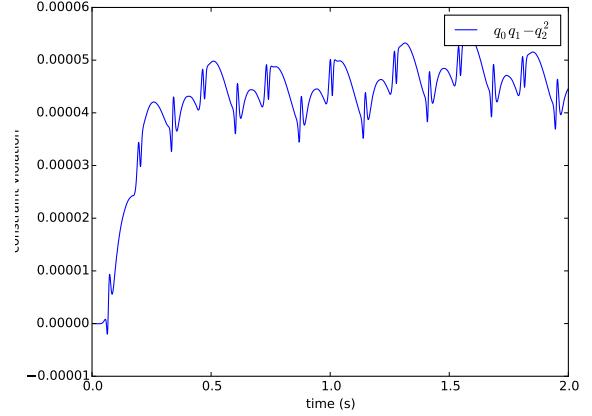
**Symbolic Computation using Sympy** To solve numerically the equations of motion one needs a code that evaluates all terms in the dynamics (9.14). We have a closed form expression for the position as a function of the curvatures  $q$  and of the curvilinear coordinates  $s = (s^1, s^2)$ :  $r = r(q, s)$ . Since this expression is relatively complex we, at first, used Sympy and PyDy [119] to generate symbolic expressions for  $M(q)$ ,  $\frac{\partial M(q)}{\partial q^i}$  and  $F_g(q)$ . Then from the symbolic expressions C code is generated using the codegen module of the PyDy package (around 8000 lines of C code are generated). By using Cython one can then call the compiled C code directly from Python which makes tasks such as implementing different numerical schemes, testing and plotting much easier. Generating the symbolic expressions in Sympy wasn't trivial, the main difficulty was getting Sympy to compute double integrals of complex expressions. We noted that all the expressions which we needed to integrate were of polynomial form in four variables,

$$\sum_{p_1, p_2, p_3, p_4} f_p(q)(s^1)^{p_1}(s^2)^{p_2}s_3^{p_3}s_4^{p_4},$$

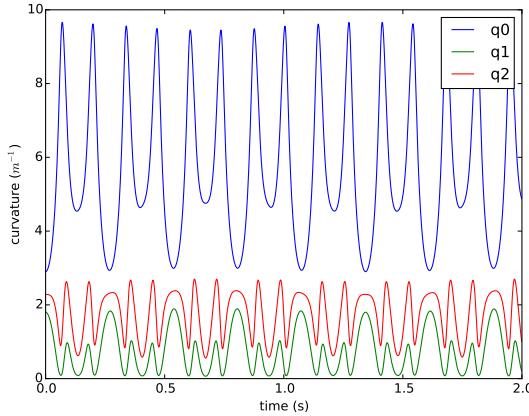
where  $(s_3, s_4)$  stand for  $(\sin(a_1(q)s^1 + a_2(q)s^2), \cos(a_1(q)s^1 + a_2(q)s^2))$ . Hence to compute the double integral of some expression it is first parsed into this polynomial form and then the total



(a) Constraint violation without stabilization.



(b) Constraint violation with stabilization.



(c) Curvature evolution with stabilization.

Figure 9.2: Numerical solution of the pure bending dynamics of the constant curvature shell patch.

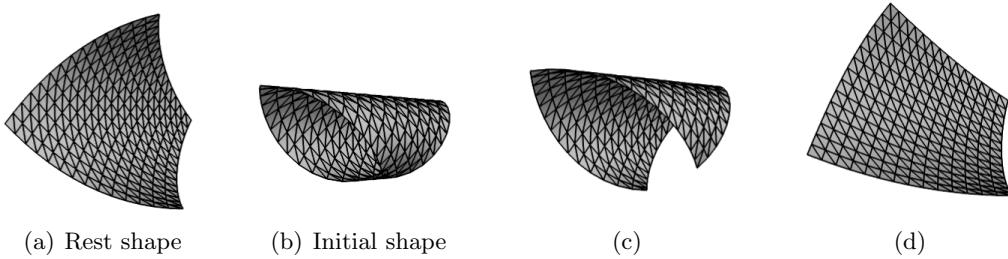


Figure 9.3: (b-d) Dynamical oscillations of our inextensible shell patch clamped at bottom right corner and subject to gravity, with rest shape (a) and initial shape (b).

integral is computed as a sum of coefficients times integrals of base functions

$$\sum_{p_1, p_2, p_3, p_4} f_p(q) \int_0^L \int_0^W (s^1)^{p_1} (s^2)^{p_2} \sin^{p_3} (a_1(q)s^1 + a_2(q)s^2) \cos^{p_4} (a_1(q)s^1 + a_2(q)s^2) ds^2 ds^1,$$

which are manageable by Sympy.

Ultimately, there were too many drawbacks using Sympy. It required the manually simplified expressions for the position in Equations (9.12) and (9.13), otherwise the resulting integrations appeared to be too complex to be handled by this computer algebra system. Sub cases have to be taken into account, depending on whether  $a_1(q)$  or  $a_2(q)$  are non zero. It takes about 5 or 6 hours (per sub case) to generate the code for the dynamics matrices using this Sympy method on a PC with a 2.6 GHz processor and 16 GB of RAM.

**Non Integrability on a circular domain.** One of the reasons we at first computed the mass matrix and its derivatives symbolically is that this was possible for the super-helix element. For a rectangular shell is also possible yet cumbersome as we saw. For a shell with a circular domain the symbolic integrations required to compute the mass matrix can no longer be resolved. If the domain  $D$  is a circle of radius 1 centred at 0, then there is no closed form expression for the mass matrix. If we parametrize the domain as  $\{(s^1, s^2) : s^1 \in [-1, 1], s^2 \in [-\sqrt{1 - (s^1)^2}, \sqrt{1 - (s^1)^2}]\}$ , then one of the integrals involved in the computation is

$$\int_{-1}^1 \int_{-\sqrt{1-(s^1)^2}}^{\sqrt{1-(s^1)^2}} (s^1)^4 \cos^2(a_1 s^1 a_2 s^2) ds^2 ds^1,$$

which according to Maple has no closed form expression. The obstacle is that there is no closed form expression for

$$\int_{-1}^1 \cos(\sqrt{1-x^2}) dx.$$

The symbolic approach features this unexpected limitation, independently of the computer algebra system which we choose to use.

**Computing the Mass matrix and derivatives with Gauss Legendre quadrature and automatic differentiation.** Because of the flexibility it provides we opted for numerical quadrature to compute the double integrals in the equations of motion. Given a function  $f(q, s^1, s^2)$  we approximate the integral  $\iint_D f(q, s^1, s^2) ds^1 ds^2$  using a Gauss Legendre method of order  $n$  as

$$\int_0^L \int_0^W f(q, s^1, s^2) ds^2 ds^1 = \frac{LW}{4} \sum_{i=1}^n \sum_{j=1}^n w_i w_j f(q, \frac{L}{2} x_i + \frac{L}{2}, \frac{W}{2} x_j + \frac{W}{2}),$$

where  $w_i$  are the weights and  $x_i$  the nodes (defined in the interval  $-1, 1$ ). In practice we use a Gauss-Legendre method of order 5. The relative difference in mass matrices computed using this method and the symbolic method is of the order  $10^{-8}$  for reasonable values of  $q$ . This number can be taken to  $10^{-16}$  by using a method of order 10. Once we have an algorithm to evaluate the mass matrix we can obtain a code for the derivatives of the mass matrix by using automatic differentiation. In practice we do this using Casadi [68]. An advantage of using numerical quadrature to compute the mass matrix is that we don't really need a closed form expression for the function  $f$ , but only a numerical method which allows to evaluate  $f$  at the required quadrature nodes. Hence one could do without the constant curvature assumption and try other shape functions for the curvatures. Using this remark we built an inextensible shell patch model with linear curvatures, the shell equivalent of a super space clothoid element [54].

### 9.2.5 Difficulties

**Singularity at  $q = 0$ .** When the shell is in a flat configuration, namely  $q = 0$ , then the gradient of the constraints vanishes. The constraint set  $\{q \in \mathbb{R}^3 : q_0 q_1 - q_2^2 = 0\}$  is a cone centred at the origin and is therefore not a manifold in any neighbourhood of the origin. The Delassus operator  $G(q)M^{-1}(q)G^T(q)$  becomes singular. This singularity is absent in the clothoidal shell patch which we propose in the next section.

**Assembling shell patches.** The difficulties in assembling shell patches comes from the fact that with curvature based coordinates one has no control over the location of the rulings. Say we want to stitch two shell patches along the coordinate line  $s^1 = L$ , so that we have  $C^1$  continuity. Then along this line both patches share the same position and frame. It follows that the curvatures  $q_1, q_2$  of the first patch are identical to the curvatures  $\tilde{q}_1, \tilde{q}_2$  of the second patch. Two situations can arise. If it happens that  $\tilde{q}_1 \neq 0$  then by the Gauss equation of the second patch  $\tilde{q}_0 = \frac{\tilde{q}_2^2}{\tilde{q}_1}$  and so all three degrees of freedom of the second patch are removed. On the other hand, if  $\tilde{q}_1 = 0$ , then  $\tilde{q}_2 = 0$  and  $\tilde{q}_0$  is left as a free variable. In this latter configuration the two patches meet at a straight line, a ruling of both patches, parallel to the vector  $e_2$ . We could thus describe only shell structures with a cylindrical symmetry.

## 9.3 Extension of the model

To go beyond the spatially constant curvature shape functions we have two options. The first is to use any arbitrary basis functions and enforce the compatibility conditions as constraints. The second is to find shape functions which automatically satisfy the compatibility conditions. In this section we explore the latter approach. First we expose how to construct solutions to the GCMP equations. In particular we will expose a set of linear shape functions which satisfy the GCMP equations and then we use them to build the dynamics of a developable shell patch with spatially linear curvatures.

