MODELING RIGID MULTI BODY DYNAMICS WITH CONTACT AND FRICTION

by

Mihai Anitescu

An Abstract

Of a thesis submitted in partial fulfillment of the
requirements for the Doctor of Philosophy
degree in Applied Mathematical and
Computational Sciences in the
Graduate College of The
University of Iowa

August 1997

Thesis supervisor: Professor Florian A. Potra

ABSTRACT

The field of multi-body dynamics simulation is expected to have a major impact on the design of complex mechanical systems. Finding realistic models for impact and friction is an important issue when dealing with such systems. Unfortunately, it is known since the late century that the acceleration setting of the differential algebraic equations involving friction does not always have a solution.

A new setting for the constrained Newton equations in presence of contact friction is proposed in which accelerations and forces are replaced by velocities and impulses. The new framework, based on a linear complementarity problem, mixes together the resolution of the velocities with the integration step and treats the contact and impact configurations in a unified manner. For all configurations, the new model is proven to have a solution that can be found by Lemke's algorithm. The model has several energy properties, also found in continuous-time systems, which make the velocities uniformly bounded as the time-step goes to zero.

A framework for integrating the geometrical information in the dynamical analysis is presented for various types of contacts. It is based on the generalized distance between two bodies, a mapping that can be defined usually only in the neighborhood of a contact configuration and provides an analytic description of the closest points on two bodies for near contact configurations. This information can be used in an integration procedure to avoid costly collision detection techniques in the moments following a contact breaking. Ways of describing, in a computationally efficient manner, configurations involving planar contact regions are also presented.

It is known that acceleration-based frictionless problems have always a solution. An important question, in light of the inconsistency of acceleration-based models for general friction configurations, is to determine which acceleration-based models have a solution for small values of the friction coefficient. A partial answer to this question is provided, by showing that disassemblable configurations, for which a force that will simultaneously break all the contacts exists, have the above property.

Abstract approved: ______________________________
Thesis supervisor

____
Title and department

____
Date

MODELING RIGID MULTI BODY DYNAMICS WITH CONTACT AND FRICTION

by

Mihai Anitescu

A thesis submitted in partial fulfillment of the
requirements for the Doctor of Philosophy
degree in Applied Mathematical and
Computational Sciences in the
Graduate College of The
University of Iowa

August 1997

Thesis supervisor: Professor Florian A. Potra

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Mihai Anitescu

has been approved by the Examining Committee for the thesis requirement for the Doctor of Philosophy degree in Applied Mathematical and Computational Sciences at the August 1997 graduation.

Thesis committee: ________________________
Thesis supervisor

____
Member

____
Member

____
Member

____
Member

To the memory of my father, Ilie Anitescu

ACKNOWLEDGEMENTS

My deepest gratitude to my advisor and Professor Florian A. Potra, for the encouragement, guidance and suggestions provided during the years spent at the University of Iowa. Many of the results in this thesis were initiated by the fruitful collaboration with Professor James F. Cremer. I am grateful to Professors Kendall E. Atkinson, Soura Dasgupta, Edward J. Haug, Joseph K. Kearney and Kenneth O. Kortanek for the discussions and suggestions that have helped shape my scientific formation.

I am grateful to my colleagues at the University of Iowa which have provided me with a challenging intellectual environment. Special thanks must go to my parents, for their love and lifelong support of my education and my wife, Magdalena, for the encouragement, motivation and love.

ABSTRACT

The field of multi-body dynamics simulation is expected to have a major impact on the design of complex mechanical systems. Finding realistic models for impact and friction is an important issue when dealing with such systems. Unfortunately, it is known since the late century that the acceleration setting of the differential algebraic equations involving friction does not always have a solution.

A new setting for the constrained Newton equations in presence of contact friction is proposed in which accelerations and forces are replaced by velocities and impulses. The new framework, based on a linear complementarity problem, mixes together the resolution of the velocities with the integration step and treats the contact and impact configurations in a unified manner. For all configurations, the new model is proven to have a solution that can be found by Lemke's algorithm. The model has several energy properties, also found in continuous-time systems, which make the velocities uniformly bounded as the time-step goes to zero.

A framework for integrating the geometrical information in the dynamical analysis is presented for various types of contacts. It is based on the generalized distance between two bodies, a mapping that can be defined usually only in the neighborhood of a contact configuration and provides an analytic description of the closest points on two bodies for near contact configurations. This information can

be used in an integration procedure to avoid costly collision detection techniques in the moments following a contact breaking. Ways of describing, in a computationally efficient manner, configurations involving planar contact regions are also presented.

It is known that acceleration-based frictionless problems have always a solution. An important question, in light of the inconsistency of acceleration-based models for general friction configurations, is to determine which acceleration-based models have a solution for small values of the friction coefficient. A partial answer to this question is provided, by showing that disassemblable configurations, for which a force that will simultaneously break all the contacts exists, have the above property.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The dynamic multi-rigid-body contact problem is concerned with predicting the trajectory of a system of rigid bodies with contact and joint constraints. Work in a number of research areas, robotics and virtual reality (VR) especially, has led to strong recent interest in this problem. Virtual reality applications, for instance, need improved physical believability. Many systems exhibit physical inconsistencies that are glaringly apparent to the user due to a poor treatment of loss of contact and multiple contacts between bodies.

The approach of classical dynamics simulation is to consider this problem as a piecewise differential algebraic equation (DAE, see Haug [10]), using the sign of the constraint force as a test to decide if the respective contact will break or not. This approach implicitly assumes that the contact forces are continuous with respect to time. Another approach initiated by Lötstedt in his Ph.D thesis (see also [12]) consists in formulating the acceleration resolution phase of the multi-rigid-body contact problem as a complementarity problem. The latter approach works, in special cases, even if the contact forces are discontinuous. The numerical schemes derived from this approach were especially successful for frictionless configurations.

The complementarity approach has been extended to various classes of applications [3, 7, 20, 22, 23].

1.1 Frictional inconsistency

The usual requirements and model assumptions for simulating multi-body dynamics in connection with virtual reality and robotics applications are

1. The bodies are assumed to be rigid. The computational cost of deformable bodies simulation would be too high for the targeted applications.

2. The bodies cannot interpenetrate. This is modeled locally by a set of inequality constraints ($\Phi_i(q) \geq 0,\ 1 \leq i \leq p$) that have to be satisfied during the simulation. The corresponding contact forces have to be nonnegative, for two bodies in contact can only push on each other, but they cannot pull each other.

3. Joint constraints have to be supported by the simulation environment. This is done by imposing a set of equality constraints ($\Theta_j(q) = 0, 1 \leq j \leq m$) [9].

4. Models have to be provided for impact resolution to accommodate for situations when two objects collide.

5. Objects in the real world exhibit contact friction. Therefore a requirement is for the model to incorporate Coulomb friction or reasonable approximations.

6. Finally, the connection between kinematics and dynamics is done with Newton laws involving accelerations. This is based on the very good experience existing with this approach for DAEs.

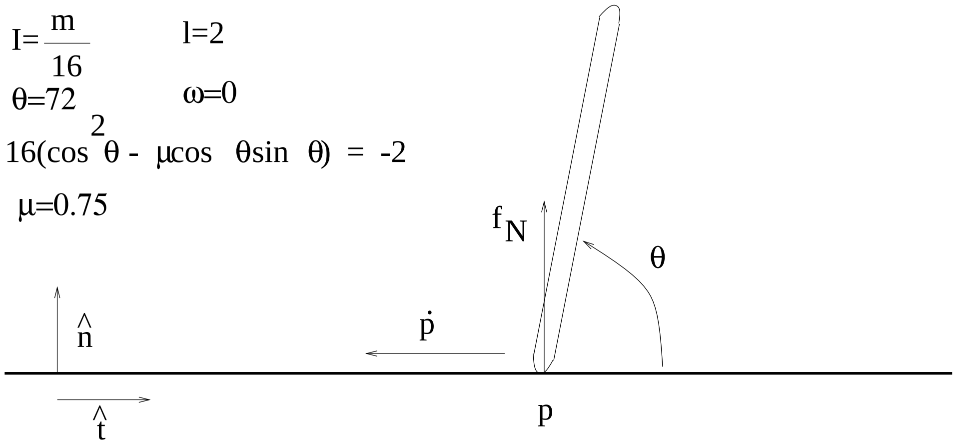

Figure 1.1: A frictionally inconsistent configuration

Unfortunately, it turns out that no model can satisfy the requirements 1,2,5 and 6 at the same time. A simple justification is provided by the following example [3]. The example consists of a thin rod in contact with a table, Figure 1.1. The rod has mass m, inertia $I = \frac{m}{16}$, length $l = 2$, the initial angle $\theta = 72$ degrees and the initial rotation velocity $\omega = 0$ with respect to the center of the mass of the rod. The friction coefficient μ is chosen such that $16(\cos^2(\theta) - \mu \sin\theta \cos\theta)) = -2$. The value of μ for which this happens is $\mu \approx 0.75$. The quantities p, $\dot{p}$ are the initial position and velocity of the tip of the rod. Let $\hat{n}$ be the normal to the table in global coordinates. We assume that initially, the tip of the rod is in perfect contact with the tabletop, $\hat{n}^T p = 0$, and the velocity of the tip of the rod is parallel to the surface of the table, $\hat{n}^T \dot{p} = 0$.

The restriction that the rod should not interpenetrate the table is translated as $\hat{n}^T p \geq 0$. This implies that $\hat{n}^T \ddot{p} \geq 0$, since all the lower derivatives are already

0. Therefore interpenetration or take-off is decided based on the sign of the second order derivative. It is immediate from Newton's equations that

$$\hat{n}^T \ddot{p} = -g + f_N(\frac{1}{m} + \frac{l}{2I}(\cos^2(\theta) - \mu \sin(\theta)\cos(\theta)))$$

where g is the gravitation and f_N is the normal contact force, $f_N \geq 0$ (the direction of the friction force is opposite to the velocity of the tip). With our choice of parameters it turns out that

$$\hat{n}^T \ddot{p} = -g - \frac{f_N}{m}. \tag{1.1}$$

If $f_N = 0$ then the normal acceleration of the tip is $-g$, which is natural (it is like the table does not exist). But if $f_N > 0$, the normal acceleration of the tip becomes even more negative, worsening the interpenetration instead of preventing it. No nonnegative force will prevent interpenetration. Thus the problem has no solution for the above requirements. The configuration is frictionally inconsistent.

The fact that problems involving friction might not have a solution has been recognized from the end of the 19th century [19]. The example used there resembles very much the example above. It must be said that if we allow impulsive reaction then there is a solution to the problem where the rod sticks to the table (the tangential velocity drops to 0 instantaneously). This kind of behavior can be noted when a chalk is driven at an angle against a blackboard, where a sequence of stick-slip movements can be observed (leaving a discontinuous trace on the blackboard).

Several approaches have been tried to avoid frictional inconsistency. Friction can be estimated from the friction history [12]. It is considered an external force in Newton's equations and the frictionless problem is solved (which is guaranteed to have a solution). Unfortunately, for starting configurations there is no friction history and the friction force cannot be estimated. The same thing happens in the impact case, which is assumed to be instantaneous.

Another approach is to look for an impulsive solution each time a solution cannot be found for the acceleration–force model [3]. A collision is generated, resolved and the simulation is then restarted. The problem is that this approach cannot apply for the collision resolution. Also the collision to be solved is arbitrarily generated, leading to non-physical results. In the rod case, the rod will jump off the table with this approach whereas the experiment suggests that the rod sticks to the table.

Finally, one can combine the acceleration resolution step with the integration step for the velocities in order to obtain a discrete problem having velocities and contact impulses as unknowns, rather than accelerations and contact forces [15, 16, 22]. The main advantage is that the new problem has a solution for a much larger class of applications (it is guaranteed to have a solution for one contact with friction, regardless of the configuration [22]). There is no guarantee that this approach yields a solution for the general case, though.

In this thesis, we will modify the model used in [22], such that the new formulation, although identical in spirit with the previous one, will be guaranteed to have a solution, *regardless of the configuration and number of contacts.* The friction model is based on a polygonal friction cone that approximates the Coulomb cone to any prescribed tolerance [22]. To our knowledge this is the first consistent discretization of the dynamic multi-body contact problem with friction that is guaranteed to have a solution for the general case. Although the solution is not be guaranteed to be unique, Lemke's algorithm [5] will always find a solution to our formulation. For a description and properties of Lemke's algorithm, see [5, §4.4]. In addition to the constraints used in [22], our model also incorporates frictionless joint-type constraints.

The discrete model in the Newton-Euler formulation in body coordinates has several properties that are also found in real dynamics systems. This is likely to increase the believability of the simulations, an important goal especially for virtual reality. We prove that the increase in kinetic energy cannot exceed the one obtained in the same configuration, with no contacts enforced. The increase in kinetic energy is at most quadratic with respect to time if the external forces are uniformly bounded and at most exponential if the external forces grow at most linearly with respect to the position and velocity.

An important problem in simulating multi-rigid-body-contact with friction is appropriate handling of impact. The impact is assumed to have two phases. In the first phase, *compression*, interpenetration is prevented by normal compression contact impulses. In the second phase, *decompression*, a fraction e of each normal compression contact impulse is restituted to the system (Poisson hypothesis, [8]). The quantity e is called *restitution coefficient* and may be different for any contact active during a certain collision.

Several models exist for problems with totally plastic collisions (zero restitution coefficient) [12, 22, 14, 15], partial elastic or elastic collisions in the absence of friction [7] (restitution coefficient between 0 and 1 or 1 respectively) as well as for partial elastic or elastic collision with Coulomb friction in two dimensions [8]. In this thesis we develop solvable models in two or three dimensions that take friction into account for collision situations with nonzero restitution coefficient.

We will also show that, for our model, an important class of impacts have a simpler resolution than the general case. For these cases we prove that the kinetic energy after impact cannot exceed the kinetic energy before impact. In [8] the authors claim that the above property is shared by all configurations of the respective two dimensional model. Unfortunately their proof is based on an incorrect proposition, as we will show at the end of Section 2.2.

1.2 Analytical description of a contact configuration

The complementarity approach appears to provide a satisfactory solution to many of the problems encountered by simulation systems [12, 3, 24]. However, several problems arise though when using such an approach . For instance, setting up the problem assumes that the contact constraints can be cast as a few easily available 5 degree-of-freedom (DOF) constraints (*e.g.* point-on-plane, edge-on-edge). But including a cylindrical glass on a table in a linear complementarity setting would require infinitely many point-on-plane constraints to handle the circular shape of the contact surface. Furthermore, consider the case of two bodies whose shapes are regular (but general) surfaces. Determining the contact point and computing the parameters of the contact constraint only as a function of the positions and the shapes is usually connected with determining the distance between the surfaces as a function of the relative position of the bodies. But, determining the distance is itself no easy matter.

It is the aim of Chapter 3 to address these two issues: planar contact with arbitrary shape and arbitrary shape contact with a unique contact point. For the former problem, we present an efficient solution for one particular case (frictionless with nonsingular Jacobian) by treating the configuration as a planar constraint, with a procedure to determine if the contact breaks, and convert the solution into a feasible solution of a linear complementarity problem if it doesn't.

The second problem is related to the contact configuration. Usually one has either a planar or linear contact region or a unique contact point. The former case is usually modeled as several point on plane constraints. If the latter is generated by point-on-plane or edge-on-edge constraints the contact points as well as the kinematic data needed to set up the dynamics equations can easily be found from the bodies' relative positions and velocities. But, for contact configurations that involve smooth shaped bodies (by the latter term meaning that it is a surface or curve sufficiently many times differentiable) we develop an explicit formulation of the differential equations governing the contact points. We then use these equations to integrate the positions of the contact points as well as to set up the basic dynamics equations.

The issue of arbitrary shaped contacts in a general setting has received significant previous attention, especially in the area of robotics. Montana [13] derived explicit contact point evolution equations for surface-surface constraints, based on the geometric parameters of the surfaces. Cai and Roth [4] developed similar equations in an implicit setting and discuss the situation where one of the shapes is a curve as a limit case. Baraff [3] has used an implicit type approach in the context of a complementarity setting of the contact problem.