### 9.3.1 Explicit solutions to the compatibility conditions

Let us denote  $X$  and  $Y$  the column vectors of the second fundamental form:  $X := (b_{11}, b_{12})^T$  and  $Y := (b_{12}, b_{22})^T$ . Then the curl of  $X$  reads

$$\text{curl}X = \partial_1 b_{12} - \partial_2 b_{11}$$

and the curl of  $Y$  reads

$$\text{curl}Y = \partial_1 b_{22} - \partial_2 b_{12}.$$

Hence the GCMP equations (8.8) for a developable shell may be written equivalently as

$$\begin{cases} \text{curl}X &= 0 \\ \text{curl}Y &= 0 \\ X \times Y &= 0 \end{cases}. \quad (9.16)$$

Since the domain  $D$  is simply connected, a vector field on  $D$  is of vanishing curl if and only if it derives from a potential (i.e., it is a gradient field). Hence the conditions (9.16) hold if and only if there exist two scalar functions  $\Phi, \Psi$  defined on the simply connected domain  $D$  such that

$$\begin{cases} X &= \nabla\Phi \\ Y &= \nabla\Psi \\ \nabla\Phi \times \nabla\Psi &= 0 \end{cases}. \quad (9.17)$$

The last equation says that  $\nabla\Phi$  and  $\nabla\Psi$  are collinear so that there is no need for 2 potentials but rather one potential and a collinearity factor. Hence we can restate (9.17) as

$$\begin{cases} \exists\Phi/ & X = \nabla\Phi \\ \exists\Psi/ & Y = \nabla\Psi \\ \exists\alpha, \beta/ & \nabla\Psi = \alpha\nabla\Phi \text{ or } \nabla\Phi = \beta\nabla\Psi \end{cases}. \quad (9.18)$$

Then necessarily one has either the following (expressing everything in terms of  $\Phi, \alpha$ ) or its counterpart (expressed in  $\Psi, \beta$ )

$$\begin{cases} X = \nabla\Phi \\ Y = \alpha\nabla\Phi \\ \partial_2\Phi = \partial_1\Psi = \alpha\partial_1\Phi \end{cases}.$$

More explicitly we can write

$$\begin{cases} \partial_1\Phi = b_{11} \\ \partial_2\Phi = b_{12} \\ b_{12} = \alpha b_{11} \\ b_{22} = \alpha b_{12} \end{cases}.$$

It follows that the potential  $\Phi$  obeys a **transport equation**

$$\partial_2\Phi - \alpha\partial_1\Phi = 0. \quad (9.19)$$

Writing the condition that the curl of  $Y$  vanishes and taking into account that  $Y = \alpha\nabla X$  we can deduce that

$$(\partial_2\alpha - \alpha\partial_1\alpha)\partial_1\Phi = 0. \quad (9.20)$$

Indeed we have

$$\begin{aligned} \operatorname{curl}Y = 0 \iff & \partial_2(\alpha\partial_1\Phi) - \partial_1(\alpha\partial_2\Phi) = 0 \\ & \partial_2\alpha\partial_1\Phi + \alpha\partial_{21}\Phi - \partial_1\alpha\partial_2\Phi - \alpha\partial_{12}\Phi = 0 \\ & \partial_2\alpha\partial_1\Phi - (\partial_1\alpha)(\alpha\partial_1\Phi) = 0. \end{aligned}$$

The regions where  $\partial_1\Phi = 0$ , in which case  $b = 0$ , are planar regions. Outside of planar regions  $\alpha$  must satisfy then the **inviscid Burgers' equation**

$$\partial_2\alpha - \alpha\partial_1\alpha = 0. \quad (9.21)$$

Once  $\alpha$  is known,  $\Phi$  is obtained as a solution of a linear transport equation  $\partial_2\Phi - \alpha\partial_1\Phi = 0$ , and all of the coefficients of the second fundamental form are thus determined. It is known that all the solutions of the inviscid Burger's equation can be represented as

$$\alpha(s^1, s^2) = F(s^1 + \alpha(s^1, s^2)s^2)$$

for any arbitrary function  $F : \mathbb{R} \rightarrow \mathbb{R}$ . And the solutions of the linear transport equation are obtained as

$$\Phi(s^1, s^2) = G(s^1 + \alpha(s^1, s^2)s^2)$$

for any arbitrary function  $G : \mathbb{R} \rightarrow \mathbb{R}$ .

**The case of constant  $\alpha$ .** Let us now turn our attention to the case where  $\alpha$  is constant. The inviscid Burgers' equation (9.21) is then automatically satisfied. The second fundamental form is then given by

$$b = G'(s^1 + \alpha s^2) \begin{bmatrix} 1 & \alpha \\ \alpha & \alpha^2 \end{bmatrix}.$$

One can see that such a tensor field  $b$  is clearly symmetric and satisfies  $\det(b) = 0$ . Furthermore one can check that  $b$  satisfies the other two Codazzi-Mainardi-Peterson equations since

$$\partial_1 b_{12} - \partial_2 b_{11} = \alpha G'' - \alpha G'' = 0$$

and

$$\partial_1 b_{22} - \partial_2 b_{12} = \alpha^2 G'' - \alpha^2 G'' = 0.$$

If we take  $G'$  to be an affine function, such that  $G'(x) = \beta x + \gamma$  for arbitrary real numbers  $\beta, \gamma$  then the resulting second fundamental form is affine in  $s^1$  and  $s^2$ , its coefficients read

$$\begin{aligned} b_{11}(s^1, s^2) &= \beta s^1 + \alpha \beta s^2 + \gamma \\ b_{22}(s^1, s^2) &= \alpha^2 \beta s^1 + \alpha^3 \beta s^2 + \alpha^2 \gamma \\ b_{12}(s^1, s^2) &= \alpha \beta s^1 + \alpha^2 \beta s^2 + \alpha \gamma, \end{aligned}$$

**Remark** It is easy to see that the resulting geometry has a cylindrical symmetry. The eigenvector of  $b$  associated to the zero eigen value is  $(\alpha, -1)$ . Since  $\alpha$  is constant the latter implies that the rulings (lines of zero curvature) are all parallel and hence that the surface has a cylindrical symmetry. The other lines of curvature, orthogonal to the rulings, are planar curves with affine curvature, they are clothoids. Hence we will speak of a 'clothoidal shell patch'.

In the following section we consider the Lagrangian dynamics of a shell patch with such a second fundamental form by letting the coefficients  $\alpha, \beta, \gamma$  vary in time and be the generalized coordinates of the system.

### 9.3.2 Dynamics of a clothoidal shell patch

Consider the following shape functions for the second fundamental form:

$$\begin{aligned} b_{11}(s^1, s^2) &= \beta s^1 + \alpha \beta s^2 + \gamma \\ b_{22}(s^1, s^2) &= \alpha^2 \beta s^1 + \alpha^3 \beta s^2 + \alpha^2 \gamma \\ b_{12}(s^1, s^2) &= \alpha \beta s^1 + \alpha^2 \beta s^2 + \alpha \gamma, \end{aligned}$$

where  $\alpha, \beta, \gamma$  are any real numbers. Then the GCMP equations are automatically satisfied. We are now going to take these three numbers to be the generalized coordinates of the shell patch,  $q := [\alpha, \beta, \gamma]^T$ . There are two main differences with the constant curvature case, the first is that there is no closed form expression for the position of the surface patch as a function of the generalized coordinates. Instead, we have a numerical algorithm which given any  $q$  and any point  $(s^1, s^2)$  will compute the position  $r(q, s^1, s^2)$ . We use an explicit Runge-Kutta-Munthe-Kaas method of order 2 (see equation (6.23)) to solve the reconstruction equations (9.3) numerically. To compute the derivatives  $\frac{\partial r}{\partial q}$ , automatic differentiation is used. For the computation of the mass matrix we need the values of  $\frac{\partial r}{\partial q}$  at the quadrature nodes. The second difference with the constant curvature case is that the resulting dynamics is not a DAE but just an ODE since our particular choice of spatially linear shape functions satisfies all 3 of the compatibility equations at once. Geometrically the surface is a patch of a cylinder with a clothoidal base, as seen in Figure 9.4 (a). In Figure 9.4 (b) we plot the oscillations of the generalized coordinates  $(\alpha, \beta, \gamma)$  throughout the simulation. The shell is clamped at the origin with the frame  $R_0 = I$ .

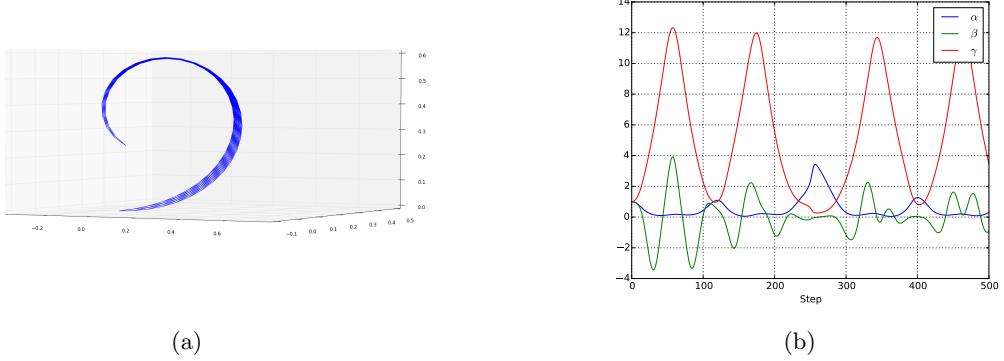


Figure 9.4: (a) Profile view of the shell patch (b) Dynamic oscillations of  $q$

## 9.4 Conclusion

**Recapitulation** In this Part III of our thesis we have given an exposition of the strain approach for shells. We have extended the construction of the strain based elements [1, 54] for inextensible and unshearable rods to the case of developable shells. Our approach to simulating the pure bending dynamics of developable shells consists in employing a method of lines discretization. The action functional of the shell is first discretized in space so as to obtain a classical Lagrangian mechanical system. The spatial discretization is achieved by introducing the components of the second fundamental form of the mid surface (its bending strains) as the degrees of freedom of the system. This results in motions of the developable shell which intrinsically keep its mid surface from stretching or shearing, without having to rely on extremely stiff membrane forces.

It was unclear in the literature how to handle situations where one cannot separate the dynamic evolution of the strains from that of the displacements. In our approach this problem is dealt with by having a numerical algorithm which given the strains  $q$  can compute the position and orientation of the mid surface  $(R, r)(q, s^1, s^2)$  at any point  $(s^1, s^2)$  in the domain. The algorithm solves the surface reconstruction equations (9.3) by employing a Runge-Kutta-Munthe-Kaas (RKMK) method. In our spatially constant curvature patch an RKMK method of order 1, the Lie-Euler method, yields the exact expression for the frame and position: just as in the case of the super-helix element for Kirchhoff rods. While for the spatially linear shell patch an RKMK method of order 2 yields a numerical approximation of the surface position and orientation.

Though limited to cylindrical configurations, our inextensible shell patches already feature some rich motions and could serve as interactive models for simulating small flexible surfaces such as leaves or feathers in computer graphics applications.

**Limitations.** Two main obstacles prevent us from employing our methods for a general strain based finite element analysis of developable shell structures. The first is the fact that patching strain based shell elements together is not as straightforward as in the case of rods. Developable surface patches should be patched together along a common ruling. Since in our formulation we have no control on the placement of the rulings, a smooth enough junction of our surface patches is in practice impossible to achieve. Only in cases with particular symmetries is it possible. The second limitation is the need to solve the Gauss-Codazzi-Mainardi-Peterson equations. Even if these compatibility equations are satisfied on each patch they aren't necessarily satisfied across the patches.

**Prospective Work.** In our opinion, the most straight forward way to extend our methods in order to tackle more challenging scenarios would be to develop a spectral method. That is, instead of trying to have many surface patches with a low number of degrees of freedom each, we could keep just one patch with a high number of degrees of freedom. It would also be interesting to include the angular and linear velocities of the mid-surface in time within the formulation. This would lead to an intrinsic formulation in both space and time. As noted in [116], one can then avoid having to impose the spatial compatibility equations, at the cost of imposing two easier space-time compatibility equations.

# Perspectives

In this thesis we have tackled some of the challenges posed by the numerical simulation of slender elastic structures in contact. The problems considered in each of the three parts of this manuscript, stem from multiple motivations and lead to conclusions and perspectives already resumed in the introduction and conclusions of each part. Our goal in this final section is not to repeat ourselves but to give some final perspectives on the topics considered. A driving motivation of ours during this thesis was the desire to better understand and to improve the numerical methods developed in our team for the simulation of thin elastic rods in contact [1, 61, 54].