In this thesis we extend the formulations from [13] and [17] to find the equations of the closest points on the bodies for the case the bodies are not necessarily in contact. We extend the analysis to cases where one or both of the contacting regions are curves. The equations of the closest points are obtained in an explicit manner. From the implementation point of view, there is not a major difference between the implicit and explicit setting since the required matrix inversion is not too expensive in either case (for each contact the matrix is limited in size). However, the explicit setting allows a better analysis of the singularities. We also present a very general form of the contact constraints based on a general definition of the contact normal.

In what follows, we analyze only the frictionless case. For smooth shaped bodies the use of the approach is straightforward in the friction case. However, for arbitrarily shaped planar contact, extension to the friction case is more difficult and we defer discussion of it to a future paper. We also do not include discussion of the geometric issues of collision detection and contact classification, nor our methods of modeling collision dynamics.

1.3 The limits of consistency for the acceleration-based model

Recently, there has been interest in determining whether an acceleration-force based contact problem has a solution for sufficiently small values of the friction coefficient [24]. This is important for understanding the limits of the acceleration based model.

In Chapter 4, we first briefly describe the acceleration based LCP model, and then give a necessary and sufficient condition under which the matrix of the LCP has a special property called strict copositivity. This property is maintained for small values of the friction coefficient and guarantees that the LCP has a solution.

The condition is the *disassemblability* of a frictionless configuration and essentially corresponds to existence of an external force that can break all contacts simultaneously. The notion is close to that of *strongly detachable fixture* from [21] but focuses on multibody systems rather than a unique manipulated object. In our work, we use techniques from linear programming to identify which part of the system is disassemblable and prove that a nondisassemblable system has a solution with infinite contact force. This approach is similar to the one in [23], where the notions of form closure, force closure, and strong force closure were analyzed in a quasi-static setting. The two earlier papers maintain a robotics perspective in which the torques of the manipulator joints are known and the contacts provide all of the kinematic constraints.

Since we are concerned with simulating more general dynamical systems, we consider the effects of joint constraints when discussing disassemblability; joint constraints are explicitly included in the system rather than incorporated into special manipulator subsystems.

Although disassemblability is usually presented in conjunction with the stability of the assembly of a mechanical system, we do not address stability issues in this thesis.

Note that analysis of disassemblability is not the goal of this thesis. Instead, it provides means to characterize the existence of robustness of the LCP for small friction coefficients.

CHAPTER 2

FORMULATING DYNAMIC MULTI RIGID BODY CONTACT PROBLEMS WITH FRICTION AS SOLVABLE LINEAR COMPLEMENTARITY PROBLEMS

Some of the acceleration–force based models do not have a solution for configurations with friction. Previous results [3, 14, 16, 22] suggest that a way to avoid frictional inconsistency is to consider impulsive reaction instead of usual force-based reaction. The purpose of this section is to develop a solvable model for multi rigid body dynamics and collision with Coulomb friction and to integrate it in a time-stepping algorithm.

2.1 The contact model

2.1.1 The constraints and the complementarity condition

In our model, a contact constraint is expressed as $\Phi(q) \geq 0$, where $\Phi(q)$ is a continuously differentiable function and q is a vector parameterizing the position (state) of the multi-body system. If the contact has been established for a while, then $\frac{d}{dt}\Phi(q(t)) \geq 0$. In case this quantity is (strictly) is positive, we have take-off, or breaking of the contact. In other words the contact becomes inactive. Assuming that the position is a differentiable function, one gets $\Phi_q(q(t))\dot{q} \geq 0$, or

$$\Phi_q(q(t))v(t) \geq 0 \tag{2.1}$$

with a strict inequality meaning that the contact breaks.

Let us assume now that we perform a numerical simulation at discrete moments in time and that $q^{(l)}$,$v^{(l)}$ and k are the computed positions, velocities and external forces at time step l. Let h be the corresponding integration step-length. Then an explicit Euler step defines the computed positions at time step $l+1$ as $q^{(l+1)} = q^{(l)} + hv^{(l)}$. Inequality (2.1) gives

$$(n(q^{(l+1)}))^T v^{(l+1)} \geq 0, \tag{2.2}$$

where $n = n(q) = \Phi_q(q)^T$ denotes the Jacobian of the constraint or the contact normal. This definition of the contact normal can be extended even for a reasonably small interpenetration [1]. We note that in [22] inequality (2.2) is replaced by a similar inequality which involves only position values

$$(n(q^{(l+1)}))^T q^{(l+1)} \geq \alpha_0, \tag{2.3}$$

where α_0 is chosen to make this inequality a good approximation of the contact constraint. It turns out that the ability of working with (2.2) rather than with (2.3) is going to be essential for the developments in the present paper.

It can be easily seen that the left-hand side of (2.2) is the normal velocity at the contact point, v_n. Let $c_n n$ be the normal force exercised by the constraint, Figure 2.1. Since the contact can only assume compressive force we have $c_n \geq 0$. In

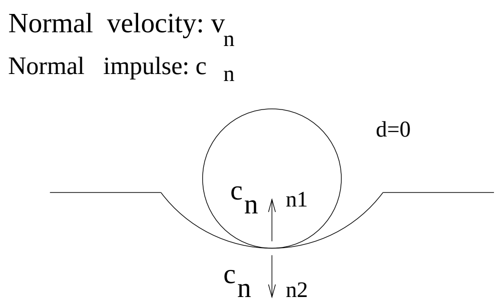

Figure 2.1: Generic contact situation between two bodies. d is the distance between the bodies, The ni are the surface normals at the contact point, and c_n is the contact force.

case of take-off we have $(n(q^{(l+1)}))^T v^{(l+1)} > 0$ and there is no contact force ($c_n = 0$). This generates a complementarity condition: both the normal force and the normal velocity are nonnegative, and if one of them is positive then the other one must vanish. It is convenient to denote this complementarity condition by writing

$$(n(q^{(l+1)})^T v^{l+1} \geq 0 \qquad \text{compl. to} \qquad c_n \geq 0.$$

A multi-body dynamic system has also joint constraints, which are determined by equality constraints of the type $\Theta(q) = 0$, where $\Theta(q)$ is a continuously differentiable function of q. The corresponding velocity constraint becomes

$$(\nu(q^{(l+1)}))^T v^{(l+1)} = 0; \tag{2.4}$$

where $\nu = \nu(q) = \Theta_q(q)^T$ is the Jacobian of the constraint. The force exercised by such a constraint is $c_\nu \nu$, [9], where c_ν can be any real number (positive or negative).

In considering friction we use the same approach as in [22]. Namely we consider Coulomb friction and we assume that the contact force has two components: a

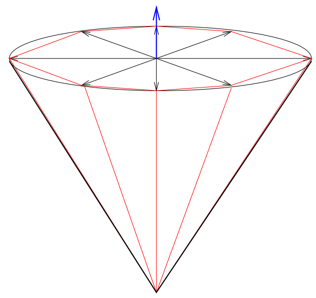

Figure 2.2: The polygonal approximation to the Coulomb cone (in 3d).

normal component c_n and a tangential(friction) component. The Coulomb friction law requires the contact force to lie in a circular cone (or elliptic cone for anisotropic friction). This is a difficult condition to deal with since it leads to a nonlinear complementarity problem which is much more difficult to analyze and/or to solve numerically.

Therefore the cone is approximated by a polygonal cone [22, 24], Figure 2.2,

$$\widehat{FC}(q) = \{c_n n + D\beta \mid c_n \geq 0, \beta \geq 0, e^T \beta \leq \mu c_n\} \tag{2.5}$$

where $e = [1, 1, ..., 1]^T \in \mathbb{R}^u$ and u is the number of edges of the polygonal approximation and μ is the friction coefficient. The u columns of D represent vectors of the space tangent to the contact Also it is assumed that if a vector d is among the columns of D, so is $-d$. In two dimensional simulations the tangent space to a contact is just a line (the line separating the two bodies in contact, passing through the contact region). Therefore the matrix D has only two columns: d corresponding to a direction on the tangent line and $-d$. The particular entries of the matrix D depend on the coordinates used and the mapping between the two dimensional coordinates and the generalized coordinates.

The new complementarity conditions can be written as

$$\lambda e + D^T v^{(l+1)} \geq 0 \quad \text{compl. to} \quad \beta \geq 0 \tag{2.6}$$

$$\mu c_n - e^T \beta \geq 0 \quad \text{compl. to} \quad \lambda \geq 0 \tag{2.7}$$

The complementarity of the vectors of (2.6) is understood componentwise. The additional variable λ is in most cases an approximation of the magnitude of the relative contact velocity [22].

If $\mu c_n - e^T\beta > 0$ then $\lambda = 0$ and therefore $D^T v^{(l+1)} \geq 0$. But if there is a column d of D such that $d^T v^{(l+1)} > 0$, then for the column corresponding to $-d$ we have $-d^T v^{(l+1)} < 0$ contradicting the non negativity of $D^T v^{(l+1)}$. Therefore $D^T v^{(l+1)} = 0$. Since the columns of D^T span the tangent plane at the contact, this means precisely that the tangential velocity is zero (in particular, this shows that $\lambda = 0$ means that there is no sliding motion at the contact, proving in this case that λ is indeed a measure of the magnitude of the relative contact velocity). Therefore if the contact force is inside the cone $\widehat{FC}(q)$ the bodies involved in contact do not exhibit relative motion.

By negating the above implication it follows that, if the bodies in contact are in relative motion, then at least one of the elements of $D^T v^{(l+1)}$ is positive, forcing $\mu c_n - e^T\beta = 0$. Also, from the complementarity condition (2.6) we have $\beta^T D^T v^{(l+1)} \leq 0$, so that the friction force exhibits negative work, as expected. It can be shown that the friction impulse $D\beta$ that satisfies the complementarity friction model maximizes the dissipation $\beta^T D^T v^{(l+1)}$ over all frictional contact impulses, given the normal reaction impulse, c_n, and the velocity v^{l+1}. [22, pg. 2679]. Thus, for isotropic friction laws (circular Coulomb cones) and sufficiently many well-spaced columns of the matrix D, the friction impulse is close to being directly opposite to the component of the sliding velocity v^{l+1} in the friction plane. Therefore in the limit, the above friction model behaves like the Coulomb friction model.

2.1.2 The Newton law and the LCP

We will assume that our multi-body system has p contacts and m joint constraints. A superscript (j) will denote an object associated with the contact (j). The superscript (l) on v and q denotes the velocities and positions at time-step (l). The Newton Law for forces is

$$M\frac{dv}{dt} = F_c + F_{ext}$$

where F_c is the force due to the constraints and F_{ext} represents Coriolis and external forces. M is the mass matrix. In an explicit Euler integration setting, this becomes:

$$M(v^{(l+1)} - v^{(l)}) = hF_c + hF_{ext}$$

The constraint forces are composed of the forces presented in the previous subsection. Therefore adding the Newton Law to the complementarity conditions from the previous sections, we get the following (generalized) LCP:

$$
\begin{aligned}
M(v^{l+1} - v^{(l)}) - \sum_{i=1}^{m} \quad & \nu^{(i)} c_{\nu}^{(i)} - \\
- \sum_{j=1}^{p} \quad & (n^{(j)} c_n^{(j)} \quad + D^{(j)} \beta^{(j)}) = hk && (2.8) \\
{\nu^{(i)}}^T v^{l+1} = 0, \quad & \qquad\qquad i = 1..m && (2.9) \\
{n^{(j)}}^T v^{l+1} \geq 0, \quad & \text{compl. to} \quad c_n^{(j)} \geq 0,\ j = 1..p && (2.10) \\
\lambda^{(j)} e^{(j)} + D^{(j)T} v^{l+1} \geq 0, \quad & \text{compl. to} \quad \beta^{(j)} \geq 0,\ j = 1..p && (2.11) \\
\mu^{(j)} c_n^{(j)} - {e^{(j)}}^T \beta^{(j)} \geq 0, \quad & \text{compl. to} \quad \lambda^{(j)} \geq 0,\ j = 1..p && (2.12)
\end{aligned}
$$

Here k represents the external and Coriolis forces. Note that the variables c_n, c_ν, β represent impulses rather than forces (they are forces multiplicated by h, the step-length). The (generalized) LCP can be written in the following matrix form:

$$
\begin{bmatrix}
M & -\tilde{\nu} & -\tilde{n} & -\tilde{D} & 0 \\
\tilde{\nu}^T & 0 & 0 & 0 & 0 \\
\tilde{n}^T & 0 & 0 & 0 & 0 \\
\tilde{D}^T & 0 & 0 & 0 & \tilde{E} \\
0 & 0 & \tilde{\mu} & -\tilde{E}^T & 0
\end{bmatrix}
\begin{bmatrix}
v^{(l+1)} \\ \tilde{c}_\nu \\ \tilde{c}_n \\ \tilde{\beta} \\ \tilde{\lambda}
\end{bmatrix}
+ b =
\begin{bmatrix}
0 \\ 0 \\ \tilde{\rho} \\ \tilde{\sigma} \\ \tilde{\zeta}
\end{bmatrix}
\tag{2.13}
$$

$$
\begin{bmatrix} \tilde{c}_n \\ \tilde{\beta} \\ \tilde{\lambda} \end{bmatrix}^T
\begin{bmatrix} \tilde{\rho} \\ \tilde{\sigma} \\ \tilde{\zeta} \end{bmatrix} = 0, \qquad
\begin{bmatrix} \tilde{c}_n \\ \tilde{\beta} \\ \tilde{\lambda} \end{bmatrix} \geq 0, \qquad
\begin{bmatrix} \tilde{\rho} \\ \tilde{\sigma} \\ \tilde{\zeta} \end{bmatrix} \geq 0.
$$

Here $\tilde{\nu} = [\nu^{(1)}, ..., \nu^{(p)}]$, $\tilde{c}_\nu = [c_\nu^{(1)}, ..., c_\nu^{(p)}]^T$, $\tilde{n} = [n^{(1)}, ..., n^{(p)}]$, $\tilde{c}_n = [c_n^{(1)}, ..., c_n^{(p)}]^T$, $\tilde{\beta} = [\beta^{(1)T}, ..., \beta^{(p)T}]$, $\tilde{D} = [D^{(1)}, ..., D^{(p)}]$, $\lambda = [\lambda^{(1)}, ..., \lambda^{(p)}]$, $\tilde{\mu} = diag(\mu^{(1)}, ..., \mu^{(p)})^T$ and $\tilde{E} = diag(e^{(1)}, ..., e^{(p)})$ and

$$b = \left[-Mv^{(l)} - hk, 0, 0, 0, 0\right]^T. \tag{2.14}$$

To prove that such a problem has always a solution, we first prove the following solvability theorem.

Theorem 2.1.1 *Consider a (mixed) LCP of the form*

$$\begin{pmatrix} 0 \\ 0 \\ s \end{pmatrix} = \begin{pmatrix} M & -F & -H \\ F^T & 0 & 0 \\ H^T & 0 & N \end{pmatrix} \begin{pmatrix} x \\ y \\ \lambda \end{pmatrix} + \begin{pmatrix} -k \\ 0 \\ b \end{pmatrix} \tag{2.15}$$

$$s \geq 0; \quad \lambda \geq 0; \quad \lambda^T s = 0, \tag{2.16}$$

where M, N, F, H are given matrices and b, k are given vectors of the appropriate dimension.

If M is a symmetric positive definite matrix, N a copositive matrix [5, Def. 3.8.1] and b a nonnegative vector having only nonnegative elements (in particular all components of b can be 0), then the above LCP has a solution. Lemke's algorithm, with precautions taken against cycling, will always find a solution λ of the LCP obtained by eliminating x and y, and then a solution (x, y, λ) of the original LCP can be recovered by solving for x and y in the first two rows of (2.15).

Proof See Appendix A.1.