One frequently encountered problem in these simulations, in the presence of frictional contact, is that the Coulomb friction solver may fail to compute the contact forces. Our investigations in Part I of this manuscript give more insights as to why Painlevé type singularities might arise, where the contact forces cease to exist. We learned that friction introduces a perturbation to the problem defining the contact forces. If this perturbation is large enough, then the properties which make the frictionless contact problem well-posed, such as the positivity of the Delassus operator, may then be lost. From a numerical standpoint, a strategy to detect and to resolve such singular situations, so as to continue the simulation, is still needed. One option may be to adaptively add or remove degrees of freedom in the system, reflecting how the physical system may deform in situations where hard constraints and compressive contact forces are involved. Another option, proposed by Alart [120], is to modify the Coulomb friction law for large normal contact forces and to consider a Coulomb-Orowan law. This last approach is backed by experimental evidence, at least in the context of granular media.

The methods developed in [1, 54], for the simulation of Kirchhoff rods, were initially intended for computer graphics applications. Their place with respect to other methods in the computer graphics literature for the simulation of Kirchhoff rods was already well understood. In the second part of this manuscript we show that in the context of flexible multibody dynamics, these methods can be classified as strain based methods. We like to think that our study of Kirchhoff rods has brought closer together the two fields of computer graphics and multibody dynamics, at least in the intersection that is the numerical simulation of rods.

The common theme in the second and third parts of this manuscript, is the investigation of strain based discretizations of slender, inextensible and unshearable elastic structures. Part II deals with the case of rods, while Part III deals with the case of shells. One of the difficulties in dealing with a slender structure, is that even if its deformations are of small strains, so that linear elasticity can be considered, the resulting displacements are large and the geometry of the problem is non linear. This results in a non linear relationship between the strains and the generalized displacements of the structure. If we use the strains as primary degrees of freedom, then the elastic potential energy is very simple, it is convex quadratic. However, all terms in the energy or the forces, involving the position and the orientation of the structure, become more involved non linear expressions of the degrees of freedom. Hence, using strain based methods is convenient when the principal source of difficulty in the problem is due to the intrinsic elastic

forces, and not to extrinsic (position and orientation dependent) forces.

While the strain based approach for rods has been quite successful, the strain based approach for shells remains little explored. We believe however, that the strain based approach for plates and shells could prove effective in highly intrinsic problems such as those considered in soft matter physics [121]. For the non-Euclidean plates considered in [121], the second fundamental form of the plate in its reference configuration is null, while its first fundamental form is hyperbolic. The reference configuration of non-Euclidean plates is thus incompatible and cannot be an equilibrium configuration. It would be a good idea to pursue strain based methods in such a problem, where the main source of difficulty comes from the surface compatibility equations and not from extrinsic effects.

## Appendix A

# Convex Analysis and Complementarity theory

### A.1 Some convex analysis and complementarity theory tools

**Subdifferential.** Let  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  be a convex proper function. Its subdifferential at  $x \in \mathbb{R}^n$ , denoted as  $\partial f(x)$  is the set of its subgradients:

$$\partial f(x) := \{g \in \mathbb{R}^n : f(x) + g^T(y - x) \leq f(y) \quad \forall y \in \mathbb{R}^n\}.$$

**Indicator function.** Let  $K$  be a set, its indicator function is  $\Psi_K(x) = 0$  if  $x \in K$  and  $= +\infty$  if  $x \notin K$ .

**Normal and tangent cones.** If  $K \subset \mathbb{R}^n$  is non-empty closed and convex, its normal cone at  $x$  is

$$N_K(x) = \{z \in \mathbb{R}^n | z^T(y - x) \leq 0 \text{ for all } y \in K\} = \partial\Psi_K(x).$$

If  $K = \mathbb{R}^n$  then  $N_K(x) = \{0\}$  for all  $x$ .

The tangent cone to  $K$  at  $x$  is defined as the set

$$T_K(x) := \{z \in \mathbb{R}^n | z^T y \leq 0 \text{ for all } y \in N_K(x)\}.$$

If  $K = \{x \in \mathbb{R}^n | f_i(x) \geq 0\}$  for  $m$  continuously differentiable functions  $f_i : \mathbb{R}^n \mapsto \mathbb{R}$  which satisfy the Mangasarian-Fromovitz constraint qualification [122], then  $T_K(x) = \{v \in \mathbb{R}^n | v^T \nabla f_i(x) \geq 0, \text{ for all } i \in \{1, \dots, m\} \text{ such that } f_i(x) = 0\}$ .

**Linear Complementarity Problem (LCP)** A linear complementarity problem LCP( $q, M$ ) with unknown  $x \in \mathbb{R}^n$  is:  $x \geq 0$ ,  $Mx + q \geq 0$ ,  $x^T(Mx + q) = 0$ . More compactly

$$0 \leq x \perp Mx + q \geq 0.$$

An LCP is said **solvable** if it has at least one solution.

A matrix  $M$  is a **P-matrix** if and only if the LCP has a unique solution for any  $q$  [40].

**Mixed Linear Complementarity Problem (MLCP)** A mixed linear complementarity problem (MLCP) is a problem of the form: find vectors  $x, y$  such that  $M_{11}x + M_{12}y + q_1 = 0$  and  $0 \leq x \perp M_{21}x + M_{22}y + q_2 \geq 0$  for given matrices  $M_{ij}$  and vectors  $q_i$ .

**Dual set.** Let  $C$  be a (non-necessarily convex) set of  $\mathbb{R}^n$ , then its dual set is  $C^* = \{x \in \mathbb{R}^n | x^T z \geq 0 \text{ for all } z \in C\}$ , which is always a closed convex cone.

**Projection.** Let  $M \in \mathbb{R}^{n \times n}$  be a symmetric positive definite matrix defining the inner product  $x^T M x$ . With this metric, the orthogonal projection of a vector  $x \in \mathbb{R}^n$  on a convex set  $K \subset \mathbb{R}^n$  is denoted as  $\text{proj}_M[K; x] = \operatorname{argmin}_{z \in K} \frac{1}{2}(z - x)^T M(z - x)$ .

The following equivalences are useful. Let  $x \in \mathbb{R}^n$ ,  $q \in \mathbb{R}^n$ ,  $M \in \mathbb{R}^{n \times n}$ ,  $K$  is a closed convex cone.

$$\begin{aligned} Mx + q \in -N_K(x) &\Leftrightarrow K \ni x \perp Mx + q \in K^* \\ &\Leftrightarrow (\text{if } M = M^T \succeq 0) x = \operatorname{argmin}_{z \in K} \frac{1}{2}z^T Mz + q^T z \\ &\Leftrightarrow (\text{if } M = M^T \succ 0) x = \text{proj}_M[K; -M^{-1}q] \end{aligned} \quad (\text{A.1})$$

## A.2 Theorem 3.1.7 in [40] (excerpts)

**Theorem A.2.1.** Let  $M \in \mathbb{R}^{n \times n}$  be positive semi definite, and let  $q \in \mathbb{R}^n$  be arbitrary. The following statements hold:

- (a) If  $z_1$  and  $z_2$  are two solutions of the LCP( $M, q$ ) then  $(z_1)^T(q + Mz_2) = (z_2)^T(q + Mz_1)$ .
- (b) If  $M$  is symmetric (as well as positive semi definite) then  $Mz_1 = Mz_2$  for any two solutions  $z_1$  and  $z_2$  of the LCP( $M, q$ ).

## A.3 Theorem 3.8.6 in [40]

**Theorem A.3.1.** Let  $M \in \mathbb{R}^{n \times n}$  be copositive and let  $q \in \mathbb{R}^n$  be given. If the implication  $[0 \leq v \perp Mv \geq 0] \Rightarrow [v^T q \geq 0]$  is valid, then the LCP( $M, q$ ) is solvable.

Let  $\mathcal{Q}_M$  denote the solution set of the homogeneous LCP( $M, 0$ ). This theorem can be restated equivalently as: If  $M$  is copositive and  $q \in \mathcal{Q}_M^*$  then LCP( $M, q$ ) is solvable.

## A.4 Theorems 2.8 and 2.11 in [123]

Chen and Xiang [123] stated very useful criteria that guarantee that a positive definite or a P-matrix remains positive definite or P when it is subject to a small enough perturbation. We give here just an excerpt of the results in [123], and a corollary of it.

**Theorem A.4.1.** If  $M$  is a P-matrix then all matrices  $A$  such that  $\beta_2(M) \|M - A\|_2 < 1$  are P-matrices, where  $\beta_2(M) := \max_{d \in [0,1]^n} \|(I - D + DM)^{-1}D\|_2$ , and  $D = \text{diag}(d)$ . When  $M$  is symmetric positive definite,  $\beta_2(M) = \|M^{-1}\|_2$ .

**Theorem A.4.2.** Let  $M \in \mathbb{R}^{n \times n}$  be a positive definite matrix. Then every matrix  $A \in \{A | \left\| \left( \frac{M+M^T}{2} \right)^{-1} \right\|_2 \|M - A\|_2 < 1\}$  is positive definite.

The next corollary is proved in [124, Corollary 2].

**Corollary A.4.2.1.** Let  $A = B + C$ , where  $A, B$  and  $C$  are  $n \times n$  real matrices, and  $B \succ 0$ , not necessarily symmetric. If  $\|C\|_2 < \frac{1}{\left\| \left( \frac{B+B^T}{2} \right)^{-1} \right\|_2}$  then  $A \succ 0$ .

## Appendix B

# KKT system: solvability and solution uniqueness

The KKT problem in (2.2) is ubiquitous in the study of mechanical systems with bilateral holonomic constraints. In this section it is analysed from various points of view, and proves to possess some subtleties depending on which assumptions are made on the data. Let us consider the next three problems, where  $M = M^T \in \mathbb{R}^{n \times n}$  and  $M \succeq 0$ ,  $F \in \mathbb{R}^{n \times m}$ ,  $N \in \mathbb{R}^{n \times m}$ ,  $x \in \mathbb{R}^n$ ,  $y \in \mathbb{R}^m$ ,  $z \in \mathbb{R}^n$ ,  $a \in \mathbb{R}^n$ ,  $b \in \mathbb{R}^m$ :

$$\underbrace{\begin{pmatrix} M & I \\ F^T & 0 \end{pmatrix}}_{\triangleq \tilde{A} \in \mathbb{R}^{(n+m) \times 2n}} \begin{pmatrix} x \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} \quad (\text{B.1})$$

and

$$\underbrace{\begin{pmatrix} M & -F \\ F^T & 0 \end{pmatrix}}_{\triangleq \bar{A} \in \mathbb{R}^{(n+m) \times (n+m)}} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} \quad (\text{B.2})$$

and

$$\underbrace{\begin{pmatrix} M & N \\ F^T & 0 \end{pmatrix}}_{\triangleq \hat{A} \in \mathbb{R}^{(n+m) \times (n+m)}} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} \quad (\text{B.3})$$

The three systems (B.1), (B.2) and (B.3) correspond to various ways to consider the system in (2.2), where  $x$  is for  $\ddot{q}$ ,  $z$  is for  $-\nabla h_{n,b}(q)\lambda_{n,b}$  and  $y$  is for  $\lambda_{n,b}$ .