Theorem 2.1.2 *For any positive definite mass matrix M and nonnegative friction coefficients, the LCP (2.13) has a solution which can be computed by Lemke's algorithm.*

Proof Let N be the matrix made of the last three sets of rows and columns of the matrix form 2.13. Since the matrix is the sum between a positive semidefinite matrix (having $\tilde{E}$ and $-\tilde{E}^T$ as its blocks) and a matrix with nonnegative coefficients ($\tilde{\mu}$) it follows that it is copositive. By defining $H = [\tilde{n}\ \tilde{D}\ 0]$, the conclusion follows after using Theorem 2.1.1 . $\square$

We denote the solution set of LCP (2.13) by $\mathcal{L}(v^{(l)}, hk)$, so that we can write (2.13) as

$$(v^{(l+1)}, \tilde{c}_\nu, \tilde{c}_n, \tilde{\beta}, \tilde{\lambda}) \in \mathcal{L}(v^{(l)}, hk) \tag{2.17}$$

The solution set of the LCP (2.13) also depends on $q^{(l)}$ and h since $\tilde{\nu}, \tilde{n}, M, \tilde{D}$ are evaluated at $q^{(l+1)} = q^{(l)} + hv^{(l)}$ but we will focus on the dependency of (2.13) on the free term (2.8). The previous theorem proves precisely that the set $\mathcal{L}(v^{(l)}, hk)$ is nonempty and that a point in it can be found by Lemke's algorithm.

2.1.3 Energy properties of the formulation

In this subsection, we assume that the mass matrix is constant during the simulation. This happens with the Newton-Euler formulation in body coordinates

[17]. This assumption allows us to prove a property of the formulation that reproduces a similar property in real dynamics systems.

Theorem 2.1.3 *The kinetic energy at the new step, $\frac{1}{2}v^{(l+1)^T}Mv^{(l+1)}$, cannot exceed the kinetic energy $\frac{1}{2}(v^{(l)}+hM^{-1}k)^T M(v^{(l)}+hM^{-1}k)$ of the system in the same configuration, with no constraints enforced.*

Proof From (2.8) we have, after multiplying with $v^{(l+1)^T}$

$$v^{(l+1)^T}M(v^{(l+1)}-v^{(l)}) = \sum_{i=1}^{m} v^{(l+1)^T}\nu^{(i)}c_\nu^{(i)}+$$

$$+\sum_{j=1}^{p} (v^{(l+1)^T}n^{(j)}c_n^{(j)} + v^{(l+1)^T}D^{(j)}\beta^{(j)}) + hv^{(l+1)^T}k. \tag{2.18}$$

Using (2.9), (2.10) and (2.13) we get

$$v^{(l+1)^T}M(v^{(l+1)}-v^{(l)}) = \sum_{j=1}^{p}(v^{(l+1)^T}D^{(j)}\beta^{(j)}) + hv^{(l+1)^T}k \tag{2.19}$$

From (2.11), (2.12) and (2.13) it follows that $v^{(l+1)^T}D^{(j)}\beta^{(j)}$ $= -\lambda^{(j)}e^T\beta^{(j)} = -\mu^{(j)}\lambda^{(j)}c_n^{(j)}$. Therefore

$$v^{(l+1)^T}M(v^{(l+1)}-v^{(l)}) = -\sum_{j=1}^{p}(\mu^{(j)}\lambda^{(j)}c_n^{(j)})$$

$$+hv^{(l+1)^T}k \leq hv^{(l+1)^T}k \tag{2.20}$$

Let $\gamma_2 = M^{1/2}v^{(l+1)}, \gamma_1 = M^{1/2}v^{(l)}, \chi = M^{-1/2}k$. Then (2.20) can be written as

$$\gamma_2^T(\gamma_2-\gamma_1) \leq h\gamma_2^T\chi \tag{2.21}$$

or $\gamma_2^T\gamma_2 \leq \gamma_2^T(\gamma_1+h\chi) \leq \|\gamma_2\|\|\gamma_1+h\chi\|$, the last inequality resulting from the Cauchy inequality. Therefore $\|\gamma_2\| \leq \|\gamma_1+h\chi\|$. Taking the squares, dividing by 2 and replacing with the original variables yields

$$\frac{1}{2}v^{(l+1)^T}Mv^{(l+1)} \leq \frac{1}{2}(v^{(l)} + hM^{-1}k)^T M(v^{(l)} + hM^{-1}k) \tag{2.22}$$

which proves the claim. □

2.2 The collision model

Most existing models for instantaneous inelastic collision can be seen as particular cases of contact dynamics problems, for $h = 0$, cf. [12, 22]. In this section, we will present a simple partially elastic collision model with friction based on the approach used for frictionless collision in [7]. For two dimensions the approach is similar with the one in [8] but, in addition, a solution to the model is guaranteed to be found. We will not consider cases for which portions of the tangential impulse is reversible (like in the highly elastic materials used in making superballs [8]) although our model can be easily modified to include some form of tangential impulse restitution.

The collision is assumed to have two phases: compression and decompression. The collision will generate a new contact (between the bodies that collide) , which, in the compression phase, is added to those already existent on the contact list. Each contact is characterized by a restitution coefficient, e_j, $j = 1..p$, which represents the ratio between the normal contact impulse restituted at decompression to the normal contact impulse during compression (Poisson hypothesis, [8]).

2.2.1 The compression phase

Suppose that the collision is a proper collision, in the sense that it has a nonzero normal velocity. In other words, there is at least a contact, say defined by index $j = 1$, such that ${n^{(1)}}^T v^- < 0$, where v^- denotes the pre-collision velocities.

In the compression phase, the dynamic system will respond with constraint impulses (generated by joints, contacts, or friction). We denote the impulses by the same symbols as in (2.13) but with superscript c. Let v^c be the velocity at the end of the compression phase. At the end of compression, each contact from the list is either maintained $c_n^c \geq 0, n^T v^c = 0$, or is breaking, $c_n^c = 0, n^T v^c \geq 0$. Therefore we recover the same complementarity conditions as in (2.10). Conservation of impulse requires

$$M(v^c - v^-) - \sum_{i=1}^{m} \nu^{(i)} c_\nu^{c(i)} - \sum_{j=1}^{p} (n^{(j)} c_n^{c(j)} + D^{(j)} \beta^{c(j)}) = 0 \tag{2.23}$$

The conditions on joints and friction being identical with (2.9), (2.11) and (2.12) with the appropriate complementarity conditions,it follows that

$$(v^c, \tilde{c}_\nu^c, \tilde{c}_n^c, \tilde{\beta}^c, \tilde{\lambda}^c) \in \mathcal{L}(v^-, 0). \tag{2.24}$$

By Theorem 2.1.2, Lemke's algorithm will find a solution in $\mathcal{L}(v^-, 0)$.

Also, it is worth mentioning that $v^c \neq v^-$, since ${n^{(1)}}^T v^- < 0$ and ${n^{(1)}}^T v^c \geq 0$.

2.2.2 The decompression phase

Let v^+ be the velocity after the decompression phase. Each contact constraint will generate a decompression impulse, dependent on the coefficient of restitution e_j.

For instance, the contact constraint j will generate an impulse $c_n^{d(j)} = e_j c_n^{c(j)} + c_n^{x(j)}$. The additional impulse $c_n^{x(j)} \geq 0$ is necessary to prevent interpenetration. Either the contact breaks, and then $n^{(j)}v^+ \geq 0$, $c_n^{x(j)} = 0$, or it is maintained and then $n^{(j)}v^+ = 0$, $c_n^{x(j)} \geq 0$. Therefore the same type of complementarity as (2.10),

$$(n^{(j)^T})v^+ \geq 0 \qquad \text{compl. to} \qquad c_n^{x(j)} \geq 0.$$

is generated. Hence solving for the post-collision velocity will be similar to (2.13).

The difference between the compression and decompression phase is in the external impulse (free term). The Poisson hypothesis generates an external restitution impulse made of fractions of the normal impulses from the compression phase, $c_n^{(c(j))}n^{(j)}$,

$$F^r = \sum_{j=1}^{p} e_j n^{(j)} c_n^{c(j)}. \tag{2.25}$$

We have assumed here that the compression tangential impulse is irreversible and does not contribute to the restitution force above. More sophisticated collision models for highly elastic objects (superballs) could also introduce some tangential terms in (2.25). For further discussion on the properties of each choice, see [8].

The conservation of impulse requires

$$M(v^+ - v^c) - \sum_{i=1}^{m} \nu^{(i)} c_\nu^{x(i)} - \sum_{j=1}^{p} (n^{(j)} c_n^{x(j)} + D^{(j)} \beta^{x(j)}) = F^r. \tag{2.26}$$

Note that during the decompression phase some tangential frictional impulses do appear ($D^{(j)}\beta^{(j)}$), but these are linked to the normal decompression contact impulse and not to the restitution impulse.

The same friction model gives rise to equations analogue to (2.11), (2.12). By Theorem 2.1.2 the post-collision velocity, v^+ , can be found by Lemke's algorithm since

$$(v^+, \tilde{c}^x_\nu, \tilde{c}^x_n, \tilde{\beta}^x, \tilde{\lambda}^x) \in \mathcal{L}(v^c, F^r). \tag{2.27}$$

There are two special cases worth noting. One is when $e_j = 0, \ \ j = 1..p$ (totally inelastic collision). Then $F^r = 0$ and $v^+ = v^c$ is a solution in $\mathcal{L}(v^c, 0)$ with all constraint impulses being 0, and $\tilde{\lambda}^x = \tilde{\lambda}^c$. In this case we can say that there is no decompression phase.

The other case concerns frictionless contacts. Our remark is an extension of a similar one [7]. Namely, we assume

- (a) all new contacts generated by collision have the same restitution coefficient ϵ. Usually there is only one such contact, but some special cases might have several simultaneous collisions, like a block falling flat on a tabletop.

- (b) The restitution coefficients characterizing the other contacts are less than ϵ, $e_j \leq \epsilon$, $1 \leq j \leq p$.

- (c) The pre-collision velocities satisfy the contact constraints exactly, $(n^{(j)}(q^-))^T v^- = 0$, for the contacts j that were maintained immediately before the collision. (this would be the case for a suitably adapted implicit integration scheme for example).

Under these assumptions a solution of the problem (2.27) is $v^+ = (1+\epsilon)v^c - v^-$, with

$c_n^{x(i)} = (\epsilon - e_i)c_n^{c(i)}$. This means that the only contacts that break at decompression are the ones originating from collision.

A disadvantage of the decompression phase is that for each collision, two LCP's need to be solved. Based on the previous observation, a simplified model for for partially inelastic collision would be to simply use $v^+ = (1+\epsilon)v^c - v^-$, where ϵ is the restitution coefficient of the contacts generated by collision.

Lemma 2.2.1 *If, for the decompression phase, $v^+ = (1+\epsilon)v^c - v^-$, then ${v^+}^T M v^+ \leq v^- M v^-$.*

Proof Let $\gamma_2 = M^{\frac{1}{2}}v^c$, $\gamma_1 = M^{\frac{1}{2}}v^-$, $\gamma_3 = M^{\frac{1}{2}}v^+$. Then since $(v^c, \tilde{c}^c_\nu, \tilde{c}^c_n, \tilde{\beta}^c, \tilde{\lambda}^c) \in \mathcal{L}(v^-, 0)$, we have $\gamma_2^T(\gamma_2 - \gamma_1) \leq 0$ by (2.21) (since $h = 0$). Obviously, $\gamma_3 = \gamma_2 + \epsilon(\gamma_2 - \gamma_1)$. Therefore

$$\begin{aligned} \|\gamma_3\|^2 = \gamma_2^T\gamma_2 + \epsilon^2(\gamma_2 - \gamma_1)^T(\gamma_2 - \gamma_1) + \\ 2\epsilon\gamma_2^T(\gamma_2 - \gamma_1) \leq \gamma_2^T\gamma_2 + \epsilon^2(\gamma_2 - \gamma_1)^T \leq \\ \gamma_2^T\gamma_2 + (\gamma_2 - \gamma_1)^T(\gamma_2 - \gamma_1) \leq \gamma_2^T\gamma_2 + \\ (\gamma_2 - \gamma_1)^T(\gamma_2 - \gamma_1) - 2\gamma_2^T(\gamma_2 - \gamma_1) = \|\gamma_1\|^2. \end{aligned} \tag{2.28, 2.29}$$

Hence $\|\gamma_3\| \leq \|\gamma_1\|$ which proves the claim. $\square$

The property in the conclusion of the Lemma means that the kinetic energy cannot increase by resolving collisions using $v^+ = (1+\epsilon)v^c - v^-$. This is a desirable property for any collision resolution.

Remark In [8] the authors claim that the property from the above Lemma holds for all two dimensional configurations. This statement is based on the claim [8, pg 480] "Since G is symmetric and positive semidefinite and ϵ is a diagonal matrix consisting of elements $0 \le \epsilon_N \le 1$, it follows that

$$\frac{1}{2}\Lambda_C^T \epsilon G \epsilon \Lambda_C - \frac{1}{2}\Lambda_C^T G \Lambda_C \le 0.''$$

The last inequality is violated for $\Lambda_C = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ $G = \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix}$ and $\epsilon = \begin{pmatrix} 0.5 & 0 \\ 0 & 1 \end{pmatrix}$ although all these matrices satisfy the assumptions. This makes invalid the previous statement and the proof of the kinetic energy inequality in [8].

2.3 Strategy for integrating the equations of motion

Various numerical schemes can be devised based on the previous sections. In what follows we propose a simple Euler integration method to find a numerical solution for the time interval $[0, T]$.

$v = v^0$, $q = q^0$; $time = 0$;

while $(time < T)$

  $q_{new} = q + hv$;

  Find $(v_{new}, \tilde{c}_\nu, \tilde{c}_n, \tilde{\beta}, \tilde{\lambda}) \in \mathcal{L}(v, hk)$ ($\tilde{\nu}, \tilde{D}, \tilde{n}, M$ are evaluated

    at q_{new}, k is evaluated at (q, v))

  if (no collision detected between $time$ and $time + h$)

$time = time + h$, $q = q_{new}$, $v = v_{new}$;

else

Estimate the collision time $time_{new}$, collision

position and velocity q_{new} and v^-;

Find $(v^c, \tilde{c}^c_\nu, \tilde{c}^c_n, \tilde{\beta}^c, \tilde{\lambda}^c) \in \mathcal{L}(v^-, 0)$;

Find $(v^+, \tilde{c}^x_\nu, \tilde{c}^x_n, \tilde{\beta}^x, \tilde{\lambda}^x) \in \mathcal{L}(v^c, F^r)$, with F^r

defined in (2.25);

(Alternately, choose $v^+ = (1+e)v^c - v^-$)

$time = time_{new}$, $v = v^+$, $q = q_{new}$.

end if

end while

Several details need to be specified. We assume that the initial velocity v^0 is consistent with the constraints, i.e. that $(n^{(j)}(q_0))^T v^0 \geq 0$ and $(\nu^{(i)}(q_0))^T v^0 = 0$ for $i = 1..m$, $j = 1..p$. If that were not true, a collision would take place. The algorithm can be easily modified to handle an initial collision. At each step an LCP of the form (2.13) is solved. To set it up, $\tilde{\nu}, \tilde{D}, \tilde{n}, k, M$ need to be specified. Following the discussion in Subsection 2.1.1, all these qualities are evaluated at q_{new}. Variants of this scheme would be to evaluate the data or parts of it at q or $q + \frac{h}{2}v$ [22]. The integration time-step does not have to be constant during the integration an can be changed at each step. One can easily devise a variant of this algorithm which uses a more sophisticated integration method (Runge-Kutta for example).

Also, the collision detection rules need to be specified. A (theoretically) simple collision rule would be to check for $\Phi^{(i)}(q) \leq 0$ (interpenetration). One problem is that the LCP decides take-off based on the condition $\Phi_q^{(i)}(q)v > 0$ but this does not necessarily mean that there was no interpenetration, for $\Phi^{(i)}(q)$ can be negative, due to numerical error. If a new collision between the bodies occurs before $\Phi^{(i)}(q) > 0$ it cannot be detected based on the sign of $\Phi^{(i)}(q)$. A solution would be to use a Nonlinear rather than a Linear Complementarity Problem [22] which enforces the constraint $\Phi^{(i)}(q) \geq 0$ with the cost of additional computation. Another solution, in the case when interpenetration persists even though the LCP (2.13) dictates take-off, would be to check for $\Phi_q^{(i)}(q)v \leq 0$ (normal velocity becomes negative) as a collision condition. This ensures that either there is no interpenetration, or there is an instantaneous decrease in interpenetration. This aspect shows the necessity of a good integration of the geometric and dynamic algorithms, for the outcome can be contradictory due to the discrete nature of the numerical schemes.

Assume that the contact constraint (i) is exactly satisfied at the current time moment, $\Phi^{(i)}(q^{(l+1)}) = 0$ and the LCP results in
$n^{(i)^T}(q^{(l+1)})v^{(l+1)} > 0$. Then $\frac{d}{dh}\Phi^{(i)}(q^{(l)} + hv^{(l+1)}) > 0$. Therefore take-off is certain at least on a small time interval following the current moment. This is an important property which shows the robustness of the formulation in the vicinity of the constraint manifold.

There is a class of configurations for which these algorithms cannot be guaranteed to sweep the interval $[0, T]$ in finite time. One example would be a ball bouncing on the floor with a restitution coefficient strictly between 0 and 1. The ball will sustain an infinite number of collisions in finite time, and the algorithm cannot get past the moment the ball has 0 velocity. One solution to this problem is to require that the restitution coefficient is zero for normal collision velocities less than a certain value.

2.4 Uniform boundedness of the numerical solution with respect to integration time step

In this section we analyze the properties of our formulation for the case of the Newton-Euler [17] equations in body coordinates. The main advantage in using this formulation is that the mass matrix is constant throughout the simulation. The drawback is that the Coriolis force, which is quadratic in velocity, has to be taken into account, which complicates the regularity analysis, because the assumption that the force is bounded, used in [22] is not obvious in this case. For one body, the Newton-Euler equations in body coordinates are [17, §2.4]

$$m\dot{v}^b + \omega^b \times mv^b = f^b \tag{2.30}$$

$$\mathcal{I}\dot{\omega}^b + \omega^b \times \mathcal{I}\omega^b = \tau^b \tag{2.31}$$

where m is the mass of the body, $\mathcal{I}$ is the inertia matrix (which is diagonal for a suitable choice of coordinates). The vectors v^b and ω^b represent the translational and rotational velocities in body coordinates. The term $f_c = -[(\omega^b \times mv^b)^T,\ (\omega^b \times \mathcal{I}\omega^b)^T]^T$

is the Coriolis force. An important property of the Coriolis force is $[v^{b^T}, \omega^{b^T}]f_c = 0$. This reflects the fact that the Coriolis force produces no work (it is perpendicular to the velocity), and will be a significant asset in proving the uniform boundedness of velocities. It should also be mentioned that for two dimensional simulations, the Newton-Euler formulation in global coordinates [17, §2.4], generates a constant mass matrix and does not exhibit Coriolis force.

We consider a dynamical system under the action of three types of external forces: forces whose magnitudes increase at most linearly with the distances (e.g., elastic forces), forces whose magnitude increase at most linearly with the velocity (e.g., viscous friction), and forces which are bounded during the simulation time interval $[0\ T]$. The following theorem proves that the kinetic energy obtained by the algorithm in the previous section is uniformly bounded on the time-interval $[0\ T]$. The step-size is not necessarily constant during the integration, but it is bounded by h. In other words for a given h, the time steps h_l can be different during the integration but they have to be smaller than h (h is just an upper bound on the individual step–sizes, and not necessarily the time–step).

Theorem 2.4.1 *Assume that the algorithm from the previous section solves a finite number of collisions between* 0 *and* T *and that the collision resolution method ensures that the kinetic energy after the collision is no greater than the one before collision. Let* $v(0), v(T)$ *and* $q(0), q(T)$ *be the velocity and position at the beginning and the end of the simulation period.*

a) If the external force is of the form $k(v,q) = f_c(v) + k_1(v,q) + k_2(v) + k_3(q)$, where $\|k_1(v,q)\| \leq d_1$, $\|k_2(v)\| \leq d_2\|v\|$ and $\|k_3(q)\| \leq d_3\|q\|$, $\forall q, v$, then for all sufficiently small h

$$v^T(T)Mv(T) \leq (v^T(0)Mv(0) + \|q(0)\|^2)e^{2cT} + 2cTe^{2cT}. \tag{2.32}$$

b) If $k_2(v) = k_3(q) = 0$, then for all sufficiently small h

$$\sqrt{v^T(T)Mv(T)} \leq \sqrt{v^T(0)Mv(0)} + cT + 1. \tag{2.33}$$

In both cases, c is a constant that depends solely on $d_i > 0, i = 1..3$ and on the mass matrix M.

Proof See Appendix A.2.

The forces considered in the previous theorem cover most cases of external forces in real dynamical systems. Therefore we see that for fairly general situations our algorithm provides solutions that are uniformly bounded as $h \downarrow 0$ for the Newton-Euler formulation. If $k_2(v) = -d_2 v$, $d_2 > 0$ (damping force) and $k_3(q) = 0$ it can be shown that the velocities satisfy a bound like (2.33). Another interesting point is that if the external forces (except for the Coriolis force) are uniformly bounded as $T \rightarrow \infty$, then the increase in velocity is at most linear with respect to time, which is also true in real dynamical systems.

CHAPTER 3

FORMULATING 3D CONTACT DYNAMICS PROBLEMS

In order to set up a multi body dynamics with contact and friction one needs kinematic data to describe the contact constraint. In this section this information is determined for several configurations based on the geometry of the bodies involved in the contact. This information is also used to determine how the bodies in contact touch at a subsequent moment, easing up the task of the geometrical analysis system. This task is considerably more difficult for the case of planar contact region where an infinite number of point-plane constraints should be considered (one for each point on the boundary of the contact region. In this section an algorithm to represent these cases as three point on plane constraints is presented significantly reducing the complexity of the problem.

3.1 The assumptions and setting of the problem

3.1.1 The extremal distance between bodies

Assume that we have two bodies in 3-dimensional space and the system configuration can be represented by $q \in R^{12}$ a vector of coordinates, half accounting for the translational coordinates and the other half for the rotational coordinates. Also, assume for now that the bodies have a sufficiently regular shape (ultimately we will relax this constraint and require only piecewise smoothness). We are interested in finding the distance between the bodies as well as the closest points on the bodies. A large part of this paper is concerned with finding the new closest points following a small relative motion. If a pair of bodies has been in "perfect" contact and we allow for a small motion subject to some nonpenetration constraints it might happen that we get a small subsequent interpenetration due to the numerical integration. It is then useful, in order not to run into kinematic inconsistencies, to define the "extremal distance" and the "extremal points" [3, p.298]. We consider the "extremal points" to be points P_{10} on the first body and P_{20} on the other body such that for any pair of differentiable curves $P_1(\tau)$ and $P_2(\tau)$ on the first and second bodies respectively, with $P_1(0) = P_{10}$, $P_2(0) = P_{20}$, we have

$$\frac{d}{d\tau}\|P_2(\tau) - P_1(\tau)\|^2 = 0 \tag{3.1}$$

at $\tau = 0$. If two bodies are separated or in perfect contact, the extremal points coincide with the closest points on the two bodies. In the case of interpenetration, assuming sufficiently small interpenetration and sufficiently smoothly shaped bodies, the "extremal distance" d is the distance between the extremal points taken with a minus sign. This defines a unique set of contact points whose positions depend on the relative body positions in a regular manner; we define $P_1(q)$ and $P_2(q)$ such that for any q, $P_{10} = P_1(q)$ and $P_{20}P_2(q)$ are the extremal points at corresponding to body positions q.

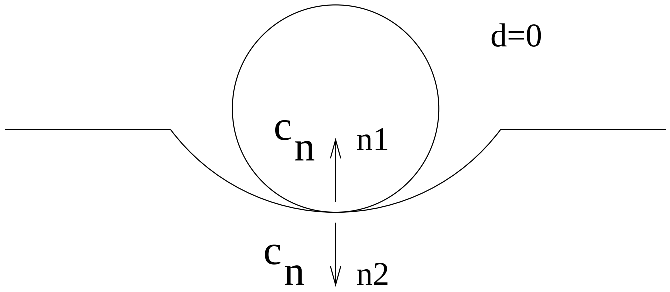

Figure 3.1: Generic contact situation between two bodies. d is the distance between the bodies, The ni are the surface normals at the contact point, and c_n is the contact force.

We have chosen a different formal definition of the extremal points than [3] because of the need to accommodate contact situations that involve curves.

As an example of defining such an "extremal distance" function take a fixed table with emerging unit normal N, p_t an arbitrary point lying on the table and p the coordinate of a body made of a unique point. Then the function $d = N \cdot (p - p_t)$ coincides with the distance (from the point p to the table) when positive, is negative at interpenetration, and is clearly differentiable in terms of coordinates. It is clear that such a distance function cannot be well defined for arbitrary interpenetration of two bodies. But since our goal is to be able to set up some geometrical constraints, the function d needs to be defined with this regularity only in neighborhoods of the current contact points.

Subsequently, we will refer to extremal points as "contact points" even when there is interpenetration or no contact between the bodies, and to function d as "the distance between the bodies". We will also assume that $P_1(q)$ and $P_2(q)$, the functions giving the contact points with respect to the body positions, as well as d, are sufficiently differentiable. We call configurations meeting these conditions "regular". We will see in the following sections that these are reasonable assumptions if the distance d between the bodies is sufficiently small and the body shapes and positions do not exhibit certain peculiarities (such as parallel edge-on-edge contact or excessive mutual flatness).

3.1.2 The complementarity condition

Consider two rigid bodies in contact as shown in Figure 3.1. Assume that the bodies have been in contact for some finite (non-infinitesimal) period of time. Therefore $d = 0$ and $\dot{d} = 0$. Since $d \geq 0$ it follows that we require $\ddot{d} \geq 0$ in order to avoid interpenetration in a subsequent instant.

Let λ be the size of the mutual contact force. If the force has the direction in the figure, it follows that $\lambda \geq 0$, for the two bodies can only push, but not pull, on one another (*i.e.*, the contact can only sustain compressive force). Thus, contact constraints are sometimes called *unilateral* constraints, while traditional joint constraints, which can assume both positive and negative reaction forces, are termed *bilateral* constraints.

Assume now that $\ddot{d} > 0$. This means that in the following instant the contact between the two bodies will break. As a result, the mutual force must vanish ($\lambda = 0$). This leads to

$$\ddot{d} \times \lambda = 0. \tag{3.2}$$

Two nonnegative variables that satisfy such a condition are said to be complementary variables. By taking the derivative of the relation $d = \phi(q)$, we get

$$\ddot{d} = \phi_q \ddot{q} + \Gamma(q, \dot{q}) \tag{3.3}$$

where Γ is a function whose form is not important for the purpose of the subsequent theoretical derivations. This equation will be important in setting up a linear relationship between $\ddot{d}$ and λ.

3.1.3 The general setting of the contact problem

Suppose that we have a system of several bodies along with a set of m contacts between the bodies. We assume that each contact is a regular contact and let $d_i = \phi_i(q)$, $1 \leq i \leq m$, be the distance functions associated with the contact set, where q is a parameterization of the position of the bodies (a list of positions and rotation parameters for each body). For our purposes each d_i has to be at least twice continuously differentiable. Let Φ be the function that maps the configuration space (the space of q's) to R^m and which is made up of the ϕ_i's, $1 \leq i \leq m$.

The functions $d_i = \phi_i$ do not have to represent the distance between the two bodies in contact. For instance if we have two bricks one on top of the other, with parallel sides but not touching, no such differentiable distance function exists. Instead we may choose three points on one of the bricks and require that the distance between each of them and the plane containing the corresponding side of the other brick be positive (point-plane constraint). Each of these three functions is differentiable. Similarly, we can model a polygonal pyramid sitting with its base on a table by requiring that each of corner of the base not penetrate the table surface. Thus, we use a point-plane constraint for each vertex of the base.

For completeness, assume that the system has, in addition to contact constraints, L joint-type constraints defined by the map $\Theta(q)$ taking the configuration space to R^L. For details concerning joint constraints see, for example, [9].

With these two maps, we can formulate the contact problem:

Problem Given the positions q and velocities $\dot{q}$ of a set of bodies, find the accelerations that define a motion consistent with the following equality and inequality constraints.

$$\Phi(q) \geq 0 \tag{3.4}$$

$$\Theta(q) = 0 \tag{3.5}$$

□

It is known [9] that for frictionless cases the direction of the constraint force is perpendicular to the manifold defining the constraint surface. Thus the forces due to active contact constraints are of the type $\Phi_q(q)^T\lambda$ where λ is the vector of the contact forces. Similarly $\Theta_q(q)^T\nu$ is the reaction force corresponding to the joint constraints (3.5), where ν is the vector of the contact forces. By writing Newton's law and computing the second derivatives of the constraints (3.4, 3.5), we get:

$$\begin{pmatrix} 0 \\ 0 \\ s \end{pmatrix} = \begin{pmatrix} M & -\Theta_q^T & -\Phi_q^T \\ \Theta_q & 0 & 0 \\ \Phi_q & 0 & 0 \end{pmatrix} \begin{pmatrix} \ddot{q} \\ \nu \\ \lambda \end{pmatrix} + \begin{pmatrix} F \\ \Gamma_1 \\ \Gamma_2 \end{pmatrix}. \tag{3.6}$$

where $s_i = \ddot{d}_i = \ddot{\phi}_i(q, \dot{q})$ $1 \leq i \leq m$ is the vector of contact accelerations. F is the sum of input forces and the Coriolis forces and Γ_1 and Γ_2 are terms involving velocities, $\dot{q}$, from the second derivatives of Φ and Θ with respect to time. (This notation is slightly abusive since in a Newton-Euler setting, $\ddot{q} = (\ddot{r}, \dot{\omega})$, where ω are angular velocities, a globally nonintegrable vector field).

For each contact the complementarity relation (3.2) must hold. Thus, $s_i\lambda_i = 0,\ 1 \leq i \leq m$. By adding up the relations (3.2) for all the contact constraints, we get

$$s^T\lambda = 0;\ s, \lambda \geq 0. \tag{3.7}$$

The relations (3.7) and (3.6) determine a linear complementarity problem [5]. It is known that the problem has a unique $\ddot{q}$ as a solution, but the forces ν and λ need not be unique. This can be seen in various easily constructed examples.

3.1.4 Problems related to the LCP setting

- To set up (3.6), Φ_q, Θ_q, Γ_1 and Γ_2 are needed. Θ_q and Γ_2 originate from derivatives of joint constraints, $\Theta(q) = 0$, that are are usually readily available explicit functions. This is not the case for Φ_q and Γ_2, though, because the functions $\phi(q)$ can be computed for a given q only by solving a minimization problem (for a general contact situation).

- For good computational results the problem should have a reasonable dimension. But if the contact region between two bodies is planar and arbitrarily shaped (non-polygonal), we might need to use infinitely many point-on-plane constraints to model it (e.g. for a problem involving a cylindrical glass standing upright on a table).

3.2 Explicit setting of the dynamics equations

In this section we focus on determining Φ_q and Γ_2 of (3.6) for a general contact situation. We restrict ourselves to the Newton-Euler formulation in global coordinates (which means that q of Section 3.1 consists of the usual translational and rotational coordinates for each body in the world frame). If B is a body, let p be the position, and R the rotation matrix of its center of mass in world coordinates.