- (i) Let us consider first (B.1) without any assumption on the form of  $z$ . From [28, Fact 2.10.22] one has  $\text{Im}(\tilde{A}) = \text{Im} \begin{pmatrix} M \\ F^T \end{pmatrix} + \text{Im} \begin{pmatrix} I \\ 0 \end{pmatrix}$ . Thus a necessary and sufficient condition for (B.1) to possess a solution  $(x, z)$  for any  $a$  and  $b$ , equivalently  $\text{rank}(\tilde{A}) = n + m$  (which also follows from [28, Proposition 6.1.7 (iii)]), is that  $F^T$  has full rank  $m$ , i.e.,  $F$  be full column rank (this implies that  $n \geq m$ ). Uniqueness of  $(x, z)$  for any  $a$  and  $b$  holds if and only if  $n = m$  and  $\text{rank}(\tilde{A}) = 2n$  (this may be proved from [28, Theorem 2.6.3 ii)]), in which case  $F$  is square and has full rank  $n$ . In this case the solution is equal to  $\tilde{A}^\dagger \begin{pmatrix} a \\ b \end{pmatrix}$ , where  $\tilde{A}^\dagger$  is the Moore-Penrose generalized inverse of  $\tilde{A}$  [28, Proposition 6.1.7 (viii)]. One sees that  $M$  plays no role in this system.

- (ii) Let us still consider (B.1) assuming that  $z = -Fy$  for some  $y$ , i.e.,  $z \in \text{Im}(F)$ . One has  $(x, z) \in \text{Ker}(\tilde{A}) \Leftrightarrow Mx = -z$  and  $x \in \text{Ker}(F^T)$ . Using  $z \in \text{Im}(F)$  and  $x \in \text{Ker}(F^T)$  and [28, Theorem 2.4.3] it follows that  $x \perp z$ . We also have  $z \in \text{Im}(M)$ . Using that  $\text{Im}^\perp(M) = \text{Ker}(M)$ , we deduce that  $x \in \text{Ker}(M)$  because  $M$  is symmetric positive semi definite<sup>1</sup>. Thus  $x \in \text{Ker}(M) \cap \text{Ker}(F^T)$  and consequently  $z = -Mx = 0$ . Thus we have shown that  $[(x, z) \in \text{Ker}(\tilde{A}) \text{ and } z \in \text{Im}(F)] \Rightarrow [z = 0 \text{ and } x \in \text{Ker}(M) \cap \text{Ker}(F^T)]$ , and the reverse implication holds also. Let  $S = \{(x, z) \in \mathbb{R}^{2n} | z \in \text{Im}(F)\}$ . Then  $\text{Ker}(\tilde{A}) \cap S = \{(x, z) \in \mathbb{R}^{2n} | z = 0 \text{ and } x \in \text{Ker}(M) \cap \text{Ker}(F^T)\}$ . One infers that  $\text{Ker}(M) \cap \text{Ker}(F^T) = \{0\} \Rightarrow \text{Ker}(\tilde{A}) \cap S = \{0\}$ . From (i) existence of solutions for system (B.1) holds for any  $a$  and  $b$  if and only if  $\text{rank}(\tilde{A}) = n + m \Leftrightarrow \text{rank}(F) = m$ , hence  $\dim(\text{Ker}(\tilde{A})) = 2n - n - m = n - m$ . From the fact that  $0 \leq \dim(\text{Ker}(\tilde{A}) \cap S) \leq \min(n - m, n + m) = n - m$  [28, Fact 2.9.14], one infers that  $\text{Ker}(M) \cap \text{Ker}(F^T) = \{0\} \Rightarrow \dim(\text{Ker}(\tilde{A}) \cap S) = 0$ , thus  $n = m$ . In this case it follows from (i) that the system has a unique solution for any  $a$  and  $b$ . Conversely the existence of solutions for arbitrary  $a$  and  $b$  and  $n = m$  imply uniqueness, as well as  $\text{Ker}(M) \cap \text{Ker}(F^T) = \{0\}$  since  $F$  is square full rank  $n$ .

Let us now pass to the system (B.2). Remark that if  $\bar{A}$  in (B.2) is invertible then  $F$  necessarily has full rank  $m$ . This follows from the fact that  $(x, y) \in \text{Ker}(\bar{A})$  implies  $x \in \text{Ker}(M) \cap \text{Ker}(F^T)$  and  $y \in \text{Ker}(F)$ , using similar arguments as in (ii). In particular if there are more constraints than degrees of freedom (i.e.,  $m > n$ ) then  $\bar{A}$  is not invertible, and likewise if  $\text{rank}(F) = r < m$ .

- (iii) Consider now the system (B.2). Let  $\text{rank}(F) = m$  (so  $m \leq n$ ) and  $M$  be positive semi definite. Then existence and uniqueness of both  $x$  and  $y$  for arbitrary  $a$  and  $b$  (equivalently, non-singularity of  $\bar{A}$ ) holds if and only if  $\text{Ker}(M) \cap \text{Ker}(F^T) = \{0\}$  (proof by direct application of [34, p.523], or using the above expression of  $\text{Ker}(\bar{A})$ ).

The rank condition on  $F$  appears to be in fact necessary and sufficient as alluded to few lines above:

- (iii') The system (B.2) has a unique solution  $(x, y)$  for arbitrary  $a$  and  $b$  if and only if  $\text{rank}(F) = m$  and  $\text{Ker}(M) \cap \text{Ker}(F^T) = \{0\}$ . This solution is equal to  $\bar{A}^{-1} \begin{pmatrix} a \\ b \end{pmatrix}$ .

The proof of (iii') follows from [28, Theorem 2.6.3, Proposition 6.1.7], noting that  $\bar{A}$  is square. It is sometimes wrongly stated that  $\bar{A}$  is non-singular if and only if  $M$  and  $F$  are both full rank matrices [31], which is only a sufficient condition. In fact one has from [28, Fact 6.4.20]:

$$\text{rank}(\bar{A}) = \text{rank}(M) + 2\text{rank}(F) - \dim[\text{Im}(M) \cap \text{Im}(F)] - \dim[\text{Im} \left( \begin{array}{c} M \\ F^T \end{array} \right) \cap \text{Im} \left( \begin{array}{c} F \\ 0 \end{array} \right)] \quad (\text{B.4})$$

The formula in (B.4) shows that one may dispense with positive definiteness conditions on  $M$ , and that the non-singularity of  $\bar{A}$  results from an interplay between the matrices ranges. Consider for instance  $M = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$  and  $F = (1 \ 0)^T$ , which yields  $\text{rank}(\bar{A}) = 3$ . In Contact Mechanics we wish to allow for situations where the constraints are redundant, so ( $\text{rank}(F) = r < \min(m, n)$ ) but which are nevertheless compatible (i.e.  $b \in \text{Im}(F^T)$ ), for otherwise the problem has no solution. Thus the most relevant problem is that of determining  $x$  and  $y$  such that (B.2) holds for arbitrary  $a$  with the additional assumption that  $b \in \text{Im}(F^T)$ . This problem is thus different from problems tackled in (ii) and (iii'), and corresponds to the problem tackled in [12, 125, 13].

---

<sup>1</sup>The conclusion does not hold without the positive definiteness condition.

- (iv) The necessary and sufficient condition for the existence of  $x$  and  $y$  with uniqueness of  $x$  and  $Fy$  such that (B.2) holds for arbitrary  $a$  is that  $b \in \text{Im}(F^T)$  and  $\text{Ker}(M) \cap \text{Ker}(F^T) = \{0\}$ .

**Proof:**

$\Leftarrow$

*Existence of  $x$  and  $y$ :* By contraposition, if not ( $\forall a, \exists x$  and  $y$  such that (B.2) holds) then there exists an  $a \in \mathbb{R}^n$  such that for all  $x, y, F^T x \neq b$  (so  $b \notin \text{Im}(F^T)$ ) or  $Mx - Fy \neq a$  (so  $\text{Im}([M \ F]) \neq \mathbb{R}^n$ , i.e.,  $\text{Ker}(M) \cap \text{Ker}(F^T) \neq \{0\}$ ).

*Uniqueness of  $x$  and  $Fy$ :* Suppose  $b \in \text{Im}(F^T)$  and  $\text{Ker}(M) \cap \text{Ker}(F^T) = \{0\}$ .

If  $(x_1, y_1)$  and  $(x_2, y_2)$  are two solutions of (B.2) then  $F^T(x_1 - x_2) = 0$  and  $M(x_1 - x_2) = F(y_1 - y_2)$ . Hence  $(x_1 - x_2) \in \text{Ker}(F^T)$  and  $M(x_1 - x_2) \in \text{Im}(F) = \text{Ker}(F^T)^\perp$ . Hence  $(x_1 - x_2)^T M(x_1 - x_2) = 0$  and since  $M$  is symmetric positive semi definite this means  $(x_1 - x_2) \in \text{Ker}(M)$ . By hypothesis  $\text{Ker}(M) \cap \text{Ker}(F^T) = \{0\}$  so one concludes that  $x_1 - x_2 = 0$ , that is,  $x$  is unique. It follows that  $Fy_1 = Fy_2$  so  $Fy$  is unique as well.

$\Rightarrow$

Suppose that for arbitrary  $a$ , there exists  $x$  and  $y$  such that (B.2) holds with uniqueness of  $x$  and  $Fy$ . Then in particular  $b \in \text{Im}(F^T)$ . Let  $x \in \text{Ker}(M) \cap \text{Ker}(F^T)$  and let  $(x^*, y)$  be the unique solution of (B.2). Then  $(x^* + x, y)$  is also a solution of (B.2). Hence  $x = 0$ , so that  $\text{Ker}(M) \cap \text{Ker}(F^T) = 0$ . ■

In [12, p.319] the condition  $b \in \text{Im}(F^T)$  is stated as: *acceleration constraints are compatible*. In conclusion, four types of systems are considered: system (B.2) with unknowns  $x$  and  $y$  in (iii), system (B.1) with unknowns  $x$  and  $z$  in (i), system (B.1) with unknowns  $x$  and  $z$  and the constraints that  $z \in \text{Im}(F)$  in (ii), system (B.2) with unknowns  $x$  and  $y$  with uniqueness of  $x$  and  $Fy$  and arbitrary  $a$  in (iv). To complete the picture let us note that  $\text{Ker}(M) \cap \text{Ker}(F^T) = \{0\} \Leftrightarrow (\text{Im}(M) + \text{Im}(F)) = \mathbb{R}^n \Leftrightarrow \text{Im}([M \ F]) = \mathbb{R}^n$  using [28, Fact 2.9.10], and we recover directly an alternative way to formulate the condition involving the kernels, sometimes used in the literature [13] [125, Equation (9)].

Problems like in (B.2) and (B.1) occur in frictionless systems. The next step is to consider systems of the form (B.3) for some matrix  $N$ . Such problems arise in the presence of Coulomb's friction, see (3.2). Using [28, Fact 6.4.20] one gets an extension of (B.4):

$$\text{rank}(\hat{A}) = \text{rank}(M) + \text{rank}(F^T) + \text{rank}(N) - \dim[\text{Im}(M) \cap \text{Im}(F)] - \dim[\text{Im}\left(\begin{array}{c} M \\ F^T \end{array}\right) \cap \text{Im}\left(\begin{array}{c} N \\ 0 \end{array}\right)] \quad (\text{B.5})$$

System (B.3) has a unique solution  $(x, y)$  for any  $a$  and  $b$  if and only if  $\text{rank}(\hat{A}) = n + m$ . One has:

$$\text{Im}\left(\begin{array}{c} M \\ F^T \end{array}\right) \cap \text{Im}\left(\begin{array}{c} N \\ 0 \end{array}\right) = \{z \in \mathbb{R}^{n+m} \mid \exists y_1 \in \text{Ker}(F^T), \exists y_2 \in \mathbb{R}^m, \text{such that } z = \left(\begin{array}{c} My_1 \\ 0 \end{array}\right) = \left(\begin{array}{c} Ny_2 \\ 0 \end{array}\right)\} \quad (\text{B.6})$$

and  $\text{Im}(M) \cap \text{Im}(F) = \{z \in \mathbb{R}^n \mid \exists y_1 \in \mathbb{R}^n, \exists y_2 \in \mathbb{R}^m, \text{such that } z = My_1 = Fy_2\}$ . It is clear from (B.5) that the system's well-posedness depends on the interplay between  $M$ ,  $F$  and  $N$ . Even if all three matrices have full rank, one may have  $\text{rank}(\hat{A}) < n + m$ . Suppose that  $\text{rank}(M) = n$ , and  $\text{rank}(F^T M^{-1} N) = m$  (implying that  $m \leq n$ ). Then  $\text{Im}(M) \cap \text{Im}(F) = \text{Im}(F)$  and  $\text{Im}\left(\begin{array}{c} M \\ F^T \end{array}\right) \cap \text{Im}\left(\begin{array}{c} N \\ 0 \end{array}\right) = \{0\}$  (since  $F^T y_1 = 0 = F^T M^{-1} N y_2$ ). Therefore from (B.5) one has  $\text{rank}(\hat{A}) = n + m$ .