We analyze a regular configuration of two bodies B_1 and B_2. Let $d = \phi(q_1, q_2)$ be the distance between the bodies as defined in Section 3.1.1. Let (P_{10}, P_{20}) be the pair of contact points, with $P_{10} \in B_1$, $P_{20} \in B_2$. If $d \neq 0$, we define the "normal"

$$N = \frac{P_{10} - P_{20}}{d}. \tag{3.8}$$

For the case $d(q_1, q_2) = 0$, let q_0 be the concatenation of the body coordinates q_1 and q_2. Because of our regularity assumption, for each q in a neighborhood of q_0 there is a unique pair of contact points $(P_1(q), P_2(q))$. Assuming the limit exists, we define the normal in this case as

$$N = \lim_{q \to q_0} \frac{P_2(q) - P_1(q)}{d(q)}. \tag{3.9}$$

If it exists, it clearly defines a unit vector N that emerges from B_1. If a regular surface is involved in the contact, then the definition coincides with the surface normal. If the contact is an edge-on-edge [1] or curve-on-curve, then it can be shown that this normal is the cross-product of the tangents to the curves at the contact points (with the appropriate sign). It is clear that if we have a corner-on-corner contact, then we obtain different limits if we converge from different directions to q_0; thus, such configurations are singular. Finally, let l_1, l_2 be the local coordinates of $P_1(q)$ and $P_2(q)$ relative to B_1 and B_2 respectively. Let p_1, R_1 and p_2, R_2 be the coordinates of the bodies with respect to the world. At the moment of the contact we have $P_1(q) = P_2(q) = p_1 + R_1 l_1 = p_2 + R_2 l_2$.

Lemma 3.2.1 *The Jacobian row $J_{12} = d_q = \phi_q$ (in world coordinates) of the contact between the bodies at a specified time moment is:*

$$J_{12} = (-N^T, -(R_1 l_1 \times N)^T, N^T, (R_2 l_2 \times N)^T). \tag{3.10}$$

Proof Assume that the bodies are separated at a distance $d \neq 0$. The result for $d = 0$ follows as the limit of the general case. The distance is continuously differentiable with respect to time because of the regularity assumptions. Let (P_1, P_2) be the pair of closest points. Then we have $|d| = \|P_2 - P_1\|$, hence $d = \frac{|d|}{d}\|P_2 - P_1\|$

[1] We use "edge-on-edge" to refer to contacts involving straight (non-curved) line segments.

and

$$\dot{d} = \frac{d}{dt}\|P_2 - P_1\|\frac{|d|}{d} = \frac{d}{dt}\langle (P_2 - P_1), (P_2 - P_1)\rangle^{\frac{1}{2}}\frac{|d|}{d} =$$

$$\frac{1}{|d|}\frac{|d|}{d}\langle \frac{d}{dt}(P_2 - P_1), (P_2 - P_1)\rangle = \frac{1}{d}\langle \frac{d}{dt}(P_2 - P_1), dN\rangle = \langle \frac{d}{dt}(P_2 - P_1), N\rangle, \quad (3.11)$$

where by $\langle x, y\rangle$ we mean the scalar product $x^T y$. It is now clear that the computation of the Jacobian row is no longer singular at $d = 0$. Replacing P_i by $p_i + R_i l_i,\ i = 1, 2,$ yields

$$\dot{d} = \langle \dot{p}_2 - \dot{p}_1, N\rangle + \langle R_2\dot{l}_2 - R_1\dot{l}_1, N\rangle + \langle \dot{R}_2 l_2 - \dot{R}_1 l_1, N\rangle. \quad (3.12)$$

We focus now on the middle term. We fix the position of the bodies to be the one at time $t = t_0$. For this position, $l_1(t)$ and $l_2(t)$ are two curves on the bodies that trace the contact points. $l_1(t_0)$ and $l_2(t_0)$ are the current contact points. So, condition (3.1) applies and we get

$$\langle R_2\frac{d}{dt}l_2(t_0) - R_1\frac{d}{dt}l_1(t_0), P_2 - P_1\rangle = 0. \quad (3.13)$$

But now, dividing by d and using the definition (3.8) of the normal, we get

$$\langle R_2\dot{l}_2(t_0) - R_1\dot{l}_1(t_0), N\rangle = 0. \quad (3.14)$$

Thus the middle term in 3.12 vanishes and we obtain:

$$\begin{aligned}
\dot{d} &= \langle \dot{p}_2 - \dot{p}_1, N\rangle + \langle \dot{R}_2 l_2 - \dot{R}_1 l_1, N\rangle \\
&= \langle \dot{p}_2 - \dot{p}_1, N\rangle + \langle \dot{R}_2 R_2^T R_2 l_2 - \dot{R}_1 R_1^T R_1 l_1, N\rangle \\
&= \langle \dot{p}_2 - \dot{p}_1, N\rangle + \langle \omega_2 \times R_2 l_2 - \omega_1 \times R_1 l_1, N\rangle \\
&= \langle \dot{p}_2 - \dot{p}_1, N\rangle + \langle R_2 l_2 \times N, \omega_2\rangle - \langle R_1 l_1 \times N, \omega_1\rangle. \quad (3.15)
\end{aligned}$$

Since this is true for any velocities, and we must have

$$\dot{d} = J_{12}(\dot{p_1}, \omega_1, \dot{p_2}, \omega_2)^T \tag{3.16}$$

and the claim follows. □

Remark It might seem strange that we are able to get continuous derivatives of the distance function, for this is a non-differentiable function in R^3. The point is that we have a privileged direction in the space, the normal N, and we have defined a signed distance which roughly corresponds to measuring the distance along the direction of N. This function has the required regularity (like the function $f(x) = x$ for $x \in R$). Obviously, such a definition can be valid only locally.

As a result of the previous lemma, we can completely define the terms involving contacts in (3.6).

Corollary 3.2.2 *For each contact i, $1 \le i \le m$ let l_{1i} and l_{2i} be the contact point coordinates for the bodies, B_{1i} and B_{2i}, involved in the contact. And let p_{1i}, R_{1i} and p_{2i}, R_{2i} be the coordinates of the bodies, and N_i the contact normal (emerging from body B_{1i}). Then*

$$\Phi_i = (0, \cdots 0, -N_i^T, -(R_{1i}l_{1i} \times N_i)^T, 0, \cdots 0, N_i^T, (R_{2i}l_{2i} \times N_i)^T, 0, \cdots, 0) \tag{3.17}$$

is the i-th row of the matrix Φ, the nonzero entries corresponding to the coordinates of B_{1i} and B_{2i}, and

$$\begin{aligned}\Gamma_{2i} &= \langle \dot{N}_i, \dot{p}_{2i}\rangle - \langle \dot{N}_i, \dot{p}_{1i}\rangle + \langle (R_{2i}l_{2i} \times \dot{N}_i), \omega_{2i}\rangle - \langle (R_{1i}l_{1i} \times \dot{N}_i), \omega_{1i}\rangle \\ &\quad + \langle (\omega_{2i} \times R_{2i}l_{2i}), (\omega_{2i} \times N_i)\rangle - \langle (\omega_{1i} \times R_{1i}l_{1i}), (\omega_{1i} \times N_i)\rangle + \langle (R_{2i}\dot{l}_{2i} \times N_i), \omega_{2i}\rangle \\ &\quad - \langle (R_{1i}\dot{l}_{1i} \times N_i), \omega_{1i}\rangle \end{aligned} \tag{3.18}$$

is the i-th element of Γ_2.

Proof The relation for Φ_i is a mere restatement of the previous lemma, whereas for Γ_{2i} we compute another derivative at (3.15) and take the terms that do not involve accelerations. □

This corollary can be used to set up the dynamics equations, provided $\dot{N}_i, \dot{l}_{1i}, \dot{l}_{2i}$ can be expressed in terms of the velocities of the bodies in contact. All these derivatives depend strongly on the particular geometry of the contact; therefore each contact type requires a different approach. If we have a simple geometry (point-on-plane, edge-on-edge) $d_i = \phi_i(q)$ is a simple algebraic expression, and we do not need to go through the procedure of Lemma 3.2.1 . But for surface-on-surface, curve-on-surface and curve-on-curve constraints there is no easy-to-derive algebraic form of $d_i = \phi_i(q)$; Corollary 3.2.2 helps in setting up the equations for those cases. Its main advantage is that the only thing required to explicitly set up 3.6 is some kinematic information (such as $\dot{N}$ and $\dot{l}_i$), which we investigate in the next section.

3.3 The kinematic description of the contact

of smooth-shaped bodies

In this section we analyze the kinematics of point contact between an arbitrary but smooth-shaped bodies. Here by smooth we mean that the position of the contact points of the bodies varies smoothly with the relative positions of the bodies. Even if the contact points on the bodies are on a curved edge, the contact is still smooth under some mild non-degeneracy assumptions. We confine ourselves to three cases of interest: surface-on-surface, curve-on-curve, and curve-on-surface. We could also have analyzed the point-on-surface case, but if the surface has an algebraic description, the contact ($d = 0$) case is treated very much like a joint constraint; for it, the analysis would be tedious and not particularly useful.

There are various ways one can perform the kinematic description of an arbitrary but smooth shaped bodies contact. In the robotics literature there are two major approaches. The first describes the shapes of the bodies in contact by functions of two variables and determines the kinematic parameters from the derivatives of these functions and the velocities of the bodies involved in contact [4]. A related approach has been used in 3D arbitrary shaped contact simulation where the motion of the contact points is described as constrained motion and the contact velocities are retrieved via the implicit function theorem [3].

The other approach, which we will use in this paper, assumes that the contact configurations are described with local charts and, if a surface is involved, a local orthonormal frame is at hand [13, 17, 18]. The resulting kinematic relations are obtained in terms of the intrinsic parameters (curvature, torsion, etc.). In robotics, this approach is of special interest in contour following [13]. The analysis done here is within the framework described in [17, chapter2]. The basic object in this setting, V, the relative velocity between two rigid bodies, is a six dimensional vector. Since we will make extensive use of it and its algebra, we present a brief review of its characteristics in Appendix A.3. The proofs of the kinematic descriptions are rather tedious and are deferred to the Appendices A.4 through A.7.

The bodies considered below are not necessarily in contact, but can be at some distance $d \neq 0$ in the sense of Subsection 3.1.1. The analysis extends relatively easily for this case. For each of the three types of contact analyzed we will give conditions under which the contact velocities can be uniquely and differentially determined from the relative velocities of the bodies in contact. As a consequence, any nonsingular configuration will be a regular configuration, because any relative configuration in the neighborhood of the given configuration will give a differentiable contact pair.

3.3.1 The surface-surface contact

3.3.1.1 The equations

We consider two bodies whose center of mass frames are O and F respectively. On each body let C_o and C_f be the instantaneous frames of contact on each body, and $L_o(\tau)$ and $L_f(\tau)$ respectively be the family of frames associated with each body, such that at time t we have $L_o(t) = C_o$ and $L_f(t) = C_f$. Let d be the distance between the two bodies, which is measured along the z direction of the contact frames, and

ψ the angle of rotation with respect to the contact frame around the z axis. ψ is chosen such that a rotation with angle ψ will make the axis x of C_o and C_f parallel (see Figure3.2).

We will assume that all calculations are done within a single map on each body, so that we will not worry about globality problems. The surfaces are given by means of the local coordinates, (u_1, v_1) and (u_2, v_2) respectively. $c_o(\alpha_o)$ and $c_f(\alpha_f)$ represent the contact points in local coordinates, where $\alpha_o = (u_1, v_1)$ and $\alpha_f = (u_2, v_2)$ at the contact point. For each of the surface we denote by M, K, T the metric, curvature and respectively torsion tensor (with the appropriate subscripts).

This case is a straightforward generalization of the case $d = 0$ which has been addressed in [13, 17, 18].

Theorem 3.3.1 Kinematic equations of contact *The motion of the contact coordinates as a function of the relative motion is given by*

$$\dot{\alpha_f} = M_f^{-1}(K_f + \tilde{K_o} + d\tilde{K_o}K_f)^{-1}((I + d\tilde{K_0})\begin{pmatrix} -\omega_y \\ \omega_x \end{pmatrix} - \tilde{K_0}\begin{pmatrix} v_x \\ v_y \end{pmatrix}) \tag{3.19}$$

$$\dot{\alpha_o} = M_o^{-1}R_\psi(K_f + \tilde{K_o} + dK_f\tilde{K_o})^{-1}(\begin{pmatrix} -\omega_y \\ \omega_x \end{pmatrix} + K_f\begin{pmatrix} v_x \\ v_y \end{pmatrix}) \tag{3.20}$$

$$\dot{\psi} = \omega_z + T_f M_f \dot{\alpha_f} + T_o M_o \dot{\alpha_o} \tag{3.21}$$

$$\dot{d} = v_z \tag{3.22}$$

where $(v_x, v_y, v_z, \omega_x, \omega_y, \omega_z) = V^b_{l_o l_f}$ *(see Appendix A.7),* $R_\psi = \begin{bmatrix} \cos\psi & -\sin\psi \\ -\sin\psi & -\cos\psi \end{bmatrix}$, *and* $\tilde{K_o} = R_\psi K_o R_\psi$.

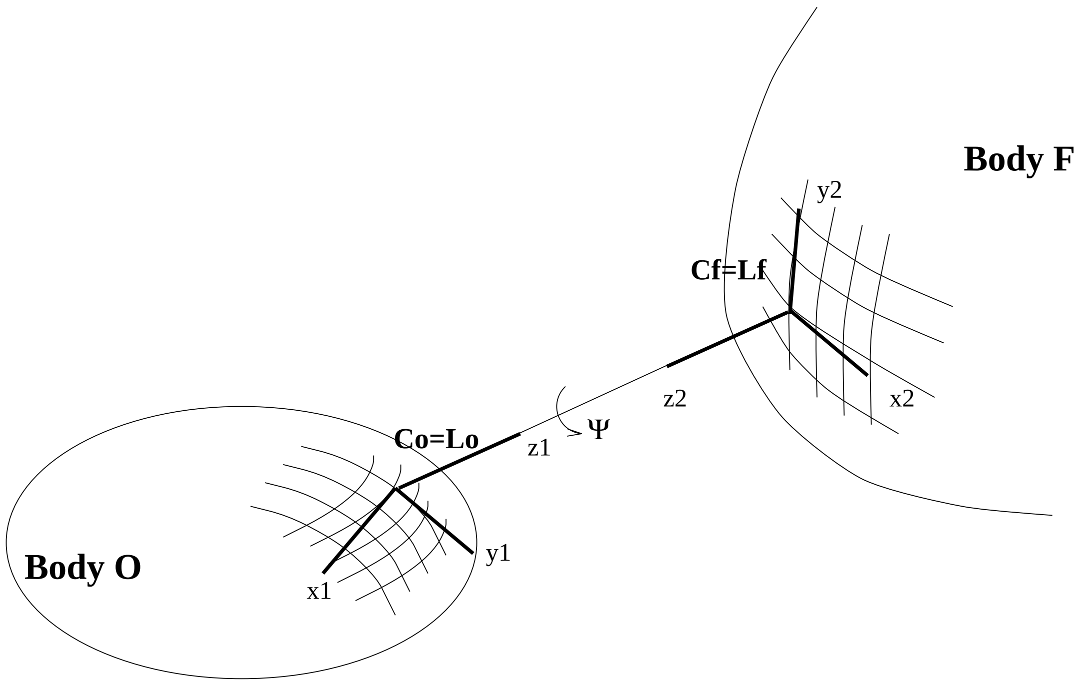

Figure 3.2: A general contact configuration between two arbitrarily shaped surfaces. ui and vi are the local parameterizations, $[xi, yi, zi]$ is the orthogonal coordinate frame on each body and Ψ is the angle between the xi's .

Proof See Appendix A.4.

The only thing still required in order to set up the dynamics equations for this configuration is to determine the derivative of the contact normal.

Lemma 3.3.2 *For the surface-surface contact, the derivative of the normal in world coordinates is*

$$\dot{N} = \omega_o \times N + H_o K_o M_o \dot{\alpha}_o, \tag{3.23}$$

where H_o is the orientation matrix , ω_o is the angular velocity of the body O with respect to the world, and N is the normal with respect to the world. K_o, M_o, α_o are the curvature tensor, the metric tensor and the parameterization of the surface on the body O. $\dot{\alpha}_o$ is obtained from the previous theorem.

Proof We have $N = H_o N_o$, where N_o is the normal in the coordinates associated with body O. We also have $\frac{dN_o}{d\alpha_o} = K_o M_o$ from the definition of the curvature tensor. But since

$$\dot{N} = \dot{H}_o N_o + H_o \dot{N}_o = \dot{H}_o H_o^T H_o N_o + H_o \dot{N}_o = \omega_o \times N + H_o \dot{N}_o, \tag{3.24}$$

the conclusion follows. □

3.3.1.2 The singularities

Clearly, the correct definition of the previous equations depends on the invertibility of the matrix $(K_f + \tilde{K}_o + d\tilde{K}_o K_f)$. For $d = 0$ we obtain the so-called "relative curvature" [17]. In this section we discuss the singularity to get some intuition about

Figure 3.3: A singular configuration for surface-surface contact

it.

If both bodies are locally strictly convex, the matrix is always positive definite, thus nonsingular. If both are locally strictly concave, then the pair of points cannot be a minimum-distance pair. Consider the case where one is strictly convex and one is strictly concave. To simplify , assume that the bodies are locally spherical, thus

$$K_o = \begin{pmatrix} \frac{1}{r} & 0 \\ 0 & \frac{1}{r} \end{pmatrix}, \tag{3.25}$$

$$K_f = \begin{pmatrix} -\frac{1}{R} & 0 \\ 0 & -\frac{1}{R} \end{pmatrix}. \tag{3.26}$$

Clearly, the matrix $(K_f + \tilde{K}_o + d\tilde{K}_o K_f)$ is singular if and only if

$$\frac{1}{r} - \frac{1}{R} - d\frac{1}{Rr} = 0; \tag{3.27}$$

or, multiplying by Rr, $R = r + d$. Thus the matrix is singular if and only if the two spheres are concentrical. But this means precisely that there are infinitely many minimum-distance pairs (any pair collinear with the common center. See Fig. 3.3) which confirms that the matrix and the configuration should be singular.

3.3.2 The curve-curve contact

3.3.2.1 The equations

The approach in setting up the curve-curve contact equations is essentially the same, except for the fact that in order to describe a contact point on a curve we need only one parameter, whereas two parameters were necessary in the surface case. To make up for the second parameter, we consider the angle of rotation around the local tangent to the curve.

Once again we consider two bodies whose respective center of mass frames are O and F respectively. On each body let C_o and C_f be the instantaneous frames of contact on each body. While for the surface-on-surface contact the definition was obvious, here we define C_o to be that frame that has the x axis along the tangent to the curve on body O and the z axis along the minimum distance line between the bodies. C_f is defined analogously. $L_o(\tau)$ and $L_f(\tau)$ respectively are the families of frames associated with each body, such that at time t $L_o(t)$ coincides with the Frenet trihedron [6] on the body O's curve. $L_f(t)$ coincides with the Frenet trihedron on body F's curve. ϕ_o is the angle defined by $L_o(t)$ and C_o, measured around the common x axis. Similarly, ϕ_f is the angle defined by $L_f(t)$ and C_f. Let d be the distance between the two bodies, which is measured along the z direction of the contact frames, and ψ the angle of rotation with respect to the contact frame around the z axis. ψ is chosen such that a rotation with angle ψ will make the x axes of C_o and C_f coincide.

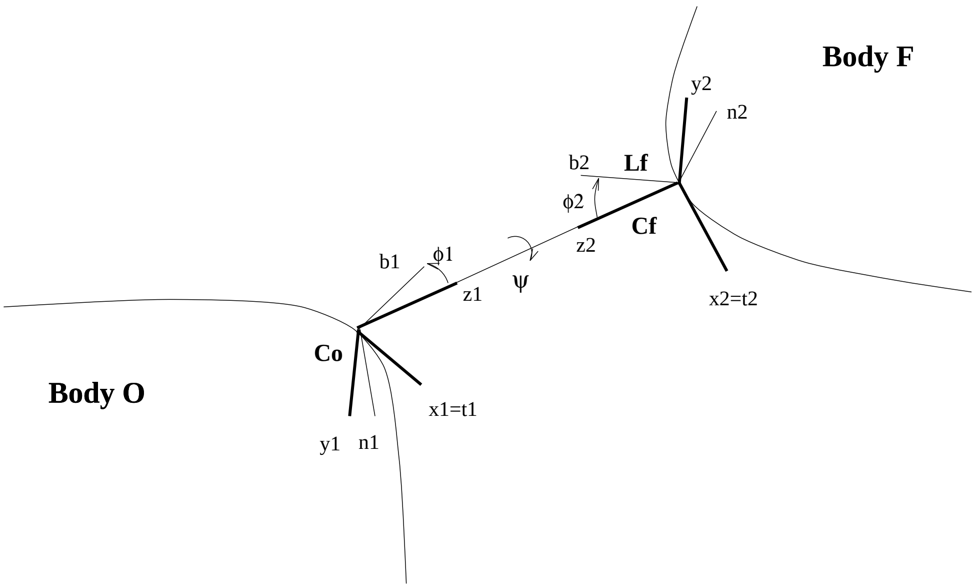

Figure 3.4: A curve curve configuration. $[ti, ni, bi]$ represents the natural frame for each curve (tangent, normal, binormal). The angle Ψ is the angle between the tangents, measured around the common perpendicular. $[xi, yi, zi]$ are chosen such that the z_i's are along the common perpendicular and the x_i's along the tangents. The angles ϕi measure the rotation between $[xi, yi, zi]$ and $[ti, ni, bi]$.

We assume that both curves are parameterized with respect to their natural parameter, the length, which allows for the Frenet description (see Fig. 3.4). We denote by $c_o(l_o)$ and $c_f(l_f)$ the contact points on each body in local coordinates.

Theorem 3.3.3 Kinematic equations of contact for the curve-curve configuration *The motion of the contact coordinates as a function of the relative motion* *is given by*

$$\begin{pmatrix} \dot{l}_f \\ 0 \end{pmatrix} + \begin{pmatrix} \tilde{v}_x \\ \tilde{v}_y \end{pmatrix} = R_\psi \begin{pmatrix} \dot{l}_o \\ 0 \end{pmatrix} - R_\psi \begin{pmatrix} -dk_o \sin\phi_o \dot{l}_o \\ d\tau_o \dot{l}_o + d\dot{\phi}_0 \end{pmatrix} \tag{3.28}$$

$$\begin{pmatrix} k_f \sin\phi_f \dot{l}_f \\ -\tau_f \dot{l}_f - \dot{\phi}_f \end{pmatrix} + \begin{pmatrix} \tilde{\omega}_y \\ -\tilde{\omega}_x \end{pmatrix} = R_\psi \begin{pmatrix} k_o \sin\phi_o \dot{l}_o \\ -\tau_o \dot{l}_o - \dot{\phi}_o \end{pmatrix} \tag{3.29}$$

$$\dot{d} = \tilde{v}_z \tag{3.30}$$

$$\dot{\psi} = \tilde{\omega}_z + k_o \cos\phi_o \dot{l}_o + k_f \cos\phi_f \dot{l}_f \tag{3.31}$$

where $(\tilde{v}_x, \tilde{v}_y, \tilde{v}_z, \tilde{\omega}_x, \tilde{\omega}_y, \tilde{\omega}_z) = Ad_{g^{-1}_{l_f c_f}} V^b_{l_o l_f}$, *(see Appendix A.7), and* $R_\psi = \begin{bmatrix} \cos\psi & -\sin\psi \\ -\sin\psi & -\cos\psi \end{bmatrix}$.

Proof See Appendix A.5. □

In the case of curve-on-curve constraint we have $N = \frac{T_1 \times T_2}{\|T_1 \times T_2\|}$ where N is the contact normal in the sense of (3.8) and T_1, T_2 are the tangents of the two curves involved in the contact, for an appropriate choice of the tangent directions. For setting up the dynamics equations, we need the derivative of the normal, given by the following result.

Lemma 3.3.4 *The derivative of the normal in world coordinates is*

$$\dot{N} = \frac{\dot{u}}{\|u\|} - \frac{u}{\|u\|^3}\langle \dot{u}, u \rangle \tag{3.32}$$

where

$$u = T_1 \times T_2 \tag{3.33}$$

$$\dot{u} = (\omega_o \times T_1) \times T_2 + T_1 \times (\omega_f \times T_2) + k_1 \dot{l}_o N_1 \times T_2 + k_2 \dot{l}_f T_1 \times N_2 \tag{3.34}$$

and ω_o and ω_f respectively are the angular velocities of the two bodies. T_1, N_1, k_1 and T_2, N_2, k_2 are the tangents, the normals and the curvatures of the curves involved in contact on bodies O and F respectively. $\dot{l}_o$ and $\dot{l}_f$ are the derivatives of the length of the curves in time, computed with the help of the previous theorem. All vectors are expressed in world coordinates.

Proof We have $T_1 = H_o T_{1o}$, $N_1 = H_o N_{1o}$, $T_2 = H_f T_{2f}$, and $N_2 = H_f N_{2f}$, where the additional subscripts o or f on the vectors mean that they are computed with respect to the local frames on each body. We then get

$$\dot{T}_1 = \dot{H}_o T_{1o} + H_o \dot{T}_{1o} = \dot{H}_o H_o^T H_o T_{1o} + H_o k_1 N_{1o} \dot{l}_o = \omega_o \times T_1 + k_1 N_1 \tag{3.35}$$

and similarly

$$\dot{T}_2 = \omega_f \times T_2 + k_2 N_2 \tag{3.36}$$

and the claim follows after taking the derivative of $N = \frac{u}{\|u\|}$. □

3.3.2.2 The singularities

For the important case of $d = 0$ we get the following specialization of the previous theorem:

Corollary 3.3.5 *For the case d=0 we have:*

$$\dot{l}_o = -\frac{\tilde{v}_y}{\sin\psi} \tag{3.37}$$

$$\dot{l}_f = -\tilde{v}_x + \cot\psi\tilde{v}_y \tag{3.38}$$

$$\dot{\phi}_o = -\frac{\tilde{\alpha}_x}{\sin\psi} \tag{3.39}$$

$$\dot{\phi}_f = -\tilde{\alpha}_y + \cot\psi\tilde{\alpha}_x \tag{3.40}$$

$$\dot{d} = \tilde{v}_z \tag{3.41}$$

$$\dot{\psi} = \tilde{\omega}_z + k_o\cos\phi_o\dot{l}_o + k_f\cos\phi_f\dot{l}_f \tag{3.42}$$

where

$$\begin{pmatrix} \tilde{\alpha}_x \\ \tilde{\alpha}_y \end{pmatrix} = \begin{pmatrix} \tilde{\omega}_y \\ -\tilde{\omega}_x \end{pmatrix} + \begin{pmatrix} k_f\sin\phi_f\dot{l}_f \\ -\tau_f\dot{l}_f \end{pmatrix} - R_\psi\begin{pmatrix} k_o\sin\phi_o\dot{l}_o \\ -\tau_o\dot{l}_o \end{pmatrix}. \tag{3.43}$$

We can see that for this case, a singularity occurs if $\sin(\psi) = 0$, which happens only when the tangents to the two curves coincide. This is an obvious result for the case the curves are straight lines. It is also intuitive in the general case, for in this case we cannot define a contact normal (which is the vector product of the two tangents).

3.3.3 The curve-surface contact

3.3.3.1 The equations

We now consider the case where the contact involves a curve on body F and a surface on body O. Define the contact frame C_f to be that frame that has the x axis along the tangent to the curve on body F and the z axis along the minimum distance line between the bodies. C_o is the same frame on body O as for the surface-surface

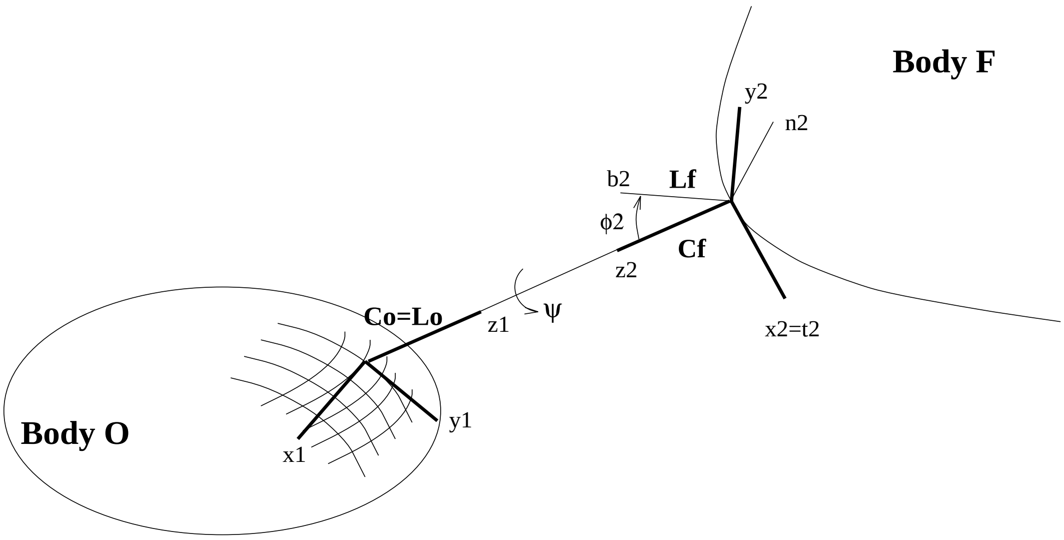

Figure 3.5: A curve-surface configuration

contact. $L_f(\tau)$ and $L_o(\tau)$ now have a clear meaning, with the former being the same as its homologue from the curve-curve case, and the latter the same as its homologue from the surface-surface case. Let ϕ_f be the angle defined by $L_f(t)$ and C_f, measured around the common x axis. d and ψ have exactly the same meaning as in the previous sections. With these conventions, the left-hand side of the equations will be the same as the one of curve-curve contact, whereas the right-hand sides will be the same as for surface-surface contact (see Fig. 3.5).

Theorem 3.3.6 Kinematic equations of contact for the curve-surface configuration *The motion of the contact coordinates as a function of the relative motion is given by*

$$\begin{pmatrix} \dot{l}_f \\ 0 \end{pmatrix} + \begin{pmatrix} \tilde{v}_x \\ \tilde{v}_y \end{pmatrix} = R_\psi (1 + dK_o) M_o \dot{\alpha}_o \tag{3.44}$$

$$\begin{pmatrix} k_f \sin\phi_f \dot{l}_f \\ -\tau_f \dot{l}_f - \dot{\phi}_f \end{pmatrix} + \begin{pmatrix} \tilde{\omega}_y \\ -\tilde{\omega}_x \end{pmatrix} = -R_\psi K_o M_o \dot{\alpha}_o; \tag{3.45}$$

$$\dot{d} = \tilde{v}_z \tag{3.46}$$

$$\dot{\psi} = \tilde{\omega}_z + T_o M_o \dot{\alpha}_o + k_f \cos\phi_f \dot{l}_f \tag{3.47}$$

where $(\tilde{v}_x, \tilde{v}_y, \tilde{v}_z, \tilde{\omega}_x, \tilde{\omega}_y, \tilde{\omega_z}) = Ad_{g^{-1}_{l_f c_f}} V^b_{l_o l_f}$ *(see Appendix A.7) and* $R_\psi = \begin{bmatrix} \cos\psi & -\sin\psi \\ -\sin\psi & -\cos\psi \end{bmatrix}$.

Proof See Appendix A.6. □

For this case the derivative of the normal is exactly the same as for the surface-surface contact (Lemma 3.3.2), with the difference that $\dot{\alpha_o}$ is now computed using Theorem 3.3.6.