Finally we may rewrite (B.3) as (B.1) posing  $z = Ny$ . Then (i) applies, but (ii) usually does not expect if  $\text{Im}(N) \subseteq \text{Im}(F)$ . Then given  $z$ , there exists a unique  $y$  if and only if  $N$  has full column rank  $m$  ( $\Rightarrow m \leq n$ ). Let  $N = -F + P$  for some matrix  $P$ .

(v) Assume that  $F$  has full column rank  $m$  (equivalently  $F^T F \in \mathbb{R}^{m \times m}$  is positive definite). Let us investigate conditions that guarantee that  $N$  has full rank  $m$  (equivalently  $N^T N \in \mathbb{R}^{m \times m}$  is positive definite). One has  $N^T N = F^T F - F^T P - P^T F + P^T P$ . A direct application of Corollary A.4.2.1, with matrices  $A, B, C$  chosen as  $A = N^T N$ ,  $B = F^T F$  and  $C = -F^T P - P^T F + P^T P$ , shows that a sufficient condition for  $N^T N$  to be positive definite is that  $\| -F^T P - P^T F + P^T P \|_2 < \frac{1}{\|(F^T F)^{-1}\|_2}$ , or equivalently  $\sigma_{\max}(-F^T P - P^T F + P^T P) < \sigma_{\min}(F^T F)$ .

Let  $M = 0$ , then using (B.5) it follows that  $\text{rank}(\hat{A}) = 2m$ , hence  $\text{rank}(\hat{A}) = n + m$  if and only if  $n = m$ . This shows that depending on the interplay between the ranges of the matrices in (B.5), the system in (B.3) may be solvable with uniqueness for any  $a$  and  $b$ , for low-rank matrices  $M$ .

(vi) Let us assume that  $\text{rank}(M) = n$  and study conditions such that the rank of  $F^T M^{-1} N = -F^T M^{-1} F + F^T M^{-1} P$  is  $m$ . Then as shown after (B.6),  $\hat{A}$  has rank  $n+m$  and the system (B.3) has a unique solution for any  $a$  and  $b$ . Using Corollary A.4.2.1 with matrices  $A, B, C$  chosen as  $A = -F^T M^{-1} N$ ,  $B = F^T M^{-1} F$  and  $C = -F^T M^{-1} P$ , one concludes that  $\text{rank}(F^T M^{-1} N) = m$  holds if  $\text{rank}(F) = m$  and  $\sigma_{\max}(-F^T M^{-1} P) < \sigma_{\min}(F^T M^{-1} F)$ .

It is noteworthy that the study of problem (B.3) may also be quite useful in the context of numerical analysis of differential algebraic equations (DAEs). Half-explicit methods involve such problems (for instance  $N$  may be the jacobian of the constraints estimated at step  $i+1$  while  $F$  is the jacobian estimated at step  $i$ ) [29, §7.1] [30, §VII.6], see also [126, 127, 128, 129] for various forms of numerical KKT systems.

# Appendix C

## Lie Groups

### C.1 Differential Geometry of $SO(3)$ and $SE(3)$

In this Appendix we give some basic definitions and concepts concerning the Lie groups  $SO(3)$  and  $SE(3)$ . The source of this material is the textbook [99].

**Definition of  $SO(3)$ :** The **special orthogonal group**  $SO(3)$  is the set of rotations of 3D Euclidean space. It can be defined as the subset of 3 by 3 matrices

$$SO(3) = \{R \in \mathbb{R}^{3 \times 3} : R^T R = I \text{ and } \det(R) = 1\}.$$

**The set  $SO(3)$  is a Lie group:** it is a group, it is a smooth manifold and the operations of product and inversion are smooth.

**The set  $\mathfrak{so}(3)$**  is defined as the tangent space of  $SO(3)$  at the identity,  $T_I SO(3)$ , which happens to be the set of skew symmetric 3 by 3 matrices:  $\mathfrak{so}(3) = \{A \in \mathbb{R}^{3 \times 3} : A^T = -A\}$ .

**The hat map** realizes an isomorphism between  $\mathfrak{so}(3)$  and  $\mathbb{R}^3$  by mapping vectors  $u = (u_1, u_2, u_3)^T$  to skew symmetric matrices as  $\hat{u} = \begin{bmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{bmatrix}$ . The hat map has an important relationship with the cross product: for any vector  $v$  in  $\mathbb{R}^3$  one has  $\hat{u}v = u \times v$ .

**The Lie algebra structure of  $\mathfrak{so}(3)$ .** The set  $\mathfrak{so}(3)$  is naturally a vector space, what makes it a 'Lie algebra' is that it is stable by the 'Lie bracket' operation: for any two skew symmetric matrices  $A$  and  $B$  of  $\mathfrak{so}(3)$ , the bracket  $[A, B]$ , defined as  $[A, B] := AB - BA$ , is also a skew symmetric matrix,  $[A, B] \in \mathfrak{so}(3)$ . The hat map is a Lie algebra homomorphism, it allows us to identify  $\mathfrak{so}(3)$  with  $\mathbb{R}^3$  as  $(\mathfrak{so}(3), [,]) \cong (\mathbb{R}^3, \times)$ .

**Left and Right representations.** Any rotation velocity  $\dot{R} \in T_R SO(3)$  can be seen as the left or right translation of an element of  $\mathfrak{so}(3)$ . That is, there exist vectors  $\kappa$  and  $\Omega$  in  $\mathbb{R}^3$  such that  $\dot{R} = R\hat{\kappa} = \hat{\Omega}R$ . For this reason the elements of  $\mathfrak{so}(3)$  are often called 'instantaneous rotations'. The corresponding vector  $\kappa$  is an instantaneous rotation vector in 'left reduced' or 'body' coordinates, while the corresponding vector  $\Omega$  represents the same instantaneous rotation but in 'right reduced' or 'spatial' coordinates. The spatial angular velocity  $\Omega$  and the body angular velocity  $\kappa$  are related by  $\Omega = R\kappa$ .

**Rodrigues Formula.** Let  $\kappa$  be a vector in  $\mathbb{R}^3$ . The exponential of the skew symmetric matrix  $\hat{\kappa} \in \mathfrak{so}(3)$  is the rotation matrix of axis  $\kappa/\|\kappa\|$  and of angle  $\|\kappa\|$ , it has a closed form expression given by the Rodrigues formula

$$\exp \hat{\kappa} = \begin{cases} I + \frac{\sin \|\kappa\|}{\|\kappa\|} \hat{\kappa} + \frac{1-\cos \|\kappa\|}{\|\kappa\|^2} \hat{\kappa}^2 & \text{if } \kappa \neq 0 \\ I & \text{if } \kappa = 0. \end{cases} \quad (\text{C.1})$$

**Definition of SE(3).** The **Special Euclidean Group SE(3)** is the set of rigid, orientation preserving, motions of Euclidean 3 dimensional space. It can be defined as the subset of 4 by 4 matrices  $SE(3) = \left\{ H \in \mathbb{R}^{4 \times 4} : H = \begin{bmatrix} R & r \\ 0 & 1 \end{bmatrix} \text{ where } R \in SO(3) \text{ and } r \in \mathbb{R}^3 \right\}$ . Elements of  $SE(3)$  will often be denoted  $H$ , alluding to the homogeneous matrix representation, or  $(R, r)$  alluding to the rotation and translation parts of the rigid motion.

**The set  $SE(3)$  is a Lie group.** It is a smooth manifold, it is a group and has smooth product and inverse operations. Note that the inverse of  $(R, r)$  is given by  $(R^T, -R^T r)$ .

**The Lie algebra  $\mathfrak{se}(3)$**  is the tangent space of  $SE(3)$  at the identity,  $T_I SE(3)$ , it is set of matrices

$$\mathfrak{se}(3) = \left\{ \xi \in \mathbb{R}^{4 \times 4} : \xi = \begin{bmatrix} \hat{\kappa} & v \\ 0 & 0 \end{bmatrix} \text{ where } \kappa \in \mathbb{R}^3 \text{ and } v \in \mathbb{R}^3 \right\}.$$

Elements of  $\mathfrak{se}(3)$  will often be denoted  $\xi$ , referring to the 4 by 4 matrix representation, or  $(\hat{\kappa}, v)$  alluding to the instantaneous rotation and translation parts. The Lie bracket of two elements  $\xi$  and  $\eta$  of  $\mathfrak{se}(3)$  is given by  $[\xi, \eta] = \xi\eta - \eta\xi$ .

**The tilde map** realizes an isomorphism between  $\mathbb{R}^6$  and  $\mathfrak{se}(3)$ . For a vector  $X = (\kappa, v) \in \mathbb{R}^6$  we denote  $\tilde{X}$  the  $\mathfrak{se}(3)$  element  $\tilde{X} = \begin{bmatrix} \hat{\kappa} & v \\ 0 & 0 \end{bmatrix}$ .

**The adjoint action of the Lie algebra  $\mathfrak{se}(3)$  on itself**  $\text{ad} : \mathfrak{se}(3) \times \mathfrak{se}(3) \rightarrow \mathfrak{se}(3)$  is just the Lie bracket  $\text{ad}_\xi \eta = [\xi, \eta] = \xi\eta - \eta\xi$ . Expanding in terms of rotation and translation parts  $\xi = (\hat{\kappa}, v)$  and  $\eta = (\hat{\omega}, u)$ , it reads

$$\text{ad}_{(\hat{\kappa}, v)}(\hat{\omega}, u) = [(\hat{\kappa}, v), (\hat{\omega}, u)] = ([\hat{\kappa}, \hat{\omega}]_{\mathfrak{so}(3)}, \hat{\kappa}u - \hat{\omega}v). \quad (\text{C.2})$$

The Lie algebra  $\mathfrak{se}(3)$  can be identified with  $\mathbb{R}^6$ . In that case the bracket of  $\mathbb{R}^6$  is given by

$$\text{ad}_{(\kappa, v)}(\omega, u) = (\kappa \times \omega, \kappa \times u - \omega \times v).$$

**The Lie coalgebra  $\mathfrak{se}(3)^*$**  is defined as the dual vector space of  $\mathfrak{se}(3)$ . In terms of mechanics the dual quantities of linear velocities  $v$  are forces  $f$  and the dual quantities of angular velocities  $\kappa$  are torques  $m$ . The duality is then just interpreted as the total work  $\langle (m, f), (\kappa, v) \rangle$ . Elements of  $\mathfrak{se}(3)^*$  will often be denoted  $\mu = (m, f)$ .