3.3.3.2 The singularities

In order to discuss the singularities, we examine the case where O has a spherical shape and radius R. We also define $\beta = R_\psi \dot{\alpha}_o$. The evolution equations become:

$$\begin{pmatrix} \dot{l}_f \\ 0 \end{pmatrix} + \begin{pmatrix} \tilde{v}_x \\ \tilde{v}_y \end{pmatrix} = (1 + \frac{d}{R})\beta \tag{3.48}$$

$$\begin{pmatrix} k_f \sin\phi_f \dot{l}_f \\ -\tau_f \dot{l}_f - \dot{\phi}_f \end{pmatrix} + \begin{pmatrix} \tilde{\omega}_y \\ -\tilde{\omega}_x \end{pmatrix} = \frac{1}{R}\beta; \tag{3.49}$$

From the structure of the system, it is obvious that the corresponding matrix is invertible if and only if

$$(1 + d/R) \neq 0, \tag{3.50}$$

$$1/R - k_f \sin\phi_f (1 + d/R) \neq 0. \tag{3.51}$$

The former ensures that β_y can be solved for, whereas the latter ensures that $\dot{l}_f$ and β_x can be solved for. Let $\tilde{r} = -\frac{1}{k_f \sin\phi_f}$ be the effective curvature radius of the curve. Then the previous relations can be rewritten as

$$d \neq -R, \tag{3.52}$$

$$\tilde{r} \neq -(R + d). \tag{3.53}$$

We have different types of singular behavior:

- $d = -R$ (with the spherical region on O concave). All points on O within the spherical region are at the same distance from the contact point on F. Thus the configuration must be singular (and there cannot be a unique contact pair).

- $\tilde{r} = -(d+R)$ is the same type of singularity as for the surface-surface constraint (3.27). Thus the discussion for the singularities of the surface-surface case applies. The only difference is that the convexity of the curve is discussed in the plane parallel to the tangent. Intuitively, for $d = 0$ this condition means that the effective curvature of the curve has to be different from the curvature of the surface along the tangent to the curve. Otherwise, the curve will have an excessively flat contact with the surface.

3.3.4 Setting up the dynamics equations

We assume that the contacts i, $1 \leq i \leq m$, are of one of the types analyzed in the previous subsections or that the function $d_i = \phi_i(q)$ can be explicitly determined (as is the case for point-on-plane, edge-on-edge or even point-on-surface for an explicitly defined surface). The latter contacts have the additional advantage that the contact points can be easily determined. The contacts analyzed in the previous subsections are described by functions η_i of six variables. For instance, in the curve-surface case $\eta_i = (u_o, v_o, l_f, \phi_f, \psi, d)$, where (u_o, v_o), is a parameterization of the surface on body O and the rest of the parameters are introduced in Section 3.3.3. The analysis gives us $\dot{\eta}_i$ as a function of η_i, the positions, and the velocities of the bodies involved in contact, say

$$\dot{\eta}_i = \nu_i(\eta_i, q, \dot{q}), \tag{3.54}$$

where by q and $\dot{q}$ we denote the positions and velocities of all the bodies in the system. If η_i is known, then it is very simple to determine the position of the contact points, l_{1i} and l_{2i}, and the associated normal, N. For instance, in the curve-surface case we have $l_{1i} = c_o(u_o, v_o)$, the local parameterization of the surface, and $l_{2i} = c_f(l_f)$, the local parameterization of the curve, as introduced in the previous subsections. For the normal we have $N = H_o N_o(u_o, v_o)$, where H_o is the rotation matrix for the body O, and $N_o(u_o, v_o)$ is the mapping that gives the exterior normal in local coordinates. Therefore we also have $\dot{l}_{1i}, \dot{l}_{2i}, \dot{N}_i$, by the chain rule, as functions of $\eta_i, \dot{\eta}_i, q, \dot{q}$. Thus we have the functions

$$\dot{l}_{1i} = \chi_{1i}(\eta_i, \dot{\eta}_i, q, \dot{q}) \tag{3.55}$$

$$\dot{l}_{2i} = \chi_{2i}(\eta_i, \dot{\eta}_i, q, \dot{q}) \tag{3.56}$$

$$\dot{N}_i = \chi_i(\eta_i, \dot{\eta}_i, q, \dot{q}). \tag{3.57}$$

Hence, having the positions and velocities of all the bodies involved in the simulation as well as the η_i's , we are able to compute $\dot{l}_{1i}, \dot{l}_{2i}, \dot{N}_i$, for $1 \leq i \leq m$. By using this data in Corollary 3.2.2 we can compute Φ_q and Γ_2 from (3.6). Since the terms Θ_q and Γ_1 can be easily computed from the joint type constraints (which are usually represented by explicit functions), the LCP (3.6) is now completely defined.

Next, assume that we are able to solve the LCP and get the accelerations. At that moment we can take an integration step and determine the new positions and velocities of the bodies. The problem is that at this new position we need to start an optimization algorithm for determining the new contact points. But we could also use (3.54) to integrate η_i which will give us the position of the new contact points, since we use only information readily available. Therefore the integration process can be implemented as follows:

- Start with q,$\dot{q}$ and assume that the η_i's are available.
- Set up and solve the linear complementarity problem (3.6, 3.7) for the accelerations $\ddot{q}$. To generate the right hand side, use (3.54) and then (3.55),(3.56) and (3.57) in Corollary 3.2.2. Each the functions $\nu_i, \chi_{1i}, \chi_{2i}, \chi_i$ is obtained from the

analysis of the geometry of each contact.

- Get the accelerations from the LCP and integrate to get the new q, $\dot{q}$.

- Integrate (3.54) to get η_i $1 \leq i \leq m$, at the new step, which gives us the new contact points. At this stage the process can be restarted.

3.4 Arbitrary-shaped planar contacts

3.4.1 The affine relation between various representations of a planar constraint

In the following, we investigate the case where the contact region is planar but has a boundary of arbitrary, not necessarily polyhedral, shape. (e.g. a cylindrical glass on a table). If the contact region were a polygon, we would simply set up the contact constraint by requiring that each vertex of the polygon stay in contact with the face on the other body. This approach has two disadvantages. First, if the contact region has more than three vertices, we will introduce a singularity in the matrix of (3.6) (because no matter how many points we introduce as a constraint, the linear subspace spanned by the corresponding columns has rank at most three, as we will see). Second, if the polygon has many edges, we can get a correspondingly large LCP. Furthermore, this approach cannot be directly applied to situations where the contact region is a closed curve since we would need to consider infinitely many constraints.

We describe how to circumvent these problems for the special case of frictionless contact and nonsingular matrix of (3.6) by using just three points of the contact region to represent the planar contact. We also assume that the external forces are continuous. In the following analysis, we confine ourselves to the Newton-Euler representation.

If we define $\ddot{q}_{12} = \{\ddot{p}_1, \dot{\omega}_1, \ddot{p}_2, \dot{\omega}_2\}^T$, then the constraint that a specific point on the planar contact region of one body is in contact with the other body is expressed as

$$J_{12}\ddot{q}_{12} + \Gamma_{12} = 0 \tag{3.58}$$

where Γ_{12} represent Coriolis-type accelerations and J_{12} is given by (3.10) with N the normal to that plane, and l_1, l_2 the local coordinates of the contact point with respect to the two bodies.

Lemma 3.4.1 *If x_1, x_2, x_3 are non-collinear points on the planar contact region and $x_4 = \sum_{i=1}^{3} \lambda_i x_i$ with $\sum_{i=1}^{3} \lambda_i = 1$, then $J_{12}(x_i)$, $i = 1, 2, 3$ are linearly independent, and*

$$J_{12}(x_4) = \sum_{i=1,3} \lambda_i J_{12}(x_i) \tag{3.59}$$

$$\Gamma_{12}(x_4) = \sum_{i=1,3} \lambda_i \Gamma_{12}(x_i). \tag{3.60}$$

Proof The linear independence follows from the form of the Jacobian (3.10). Any zero linear combination of these three Jacobians will have to have $\sum_{i=1,3} \lambda_i = 0$ in order to cancel the first component (3.10). But then $N \perp \sum \lambda_i x_i$ which makes it impossible to cancel the second component if any λ is nonzero and the points x_i non-collinear.

The only way J_{12} depends on x is through l_1 and l_2. But the dependency is clearly affine in x, since $l_1(x)$ and $l_2(x)$ are both affine functions, which proves (3.59) . On the other hand, the Coriolis accelerations depend only on the velocities, thus even if we assume that the planar constraint is bilateral, the Coriolis accelerations will be the same if we have contact at that specific moment. For bilateral constraints, we have

$$0 = J_{12}(x_4)\ddot{q}_{12} + \Gamma_{12}(x_4) = \sum_{i=1,3} \lambda_i J_{12}(x_i) + \Gamma_{12}(x_4) = \tag{3.61}$$

$$- \sum_{i=1,3} \lambda_i \Gamma_{12}(x_i) + \Gamma_{12}(x_4). \tag{3.62}$$

which proves (3.60) and the Lemma. □

Corollary 3.4.2 *Let y_i, $1 \leq i \leq n$ be a set of points on the base and μ_i, $1 \leq i \leq n$, a set of forces such that $\sum_{i=1}^n \mu_i > 0$. Then if the point $x = \frac{\sum_{i=1}^n \mu_i y_i}{\sum_{i=1}^n \mu_i}$ (which is a point on the contact plane) is inside the convex hull of the contact region, there are three points in the contact region x_1, x_2, x_3 and three positive forces $\lambda_1, \lambda_2, \lambda_3 > 0$ such that*

$$\sum_{i=1}^{n} \mu_i J_{12}(y_i) = \sum_{j=1}^{3} \lambda_j J_{12}(x_j). \tag{3.63}$$

Proof Let $\nu_i = \frac{\mu_i}{\sum_{i=1}^n \mu_i}$. Then $\sum_{i=1}^n \nu_i = 1$, thus $x = \sum_{i=1}^n \nu_i y_i$ is a point on the common contact plane (since the latter is an affine space). If x is in the interior of the convex hull of the contact region it follows that there are x_j, $j = 1, 3$ in the contact region such that $x = \sum_{j=1}^3 \theta_j x_j$ with $\theta_j > 0$. By the Lemma 3.4.1 we have:

$$\sum_{i=1}^{n} \nu_i J_{12}(y_i) = J_{12}(x) = \sum_{i=1}^{3} \theta_j J_{12}(x_j) \tag{3.64}$$

and the conclusion of the corollary follows after multiplying the latter relation with $\sum_{i=1}^{n} \mu_i$ and taking $\lambda_j = \theta_j(\sum_{i=1}^{n} \mu_i),\ j = 1, 3.$ $\square$

It is easy to see that there is no restriction on the sign of the μ_i's in the previous lemma, which leads to the idea that the contact constraint for this case (arbitrary-shaped contact regions) can be viewed as a planar joint constraint, expressed, for example, as three point-on-plane bilateral constraints.

We now use the assumption from the beginning of the section that the matrix in (3.6) is invertible and that the external forces are continuous. We see that forces corresponding to these three bilateral constraints can be computed and are continuous functions; thus, the x obtained by the procedure from the proof of the previous corollary is a continuous function. As long as this point is inside the convex hull of the contact region, the planar contact will still be acting during the next integration step, because a solution to the LCP (3.6) has been found which has only positive forces and zero left hand side. When x moves outside this convex hull the contact will break, and the continuity of x allows us to decide what type of constraint to impose on the bodies after the planar constraint breaks.

It is not necessary to require that the matrix from (3.6) be nonsingular. But, we need some procedure to ensure that the forces from (3.6) are continuous (see for instance [3]), for otherwise it becomes difficult to determine the way the contact will break.

3.4.2 Integration strategy for arbitrary-shaped

planar contacts

We assume that the bodies involved in the contact have been in contact for some time and that at the beginning of the simulation we will obtain solutions consistent with the contact being still in effect. For solving the dynamics, we adopt techniques from both the previous subsection and Section 3.3 to get the following strategy:

- We take three points on the boundary of the bottom of the contact region and solve for the accelerations, by using the linear system setting, not the LCP (set the left hand side of (3.6) to zero and solve for the forces and accelerations). This is possible, since the system matrix is now nonsingular. Among other data, we get $\mu_i,\ i = 1,3$, the forces associated to the three point on plane constraints.
- If some of the μ's are negative, we find the point x from Corollary 3.4.2 and check whether the point is inside the convex hull of the contact region (x is expressed in local coordinates).
  - If yes, that means that there is a choice of three points on the contact regions' convex hull such that $x = \sum_{i=1,3} \lambda_i x_i$ with $\lambda_i > 0$, by the same Corollary. Following (3.59) if we replace the old three point-plane constraints with the point-plane constraints associated with these new points, the forces will be positive and the contact will persist.
  - If no, we estimate by interpolation where x has crossed the boundary of the convex hull of the contact region (because the system matrix is

nonsingular, x is continuously differentiable with respect to the time) and restart the simulation there, using a curve- plane constraint initialized at x (or something similar if the boundary of the contact region contains a line segment).

To use Corollary 3.4.2 we need to have $\sum_{i=1}^{3} \mu_i > 0$. This is a reasonable assumption for most exterior forces. However, this is not true for example for a force that will lift one body from the other. For this case we need to adopt a different strategy once we have noticed that $\sum_{i=1}^{3} \mu_i = 0$. But for most cases x will get out of the convex hull of the planar contact before $\sum_{i=1}^{3} \mu_i = 0$ and the previous strategy will work.

3.4.3 Implementation details

We simulated the dynamics of a cylindrical glass on a table, subject to an external force and torque. The force and torque were defined such that the contact region was the bottom of the glass for a while at the beginning of the simulation, after which glass began to rotate and slide on the contour of its circular base. Four frames from the simulation are shown in Figure 3.6.

In the first stage the strategy from the previous subsection is applied. When the dynamics is no longer consistent with the contact region being the bottom of the glass, the point-on-plane constraints are replaced by a curve-on-surface constraints.

A problem is that the configuration is singular for the curve on surface constraint when the glass is on the bottom, since ϕ_f, the angle between the normal of the plane and the binormal is exactly π. Thus the condition (3.51) fails for $R = \infty$ (the

Figure 3.6: Four frames from a simulation of a cylindrical object rolling and sliding moving on a tabletop.

surface is a plane) and $\phi_f = \pi$ (and this is true for any shape of the bottom of the glass). We dealt with this by requiring that an imaginary edge on the glass bottom boundary, centered at the intended contact point, stay in contact with the plane until the glass achieves a pitch of at least .01 radians (which is sufficiently far from the singularity). From that moment on, we solved the dynamics via Theorem 3.3.6. For this particular case, we could have done away with the contact equations, because the distance between a circle and a plane can be explicitly computed. Nevertheless it was a useful test of the correctness for our equations.

CHAPTER 4

ON THE EXISTENCE OF SOLUTIONS TO COMPLEMENTARITY FORMULATIONS OF CONTACT PROBLEMS WITH FRICTION

It has been established that acceleration–force based models for multi-body-dynamics with friction do not always have a solution [19] although the frictionless configurations always have a solution. To understand the nature of frictional inconsistency for certain configurations it is important to understand which configurations are solvable for small coefficients of friction (are frictionally robust). This section demonstrates that these configurations are the ones that can be disassembled.

4.1 The assumptions and setting of the problem

4.1.1 The contact forces

Assume that we have two bodies in 3-dimensional space and that the system configuration can be represented by a vector $q \in R^{12}$ such that half of its coordinates account for translation and half for rotation. Assume that the bodies are in contact, as shown in Figure 4.1. Let N be the direction of the reaction force (thus, we will assume that $||N|| = 1$), T_1 and $T_2 = -T_1$ the directions of the friction force acting on the first and second body respectively, G_1 and G_2 the locations of the centers of mass of the bodies, and l_1 and l_2 the locations of the contact points in local coordinates (*i.e.*, with respect to G_1 and G_2). Let λ be the magnitude of the reaction force. Let H_1 and H_2 be the bodies' rotation matrices with respect to the world. We assume that the bodies are in relative motion; thus we consider dynamic friction only, and the direction of the friction is determined based on the direction of the relative velocity of the bodies. Also, we assume a Coulomb friction model, and let μ represent the friction coefficient.