**The coadjoint action of  $\mathfrak{se}(3)$  on  $\mathfrak{se}(3)^*$**  is defined as the map  $\text{ad}^* : \mathfrak{se}(3) \times \mathfrak{se}(3)^* \rightarrow \mathfrak{se}(3)^*$  such that

$$\text{ad}_{(\hat{\kappa}, v)}^*(\hat{m}, f) = ((-\kappa \times m - v \times f)^\wedge, -\kappa \times f). \quad (\text{C.3})$$

The linear map  $\text{ad}_\xi^*$  is the dual of  $\text{ad}_\xi$ .

**The Exponential map of any element  $(\hat{\kappa}, v)$  in  $\mathfrak{se}(3)$**  is the rigid motion of rotation  $\exp \hat{\kappa}$  and translation  $Bv$  given by

$$\exp \begin{bmatrix} \hat{\kappa} & v \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \exp \hat{\kappa} & Bv \\ 0 & 1 \end{bmatrix},$$

$$\text{where } B = \begin{cases} I + \frac{1-\cos||\kappa||}{||\kappa||^2} \hat{\kappa} + \frac{||\kappa|| - \sin||\kappa||}{||\kappa||^3} \hat{\kappa}^2 & \text{if } \kappa \neq 0 \\ I & \text{otherwise.} \end{cases}$$

# Bibliography

- [1] F. Bertails, B. Audoly, M.-P. Cani, B. Querleux, F. Leroy, and J.-L. Lévéque, “Super-helices for predicting the dynamics of natural hair,” in *ACM Transactions on Graphics (Proceedings of the SIGGRAPH conference)*, August 2006.
- [2] A. Blumentals, B. Brogliato, and F. Bertails-Descoubes, “The Contact Problem in Lagrangian Systems subject to Bilateral and Unilateral constraints with sliding Coulomb’s Friction,” in *ECCOMAS Thematic Conference on Multibody Dynamics*, (Barcelone, Spain), June 2015.
- [3] A. Blumentals, B. Brogliato, and F. Bertails-Descoubes, “The contact problem in lagrangian systems subject to bilateral and unilateral constraints, with or without sliding coulomb’s friction: a tutorial,” *Multibody System Dynamics*, vol. 38, no. 1, pp. 43–76, 2016.
- [4] A. Blumentals, F. Bertails-Descoubes, and R. Casati, “Dynamics of a developable shell with uniform curvatures,” in *The 4th Joint International Conference on Multibody System Dynamics*, (Montréal, Canada), May 2016.
- [5] R. Leine and N. van de Wouw, *Stability and Convergence of Mechanical Systems with Unilateral Constraints*, vol. 36 of *Lecture Notes in Applied and Computational Mechanics*. Springer, 2008.
- [6] V. Acary and B. Brogliato, *Numerical Methods for Nonsmooth Dynamical Systems*, vol. 35 of *Lecture Notes in Applied and Computational Mechanics*. Springer Verlag, 2008.
- [7] F. Génot and B. Brogliato, “New results on Painlevé paradoxes,” *European Journal of Mechanics - A/Solids*, vol. 18, pp. 653–677, July 1999.
- [8] P. Painlevé, *Leçons sur le Frottement*. A. Hermann, 1895.
- [9] L. X. Anh, *Dynamics of Mechanical Systems with Coulomb Friction*. Foundations of Mechanical Engineering, Springer Verlag, 2003.
- [10] R. Leine, B. Brogliato, and H. Nijmeijer, “Periodic motion and bifurcations induced by the Painlevé paradox,” *European Journal of Mechanics - A/Solids*, vol. 21, no. 5, pp. 869 – 896, 2002.
- [11] J. Pang and J. Trinkle, “Complementarity formulation and existence of solutions of dynamic rigid-body contact problems with Coulomb friction,” *Mathematical Programming*, vol. 73, pp. 199–226, May 1996.

- [12] J. G. de Jalon and M. Gutierrez-Lopez, “Multibody dynamics with redundant constraints and singular mass matrix: existence, uniqueness, and determination of solutions for accelerations and constraints forces,” *Multibody System Dynamics*, vol. 30, no. 3, pp. 311–341, 2013.
- [13] F. Udwadia and P. Phohomsiri, “Explicit equations of motion for constrained mechanical systems with singular mass matrices and applications to multi-body dynamics,” *Proc. Royal Soc. London A*, vol. 462, pp. 2097–2117, 2006.
- [14] E. Bayo and R. Ledesma, “Augmented lagrangian and mass-orthogonal projection methods for constrained multibody dynamics,” *Nonlinear Dynamics*, vol. 9, no. 1-2, pp. 113–130, 1996.
- [15] W. Blajer, “Augmented lagrangian formulation: geometrical interpretation and application to systems with singularities and redundancy,” *Multibody System Dynamics*, vol. 8, no. 2, pp. 141–159, 2002.
- [16] B. Ruzzeh and J. Kövecses, “A penalty formulation for dynamics analysis of redundant mechanical systems,” *Journal of Computational and Nonlinear Dynamics*, vol. 6, no. 2, p. 021008, 2011.
- [17] D. Dopico, F. González, J. Cuadrado, and J. Kövecses, “Determination of holonomic and nonholonomic constraint reactions in an index-3 augmented lagrangian formulation with velocity and acceleration projections,” *Journal of Computational and Nonlinear Dynamics*, vol. 9, no. 4, p. 041006, 2014.
- [18] M. Wojtyra and J. Fraczek, “Solvability of reactions in rigid multibody systems with redundant nonholonomic constraints,” *Multibody System Dynamics*, vol. 30, no. 2, pp. 153–171, 2013.
- [19] M. Wojtyra, “Joint reactions in rigid body mechanisms with dependent constraints,” *Mechanism and Machine Theory*, vol. 44, no. 12, pp. 2265 – 2278, 2009.
- [20] J. Fraczek and M. Wojtyra, “On the unique solvability of a direct dynamics problem for mechanisms with redundant constraints and coulomb friction in joints,” *Mechanism and Machine Theory*, vol. 46, no. 3, pp. 312 – 334, 2011.
- [21] F. Udwadia and R. Kalaba, *Analytical Dynamics: A New Approach*. Cambridge University Press, 1996.
- [22] V. M. Matrosov and I. Finogenko, “The theory of differential equations which arise in the dynamics of a system of rigid bodies with coulomb friction,” *Monograph of Academy of Nonlinear Sciences, Advances on Nonlinear Science*, vol. 2, pp. 16–106, 2008.
- [23] V. Matrosov and I. Finogenko, “Right-hand solutions of the differential equations of dynamics for mechanical systems with sliding friction,” *Journal of Applied Mathematics and Mechanics*, vol. 59, no. 6, pp. 837–844, 1995.
- [24] P. Lötstedt, “Coulomb friction in two-dimensional rigid body systems,” *Zeitschrift für Angewandte Mathematik and Mechanik*, vol. 61, pp. 605–615, 1981.
- [25] B. Brogliato and D. Goeleven, “Singular mass matrix and redundant constraints in unilaterally constrained Lagrangian and Hamiltonian systems,” *Multibody System Dynamics*, vol. 35, no. 1, pp. 39–61, 2015. DOI:10.1007/s11044-014-9437-4.

- [26] A. Ivanov, “Singularities in the dynamics of systems with non-ideal constraints,” *Journal of Applied Mathematics and Mechanics*, vol. 67, no. 2, pp. 185–192, 2003.
- [27] C. Glocker, *Set-Valued Force Laws: Dynamics of Non-Smooth Systems*. Springer Verlag, 2001.
- [28] D. Bernstein, *Matrix, Mathematics. Theory, Facts, and Formulas with Application to Linear Systems Theory*. Princeton University press, 2005.
- [29] B. Simeon, *Computational Flexible Multibody Dynamics. A Differential-Algebraic Approach*. Springer Verlag, 2013. Differential-Algebraic Equations Forum.
- [30] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems*, vol. 14 of *Series in Computational Mathematics*. Berlin Heidelberg: Springer Verlag, second revised ed., 1996.
- [31] A. Laulusa and O. Bauchau, “Review of classical approaches for constraint enforcement in multibody systems,” *ASME J. of Comput. and Nonlinear Dynamics*, vol. 3, no. 1, p. 011004, 2008. doi:10.1115/1.2803257.
- [32] Z. Terze, A. Müller, and D. Zlatar, “Lie-group integration method for constrained multibody systems in state space,” *Multibody System Dynamics*, 2015. DOI: 10.1007/s11044-014-9439-2.
- [33] M. Wojtyra and J. Fraczek, “Comparison of selected methods of handling redundant constraints in multibody systems simulations,” *ASME Journal of Computational and Nonlinear Dynamics*, vol. 8, p. 0210007, 2013.
- [34] L. V. S. Boyd, *Convex Optimization*. Cambridge University Press, 2004.
- [35] A. Shabana, “Euler parameters kinetic singularity,” *Proc. Inst. Mech. Eng. Part K: Journal of Multi-body Dynamics*, 2014. DOI: 10.1177/1464419314539301.
- [36] F. Udwadia and A. Schutte, “Equations of motion for general constrained systems in Lagrangian mechanics,” *Acta Mechanica*, vol. 213, no. 1-2, pp. 111–129, 2010.
- [37] J. Moreau, “Les liaisons unilatérales et le principe de Gauss,” *Comptes Rendus Académie des Sciences*, pp. 871–874, 1963. 1er semestre, T. 256, no 4.
- [38] J. Moreau, “Quadratic programming in mechanics: dynamics of one-sided constraints,” *SIAM Journal on control*, vol. 4, no. 1, pp. 153–158, 1966.
- [39] V. Acary and F. Pérignon, “Siconos: A software platform for modeling, simulation, analysis and control of nonsmooth dynamical systems,” *Simulation News Europe*, vol. 17, no. 3/4, pp. 19–26, 2007.
- [40] R. Cottle, J. Pang, and R. Stone, *The Linear Complementarity Problem*. Computer Science and Scientific Computing, Academic Press, 1992.
- [41] P. Löstedt, “Mechanical systems of rigid bodies subject to unilateral constraints,” *SIAM Journal on Applied Mathematics*, vol. 42, no. 2, pp. 281–296, 1982.
- [42] J. Pang, J. Trinkle, and G. Lo, “A complementarity approach to a quasistatic multi-rigid-body contact problem,” *Journal of Computational Optimization and Applications*, vol. 5, pp. 139–154, Mar. 1996.

- [43] B. Brogliato, “Inertial couplings between unilateral and bilateral holonomic constraints in frictionless Lagrangian systems,” *Multibody System Dynamics*, vol. 29, pp. 289–325, Mar. 2013.
- [44] F. Pfeiffer, “On non-smooth multibody dynamics,” *Proc. Inst.Mech. Eng., Part K: Journal of Multi-body Dynamics*, vol. 226, no. 2, pp. 147–177, 2014.
- [45] M. Hjiaj, G. de Saxcé, and Z. Mroz, “A variational inequality-based formulation of the frictional contact law with a non-associated sliding rule,” *European Journal of Mechanics A/Solids*, vol. 21, pp. 49–59, 2002.
- [46] G. de Saxcé and Z. Feng, “New inequality and functional for contact friction: the implicit standard material approach,” *Mech. Struct. Mach.*, vol. 19, pp. 301–325, 1991.
- [47] G. E. Coxson, “The p-matrix problem is co-np-complete,” *Mathematical Programming*, vol. 64, no. 1, pp. 173–178, 1994.
- [48] E. Jones, T. Oliphant, P. Peterson, *et al.*, “SciPy: Open source scientific tools for Python,” 2001–. [Online; accessed 2016-05-04].
- [49] J. Moreau, “Unilateral contact and dry friction in finite freedom dynamics,” *Nonsmooth mechanics and Applications*, 1988. CISM Courses Lect. 302, 1-82,.
- [50] M. Haddouni, *Algorithmes de résolution de la dynamique du contact avec impact et frottement*. PhD thesis, 2015. Thèse de doctorat dirigée par Brogliato, Bernard Matériaux, mécanique, génie civil, électrochimie Grenoble Alpes 2015.
- [51] D. Stewart, “Convergence of a time-stepping scheme for rigid-body dynamics and resolution of painlevé’s problem,” *Archives of Rational Mechanics Anal.*, vol. 145, no. 3, pp. 215–260, 1998.
- [52] V. Acary, F. Cadoux, C. Lemaréchal, and J. Malick, “A formulation of the linear discrete Coulomb friction problem via convex optimization,” *Zeitschrift für angewandte Mathematik und Mechanik*, vol. 91, pp. 155–175, Feb. 2011.
- [53] P. Alart, “Contributions and limitations of the Non Smooth Contact Dynamics for the simulation of dense granular systems,” in *PARTICLES 2015 Fourth International Conference on Particle-Based Methods*, (Barcelone, Spain), Sept. 2015.
- [54] R. Casati and F. Bertails-Descoubes, “Super space clothoids,” *ACM Trans. Graph.*, vol. 32, pp. 48:1–48:12, July 2013.
- [55] E. Lauga and T. R. Powers, “The hydrodynamics of swimming microorganisms,” *Reports on Progress in Physics*, vol. 72, no. 9, p. 096601, 2009.
- [56] C. J. Benham and S. P. Mielke, “Dna mechanics,” *Annual Review of Biomedical Engineering*, vol. 7, no. 1, pp. 21–53, 2005. PMID: 16004565.
- [57] J. Linn, “Discrete kinematics of Cosserat rods based on the difference geometry of framed curves,” in *The 4th Joint International Conference on Multibody System Dynamics, Montréal, Canada*, 2016.
- [58] S. Goyal, N. Perkins, and C. L. Lee, “Non-linear dynamic intertwining of rods with self-contact,” *International Journal of Non-Linear Mechanics*, vol. 43, no. 1, pp. 65 – 73, 2008.

- [59] D. H. Hodges, “Geometrically exact, intrinsic theory for dynamics of curved and twisted anisotropic beams,” *AIAA journal*, vol. 41, no. 6, pp. 1131–1137, 2003.
- [60] F. Demoures, F. Gay-Balmaz, M. Kobilarov, and T. S. Ratiu, “Multisymplectic lie group variational integrator for a geometrically exact beam in,” *Communications in Nonlinear Science and Numerical Simulation*, vol. 19, no. 10, pp. 3492 – 3512, 2014.
- [61] G. Daviet, F. Bertails-Descoubes, and L. Boissieux, “A hybrid iterative solver for robustly capturing coulomb friction in hair dynamics,” *ACM Trans. Graph.*, vol. 30, pp. 139:1–139:12, Dec. 2011.
- [62] N. Bou-Rabee and J. E. Marsden, “Hamilton–Pontryagin integrators on Lie groups part i: Introduction and structure-preserving properties,” *Foundations of Computational Mathematics*, vol. 9, no. 2, pp. 197–219, 2009.
- [63] D. H. Hodges, “A mixed variational formulation based on exact intrinsic equations for dynamics of moving beams,” *International Journal of Solids and Structures*, vol. 26, no. 11, pp. 1253 – 1273, 1990.
- [64] M. Borri and C. Bottasso, “An intrinsic beam model based on a helicoidal approximation—part i: Formulation,” *International Journal for Numerical Methods in Engineering*, vol. 37, no. 13, pp. 2267–2289, 1994.
- [65] H. Lang, H. Laube, and S. Leyendecker, “Various multibody dynamic models for the description of plane kirchhoff rods,” *ECCOMAS Thematic Conference on Multibody Dynamics June 29 - July 2, 2015, Barcelona, Catalonia, Spain*, 2015.
- [66] V. Jurdjevic, *Geometric Control Theory*. Cambridge Studies in Advanced Mathematics, Cambridge University Press, 1996.
- [67] S. Kehrbaum and J. H. Maddocks, “Elastic rods, rigid bodies, quaternions and the last quadrature,” *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 355, no. 1732, pp. 2117–2136, 1997.
- [68] J. Andersson, *A General-Purpose Software Framework for Dynamic Optimization*. PhD thesis, Arenberg Doctoral School, KU Leuven, Department of Electrical Engineering (ESAT/SCD) and Optimization in Engineering Center, Kasteelpark Arenberg 10, 3001-Heverlee, Belgium, October 2013.
- [69] V. Sonnevile, A. Cardona, and O. Brüls, “Geometrically exact beam finite element formulated on the special Euclidean group SE(3),” *Computer Methods in Applied Mechanics and Engineering*, vol. 268, pp. 451 – 474, 2014.
- [70] K.-J. Bathe, *Finite Element Procedures*. Klaus-Jurgen Bathe, 2006.
- [71] V. Sonnevile, A. Cardona, and O. Brüls, “Geometric interpretation of a non-linear beam finite element on the Lie group SE(3),” *Archive of Mechanical Engineering*, vol. 61, no. 2, pp. 305–329, 2014.
- [72] P. Jung, S. Leyendecker, J. Linn, and M. Ortiz, “A discrete mechanics approach to the cosserat rod theory—part 1: static equilibria,” *International Journal for Numerical Methods in Engineering*, vol. 85, no. 1, pp. 31–60, 2011.
- [73] O. Sander, “Geodesic finite elements for Cosserat rods,” *International Journal for Numerical Methods in Engineering*, vol. 82, no. 13, pp. 1645–1670, 2010.

- [74] R. Levien, “The elastica: a mathematical history,” *University of California, Berkeley, Technical Report No. UCB/EECS-2008-103*, 2008.
- [75] F. Bertails-Descoubes, “Super-clothoids,” *Computer Graphics Forum*, vol. 31, no. 2, pp. 509–518, 2012.
- [76] L. D. Landau, E. M. Lifsic, L. Pitaevskii, and A. Kosevich, *Course of Theoretical Physics: Volume 7, Theory of Elasticity*. Pergamon Press, 1986.
- [77] B. Audoly and Y. Pomeau, *Elasticity and Geometry: from Hair Curls to the Nonlinear Response of Shells*. Oxford University Press, 2010.
- [78] I. M. Gelfand and S. V. Fomin, *Calculus of Variations: Ed.: Richard A. Silverman*. Prentice-Hall, 1963.
- [79] D. A. Singer, O. J. Garay, E. García-Río, and R. Vázquez-Lorenzo, “Lectures on elastic curves and rods,” in *AIP Conference Proceedings*, vol. 1002, pp. 3–32, AIP, 2008.
- [80] E. Reissner, “On one-dimensional finite-strain beam theory: The plane problem,” *Zeitschrift für angewandte Mathematik und Physik ZAMP*, vol. 23, no. 5, pp. 795–804, 1972.
- [81] S. Antman, *Nonlinear Problems of Elasticity; 2nd ed.* Dordrecht: Springer, 2005.
- [82] E. Cosserat, F. Cosserat, M. Brocato, and K. Chatzis, “Théorie des corps déformables,” 1909.
- [83] G. Kirchhoff, “Über das Gleichgewicht und die Bewegung eines unendlich dünnen elastischen Stabes.,” *Journal für die reine und angewandte Mathematik*, vol. 56, pp. 285–313, 1859.
- [84] E. H. Dill, “Kirchhoff’s theory of rods,” *Archive for History of Exact Sciences*, vol. 44, no. 1, pp. 1–23, 1992.
- [85] H. Lang and M. Arnold, “Numerical aspects in the dynamic simulation of geometrically exact rods,” *Applied Numerical Mathematics*, vol. 62, no. 10, pp. 1411–1427, 2012.
- [86] M. P. Do Carmo, *Differential Geometry of Curves and Surfaces*. Prentice-Hall Englewood Cliffs, 1976.
- [87] R. L. Bishop, “There is more than one way to frame a curve,” *The American Mathematical Monthly*, vol. 82, no. 3, pp. 246–251, 1975.
- [88] J. Langer and D. A. Singer, “Lagrangian aspects of the Kirchhoff elastic rod,” *SIAM review*, vol. 38, no. 4, pp. 605–618, 1996.
- [89] M. Bergou, M. Wardetzky, S. Robinson, B. Audoly, and E. Grinspun, “Discrete elastic rods,” in *ACM Transactions on Graphics (TOG)*, no. 3, p. 63, ACM, 2008.
- [90] I. A. Bobenko and B. Y. Suris, “Discrete Time Lagrangian Mechanics on Lie Groups, with an Application to the Lagrange Top,” *Communications in Mathematical Physics*, vol. 204, no. 1, pp. 147–188, 1999.
- [91] M. F. Dixon, *Geometric Integrators for Continuum Dynamics*. PhD thesis, Imperial College London (University of London), 2007.

- [92] D. Zupan and M. Saje, “Finite-element formulation of geometrically exact three-dimensional beam theories based on interpolation of strain measures,” *Computer Methods in Applied Mechanics and Engineering*, vol. 192, no. 49–50, pp. 5209 – 5248, 2003.
- [93] D. Zupan and M. Saje, “The three-dimensional beam theory: Finite element formulation based on curvature,” *Computers & Structures*, vol. 81, no. 18–19, pp. 1875 – 1888, 2003. Civil Comp Special Issue.
- [94] P. Češarek, M. Saje, and D. Zupan, “Dynamics of flexible beams: Finite-element formulation based on interpolation of strain measures,” *Finite Elements in Analysis and Design*, vol. 72, pp. 47 – 63, 2013.
- [95] E. Reissner, “On One-Dimensional Large-Displacement Finite-Strain Beam Theory,” *Studies in Applied Mathematics*, vol. 52, no. 2, pp. 87–95, 1973.
- [96] M. A. Crisfield and G. Jelenić, “Objectivity of strain measures in the geometrically exact three-dimensional beam theory and its finite-element implementation,” in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 455, pp. 1125–1147, The Royal Society, 1999.
- [97] L. T. Biegler and V. M. Zavala, “Large-scale nonlinear programming using ipopt: An integrating framework for enterprise-wide dynamic optimization,” *Computers & Chemical Engineering*, vol. 33, no. 3, pp. 575–582, 2009.
- [98] J. M. Sanz-Serna, “Symplectic Runge-Kutta Schemes for Adjoint Equations, Automatic Differentiation, Optimal Control, and More,” *SIAM Review*, vol. 58, no. 1, pp. 3–33, 2016.
- [99] D. Holm, T. Schmah, and C. Stoica, *Geometric Mechanics and Symmetry, vol. 12 of Oxford Texts in Applied and Engineering Mathematics*. Oxford University Press, Oxford, 2009.
- [100] W. W. Hager, “Runge-Kutta methods in optimal control and the transformed adjoint system,” *Numerische Mathematik*, vol. 87, no. 2, pp. 247–282, 2000.
- [101] J. Bonnans, Frederic, D. Giorgi, V. Grelard, B. Heymann, S. Maindrault, P. Martinon, O. Tissot, and J. Liu, “Bocop – A collection of examples,” tech. rep., INRIA, 2017.
- [102] J. Stuelpnagel, “On the parametrization of the three-dimensional rotation group,” *SIAM Review*, vol. 6, no. 4, pp. 422–430, 1964.
- [103] H. Munthe-Kaas, “Runge-Kutta methods on Lie groups,” *BIT Numerical Mathematics*, vol. 38, no. 1, pp. 92–111, 1998.
- [104] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*; 2nd ed. Dordrecht: Springer, 2006.
- [105] E. Cerda and L. Mahadevan, “Confined developable elastic surfaces: cylinders, cones and the elastica,” in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 461, pp. 671–700, The Royal Society, 2005.
- [106] H. Ferreau, C. Kirches, A. Potschka, H. Bock, and M. Diehl, “qpOASES: A parametric active-set algorithm for quadratic programming,” *Mathematical Programming Computation*, vol. 6, no. 4, pp. 327–363, 2014.

- [107] M. Dias and B. Audoly, ““wunderlich, meet kirchhoff”: A general and unified description of elastic ribbons and thin rods,” *Journal of Elasticity*, pp. 1–18, 2014.
- [108] R. Casati, *Quelques contributions à la modélisation numérique de structures élancées pour l'informatique graphique*. PhD thesis, 2015. Thèse de doctorat dirigée par Brogliato, Bernard et Bertails, Florence Mathématiques et informatique Grenoble Alpes 2015.
- [109] P. G. Ciarlet, “An Introduction to Differential Geometry with Applications to Elasticity,” *Journal of Elasticity*, vol. 78, no. 1-3, pp. 1–215, 2005.
- [110] E. R. Phillips, “Karl M. Peterson: The earliest derivation of the Mainardi-Codazzi equations and the fundamental theorem of surface theory,” *Historia Mathematica*, vol. 6, no. 2, pp. 137 – 163, 1979.
- [111] R. S. Palais, *A modern course on curves and surfaces*. Virtual Math Museum, 2003.
- [112] W.-Z. Chien, “The intrinsic theory of thin shells and plates. I. general theory,” *Quarterly of Applied Mathematics*, vol. 1, no. 4, pp. 297–327, 1944.
- [113] D. Danielson, “Simplified intrinsic equations for arbitrary elastic shells,” *International Journal of Engineering Science*, vol. 8, no. 3, pp. 251–259, 1970.
- [114] W. T. Koiter and J. G. Simmonds, *Foundations of shell theory*, pp. 150–176. Springer, 1973.
- [115] S. Opoka and W. Pietraszkiewicz, “Intrinsic equations for non-linear deformation and stability of thin elastic shells,” *International Journal of Solids and Structures*, vol. 41, no. 11–12, pp. 3275 – 3292, 2004.
- [116] D. H. Hodges, W. Yu, and M. J. Patil, “Geometrically-exact, intrinsic theory for dynamics of moving composite plates,” *International Journal of Solids and Structures*, vol. 46, no. 10, pp. 2036 – 2042, 2009. Special Issue in Honor of Professor Liviu Librescu.
- [117] J. C. Simo and D. D. Fox, “On stress resultant geometrically exact shell model. part i: Formulation and optimal parametrization,” *Comput. Methods Appl. Mech. Eng.*, vol. 72, pp. 267–304, Mar. 1989.
- [118] V. Sonneville, *A geometric local frame approach for flexible multibody systems*. PhD thesis, Université de Liège, Liège, Belgique, 2015.
- [119] G. Gede, D. L. Peterson, A. S. Nanjungud, J. K. Moore, and M. Hubbard, “Constrained Multibody Dynamics with Python: From symbolic equation generation to publication,” in *ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, vol. 7B, 2013.
- [120] P. Alart, “Dynamique non régulière multicontacte -quelques questions liées à la simulation des milieux granulaires denses,” in *CSMA 2015*, CSMA 2015, (Giens, France), May 2015.
- [121] E. Sharon and E. Efrati, “The mechanics of non-euclidean plates,” *Soft Matter*, vol. 6, pp. 5693–5704, 2010.
- [122] F. Facchinei and J. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems, vol. I*. Operations Research, New York: Springer Verlag, 2003.

- [123] X. Chen and S. Xiang, “Perturbation bounds of P-matrix linear complementarity problems,” *SIAM Journal on Optimization*, vol. 18, no. 4, pp. 1250–1265, 2007.
- [124] B. Brogliato, “Kinetic quasi-velocities in unilaterally constrained lagrangian mechanics with impacts and friction,” *Multibody System Dynamics*, vol. 32, pp. 175–216, 2014.
- [125] J. G. de Jalon, A. Callejo, and A. Hidalgo, “Efficient solution of Maggi’s equations,” *ASME J. of Computational and Nonlinear Dynamics*, vol. 7, p. 021003, April 2012.
- [126] O. Brüls and M. Arnold, “Convergence of the generalized  $\alpha$  scheme for constrained mechanical systems,” *Multibody System Dynamics*, vol. 18, no. 2, pp. 185–202, 2007.
- [127] C. Lunk and B. Simeon, “Solving constrained mechanical systems by the family of Newmark and  $\alpha$ -methods.,” *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*,, vol. 86, no. 10, pp. 772–784, 2006.
- [128] A. Murua, “Partitioned half-explicit Runge-Kutta methods for differential-algebraic systems of index 2,” *Computing*, vol. 59, no. 1, pp. 43–61, 1997.
- [129] D. Negrut, L. Jay, and N. Khude, “A discussion of low-order numerical integration formulas for rigid and flexible multibody dynamics,” *ASME J. of Computational and Nonlinear Dynamics*, vol. 4, p. 021008, 2009.

**Abstract.** This dissertation focuses on the numerical modelling of thin elastic structures in contact. Many objects around us, either natural or man-made, are slender deformable objects. Curve-like objects such as industrial cables, helicopter blades, plant stems and hair can be modelled as thin elastic rods. While surface-like objects such as paper, boat sails, leaves and clothes can be modelled as thin elastic shells. The numerical study of the mechanical response of such structures is important in many applications of engineering, bio-mechanics, computer graphics and other fields. In this dissertation we treat rods and shells as finite dimensional multibody systems.

When a multibody system is subject to frictional contact constraints, a problem often arises. In some configurations there may exist no contact force which can prevent the system from violating its contact constraints. This is known as the Painlevé paradox. In the first part of this manuscript we analyze the contact problem (whose unknowns are the accelerations and the contact forces) and we derive computable upper bounds on the friction coefficients at each contact, such that if verified, the contact problem is well-posed and Painlevé paradoxes are avoided.

Some rod-like structures may easily bend and twist but hardly stretch and shear, such structures can be modelled as Kirchhoff rods. In the second part of this manuscript we consider the problem of computing the stable static equilibria of Kirchhoff rods subject to different boundary conditions and frictionless contact constraints. We formulate the problem as an Optimal Control Problem, where the strains of the rod are interpreted as control variables and the position and orientation of the rod are interpreted as state variables. Employing direct methods of numerical Optimal Control then leads us to the proposal of new spatial discretization schemes for Kirchhoff rods. The proposed schemes are either of the strain-based type, where the main degrees of freedom are the strains of the rod, or of the mixed type, where the main degrees of freedom are both the strains and the generalized displacements.

Very much like for Kirchhoff rods, thin surface-like structures such as paper can hardly stretch or shear at all. One of the advantages of the strain based approach is that the no extension and no shear constraints of the Kirchhoff rod are handled intrinsically, without the need of stiff repulsion forces, or of further algebraic constraints on the degrees of freedom. In the third part of this dissertation we propose an extension of this approach to model the dynamics of inextensible and unshearable shells. We restrict our study to the case of a shell patch with a developable mid-surface. We use as primary degrees of freedom the components of the second fundamental form of the shell's mid-surface. This also leads to an intrinsic handling of the no shear and no extension constraints of the shell.

**Résumé.** Cette thèse porte sur la modélisation numérique des structures élastiques minces en contact. De nombreux objets autour de nous, naturels ou artificiels, sont des objets minces et déformables. Les objets filiformes tels que les câbles industriels, les pales d'hélicoptères, les tiges des plantes et les cheveux peuvent être modélisés comme des tiges élastiques minces. Alors que les objets surfaciques tels que le papier, les voiles de bateaux, les feuilles et les vêtements peuvent être modélisés comme des coques élastiques minces. L'étude numérique de la réponse mécanique de ces structures est de la plus grande importance dans de nombreuses applications de l'ingénierie, de la biomécanique, de l'infographie et de bien d'autres domaines. Dans cette thèse, nous traitons les tiges et les coques comme des systèmes multi-corps en dimension finie.

Lorsqu'un système multi-corps est soumis à des contraintes de contact frottant, un problème se pose souvent. Dans certaines configurations, il est à craindre qu'il n'existe aucune force de contact et aucune accélération qui puisse empêcher le système de violer ses contraintes. Ce phénomène est connu sous le nom de Paradoxe de Painlevé. Dans la première partie de ce manuscrit, nous analysons le problème de contact (dont les inconnues sont les accélérations et les forces de contact) et nous obtenons des bornes supérieures calculables sur les coefficients de frottement à chaque contact, de sorte que si elles sont vérifiées, le problème de contact est bien posé et les paradoxes de Painlevé sont évités.

Certaines structures filiformes peuvent facilement se courber et se tordre, alors qu'elles peuvent difficilement s'étirer ou cisailleur. De telles structures peuvent être modélisées comme des tiges de Kirchhoff. Dans la deuxième partie de ce manuscrit, nous considérons le problème du calcul des équilibres statiques stables des tiges de Kirchhoff soumises à des conditions de bord différentes et à des contraintes de contact sans frottement. Nous formulons le problème comme un problème de Commande Optimale, où les courbures de la tige sont interprétées comme des commandes et la position et l'orientation de la tige sont interprétées comme des variables d'état. L'utilisation de méthodes directes pour la Commande Optimale numérique nous conduit alors à la proposition de nouveaux schémas de discréttisation spatiale pour les tiges de Kirchhoff. Les schémas proposés sont soit du type intrinsèque, où les principaux degrés de liberté sont les courbures de la tige, soit du type mixte, où les principaux degrés de liberté sont à la fois les courbures et les déplacements généralisés.

Similairement aux tiges de Kirchhoff, certaines structures surfaciques telles que le papier peuvent difficilement s'allonger ou cisailleur. L'un des avantages de l'approche intrinsèque pour les tiges de Kirchhoff est que les contraintes de non élongation et de non cisaillement de la tige sont traitées intrinsèquement, sans faire appel à des forces de répulsion trop raides ou à d'autres contraintes algébriques sur les degrés de liberté. Dans la troisième partie de cette thèse, nous proposons une extension de cette approche pour modéliser la dynamique des coques inextensibles et sans cisaillement. Nous limitons notre étude au cas d'un élément de coque avec une surface moyenne développable. Nous utilisons comme degrés de liberté les composantes de la seconde forme fondamentale de la surface moyenne de la coque. Cela conduit également à une gestion intrinsèque des contraintes de non extension et de non cisaillement de la coque.