Since the equations of motion are usually written with respect to the centers of mass of the bodies involved in the simulation, the force must be expressed with respect to the centers of mass. The reaction force λN at the contact point generates the force $-\lambda N$ and the torque $-H_1 l_1 \times \lambda N = -\lambda H_1 l_1 \times N$ with respect to the center of mass G_1, as well as the force λN and the torque $\lambda H_2 l_2 \times N$ with respect to the center of mass G_2. Similarly, the friction force generates a force $\mu\lambda T_1$ and a torque $\mu\lambda H_1 l_1 \times T_1$ with respect to the center of mass G_1, as well as a force $\mu\lambda T_2$ and a torque $\mu\lambda H_2 l_2 \times T_2$ with respect to the center of mass G_2. Anticipating the generalized setting, we introduce the following vectors

$$R = \begin{pmatrix} -N \\ -H_1 l_1 \times N \\ N \\ H_2 l_2 \times N \end{pmatrix} \tag{4.1}$$

and

$$F = \begin{pmatrix} T_1 \\ H_1 l_1 \times T_1 \\ T_2 \\ H_2 l_2 \times T_2 \end{pmatrix}. \tag{4.2}$$

Then we can say that the 12-dimensional vectors λR and $\mu\lambda F$ represent the generalized reaction and friction forces for this two-body configuration.

4.1.2 The kinematic contact constraint

To prevent interpenetration between bodies, we require that $d \geq 0$, where d is a distance measure as defined in [1]. This d is defined such that when there is no interpenetration d is positive and corresponds to the interbody distance. When there is interpenetration, d is negative, and when objects are in perfect contact, d is zero. In the case where $d = 0$ and we have a point contact (as in Figure 1) we require a kinematic constraint that, essentially, prohibits translation in the direction of N. To express that restriction in a workable fashion, it is customary to work in a Newton-Euler setting. Let v_1, ω_1, v_2, and ω_3 be the translational and rotational velocities of the first and second bodies, expressed in world coordinates. Then it can be shown (see, for instance, [1]) that

$$\dot{d} = (v_2 - v_1)^T N + (H_2 l_2 \times N)^T \omega_2 - (H_1 l_1 \times N)^T \omega_1. \tag{4.3}$$

For the second derivative, we get

$$\ddot{d} = (\dot{v}_2 - \dot{v}_1)^T N + (H_2 l_2 \times N)^T \dot{\omega}_2 - (H_1 l_1 \times N)^T \dot{\omega}_1 + \gamma, \tag{4.4}$$

where γ is a number that depends on the bodies' velocities and positions (but not the accelerations). Let

$$a = \begin{pmatrix} \dot{v}_1 \\ \dot{\omega}_1 \\ \dot{v}_2 \\ \dot{\omega}_2 \end{pmatrix} \tag{4.5}$$

be the acceleration vector. Then, using R from (4.1), we can rewrite (4.4) as

$$\ddot{d} = R^T a + \gamma. \tag{4.6}$$

4.1.3 The complementarity condition

Assume that the two bodies have been in contact for some time. Therefore $d = 0$ and $\dot{d} = 0$. Since we require that $d \geq 0$ it follows that we must ensure $\ddot{d} \geq 0$ in order to avoid interpenetration in a subsequent instant. If the forces have the directions shown in Figure 1, it follows that $\lambda \geq 0$ since the two bodies may push but not pull on each other (*i.e.* the contact can only sustain compressive force). This is the main difference between contact and traditional joint constraints which can assume both positive and negative reaction forces. Thus, contact constraints are sometimes called *unilateral* constraints, while joint constraints are termed *bilateral* constraints.

Assume now that $\ddot{d} > 0$. This means that in the following instant the contact between the two bodies will break. As a result, the mutual force must vanish, yielding

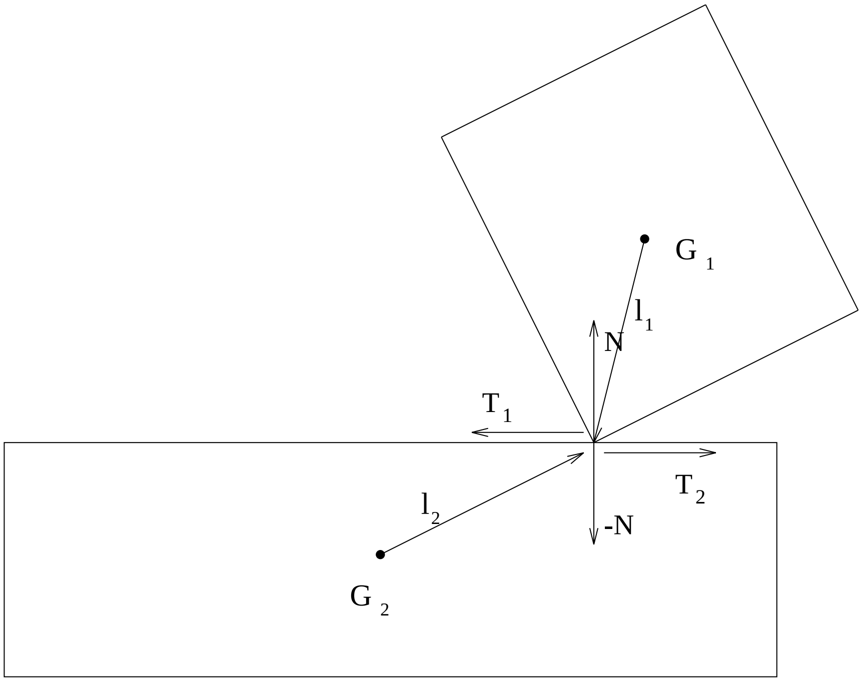

Figure 4.1: Generic contact situation between two bodies. N is the direction of the reaction force. T_1 and $T_2 = -T_1$ are the directions of the frictional force. G_1 and G_2 are the locations of the centers of mass, l_1 and l_2 are the locations of the contact points with respect to the centers of mass.

$\lambda = 0$. Combining the two cases gives

$$\ddot{d}\lambda = 0. \tag{4.7}$$

Two nonnegative variables that satisfy such a condition are said to be complementary variables.

4.1.4 The contact problem for two bodies

For the subsequent derivations we define $s = \ddot{d}$. Then by (4.7) we must have $s^T\lambda = 0$ and by (4.6) we must have $s = R^T a + \gamma$, where both s and λ are nonnegative. Now we can write Newton's law, which connects acceleration and forces:

$$Ma = Q + \lambda R + \mu\lambda F, \tag{4.8}$$

where M is the generalized mass matrix for the two bodies, a block diagonal matrix composed of 3-by-3 blocks. The blocks corresponding to $\dot{v}_1$ and $\dot{v}_2$ are diagonal having the corresponding masses as entries. The blocks corresponding to $\dot{\omega}_1$ and $\dot{\omega}_2$ are the inertia matrices written in world coordinates. Q is the sum of the external and Coriolis forces, the latter being computed once the positions and velocities are known. Finally λR and $\mu\lambda F$ are the contact reaction and friction forces.

The main problem in dynamics is to find the accelerations given the positions and velocities of the bodies involved in the contact. This problem can be restated as mixed linear complementarity problem of the form: Find (a, s, λ) such that

$$\begin{pmatrix} 0 \\ s \end{pmatrix} = \begin{pmatrix} M & -(R+\mu F) \\ R^T & 0 \end{pmatrix} \begin{pmatrix} a \\ \lambda \end{pmatrix} + \begin{pmatrix} -Q \\ \gamma \end{pmatrix} \tag{4.9}$$

$$s^T\lambda = 0; \quad s, \lambda \geq 0 \tag{4.10}$$

4.1.5 The general setting of the contact problem

Suppose that we have a system of several bodies along with a set of m contacts between the bodies. For each pair of bodies in contact use an analysis similar to that above. Each contact i, $1 \le i \le m$ will generate an s_i (slack) variable, a row R_i^T, and a generalized force $\lambda_i(R_i + \mu F_i)$. When the friction coefficients are not the same for each i, the F_i's should be scaled accordingly. For completeness, assume that the system has, in addition to contact constraints, l joint-type constraints defined by the map $\Theta(q)$ taking the configuration space to R^l. Here by q we mean the generalized position of the system of bodies, which can be written for instance as the concatenation of the translation and rotation coordinates of the bodies. We will also denote by $\dot{q}$ and $\ddot{q}$ the generalized positions and velocities. The constraints can be written as $\Theta(q) = 0$. For details concerning joint constraints and forces see, for example, [9].

By writing the joint and contact constraints together with Newton's law and the appropriate constraint forces we get the following equation

$$\begin{pmatrix} 0 \\ 0 \\ s \end{pmatrix} = \begin{pmatrix} M & -\Theta_q & -(R+\mu F) \\ \Theta_q^T & 0 & 0 \\ R^T & 0 & 0 \end{pmatrix} \begin{pmatrix} \ddot{q} \\ \nu \\ \lambda \end{pmatrix} + \begin{pmatrix} -Q \\ \Gamma_1 \\ \Gamma_2 \end{pmatrix}. \tag{4.11}$$

where $s_i = \ddot{d}_i$, $1 \le i \le m$ is the vector of contact accelerations. Q is the sum of input forces and the Coriolis forces, Γ_1 and Γ_2 are terms involving the velocities $\dot{q}$. R is the matrix whose columns are the R_i's, and F is the matrix whose columns are the F_i's. M is the mass matrix which is positive definite. The dimension of M is $n \times n$, where

n is the numbers of degrees of freedom of the system if there were no constraints (in three dimensions $n = 6k$, where k is the number of the bodies in the system). $\nu\Theta_q$ is a vector of the joint reaction forces. We can assume without loss of generality that Θ_q has full row rank, for otherwise we can reduce the dependent rows and deal with a smaller set of ν's.

For each contact the complementarity relation (4.7) must hold. Thus, $s_i\lambda_i = 0,\ 1 \le i \le m$. By adding up the relations (3.2) for all the contact constraints, we get

$$s^T\lambda = 0; \quad s, \lambda \ge 0. \tag{4.12}$$

The relations (4.12) and (4.11) determine a mixed linear complementarity problem [5].

4.2 Disassemblability

4.2.1 Definition of disassemblability

It is known that in the frictionless case ($\mu = 0$) the former linear complementarity problem has very good properties. For instance, the matrix on the left hand side of (4.11) becomes positive semidefinite, and therefore the mixed LCP is equivalent to a traditional monotone LCP (in which no equality constraints are involved) [2, 25]. For such an LCP, if nonnegative s and λ exist so that (4.11) is satisfied, then a complementary solution (satisfying (4.12)) exists [5]. Moreover it is easy to see that the accelerations $\ddot{q}$ and the slacks s are unique for any solution of the LCP. In this section, we will show that under mild assumptions the friction problem inherits some of the good behavior of the frictionless problem for a sufficiently small μ. We will consider first the frictionless case and discuss later the impact of having a nonzero friction coefficient μ.

It is useful to rewrite (4.11) such that the accelerations do not appear in the new formulation. We solve for the accelerations in the first row of (4.11) and then substitute for them in the other rows. We then get an equivalent mixed LCP

$$\begin{pmatrix} s \\ 0 \end{pmatrix} = \begin{pmatrix} R^T M^{-1} R & R^T M^{-1} \Theta_q \\ \Theta_q^T M^{-1} R & \Theta_q^T M^{-1} \Theta_q \end{pmatrix} \begin{pmatrix} \lambda \\ \nu \end{pmatrix} + \begin{pmatrix} R^T M^{-1} Q + \Gamma_2 \\ \Theta_q^T M^{-1} Q + \Gamma_1 \end{pmatrix}. \tag{4.13}$$

$$s^T \lambda = 0; \; s, \lambda \geq 0. \tag{4.14}$$

We would like to determine if there exists an external force Q such that (4.13) (hence (3.6)) is satisfied and, at the same time, the slacks are strictly positive. More specifically, we want to determine whether the polytope

$$\mathcal{P}_1 = \left\{ s \geq 0 | \; \exists \chi \in R^n, \; s = R^T \chi; \; \Theta_q^T \chi = 0 \right\} \tag{4.15}$$

has an interior point $s_0 > 0$ with a corresponding χ_0. When it does, let $\nu = -(\Theta_q^T M^{-1} \Theta_q)^{-1} \Gamma_1$, which is well-defined since M is positive definite and Θ_q has full row rank. And let $Q = tM\chi_0$ for t such that $tR^T \chi_0 + \Gamma_2 + R^T M^{-1} \Theta_q \nu = ts_0 + \Gamma_2 + R^T M^{-1} \Theta_q \nu > 0$. Then, (λ, ν, s) with $\lambda = 0$, ν defined as above and $s = ts_0 + \Gamma_2 + R^T M^{-1} \Theta_q \nu > 0$ is a solution of the LCP (4.13). Hence applying the force Q effects the immediate breakage of all contacts. The system can thus be disassembled with respect to the contact constraints (while remaining consistent with respect to the joint-type constraints). For this reason, a configuration for which the polytope $\mathcal{P}_1$ has an interior point (*i.e.*, (4.13) has a solution with strictly positive s's) is called *disassemblable.*

For s_0, ν, and Q as defined above, $s_0 > 0$, $\lambda = 0$ and ν constitute a solution of the LCP (4.13). If the system would have had friction, the only change would be that R would be replaced by $R + \mu F$ everywhere it appears to the right of M^{-1} in (4.13). It is easy to see that s_0, $\lambda = 0$, ν would be a solution of the problem with friction as well. Thus, both the friction and the frictionless configurations can be disassembled by the Q chosen above. Hence, friction is irrelevant for the disassemblability property.

Note, however, that disassemblability is perhaps less well defined defined in the friction case since for the LCP with friction, there is no guarantee, even when a solution exists, that the s's are unique. This leaves opened the possibility that a different solution exists for the same Q that does not disassemble the configuration. Below, we will consider, without loss of generality, just the case $\mu = 0$ when discussing disassemblability, but will account for the case $\mu \neq 0$ when looking for solutions of the frictional versions of the LCP (4.13).

4.2.2 Characterization of nondisassemblability

Assume that a frictionless contact configuration is not disassemblable. Then, the polytope $\mathcal{P}_1$ has no interior point and there is no χ exists such that:

$$s = R^T \chi > 0 \tag{4.16}$$

$$\Theta_q^T \chi = 0. \tag{4.17}$$

Now, let I be the set of indices i for which $\exists s \in \mathcal{P}_1$ with positive ith component ($s_i > 0$). For each $i \in I$, let s^i be a point satisfying the above property (i.e.,

$s_i^i > 0$) and let $\chi^i \in R^n$ the corresponding χ from 4.16. It follows that if we take $s^{pos} = \sum_{i \in I} s^i$, then $s_i^{pos} > 0, \forall i \in I$. And, clearly, $s^{pos} \in \mathcal{P}_1$. Thus if the only contacts were those corresponding to the inequalities i for $i \in I$, we would have an interior point. The "problem" index set is thus the complement of I with respect to the total number of inequalities and we call it $\overline{I}$. Following its definition we have that for any $s \in \mathcal{P}_1$, $s_i = 0$ for all $i \in \overline{I}$. Let χ^{pos} satisfy $\Theta_q^T \chi^{pos} = 0$ and $s^{pos} = R^T \chi^{pos}$. Such a χ^{pos} can be obtained by taking $\chi^{pos} = \sum_{i \in I} \chi^i$. Also, let W be the matrix consisting of those columns of R whose indices are in $\overline{I}$. Consider now the polytope:

$$\mathcal{P}_2 = \left\{ w \geq 0 | \quad \exists \chi \in R^n \quad w = W^T \chi, \quad \Theta_q^T \chi = 0 \right\}. \tag{4.18}$$

Lemma 4.2.1 $\mathcal{P}_2 \equiv \{0\}$.

Proof Assume there is a $w \in \mathcal{P}_2$ containing some positive elements. Then, there must be a corresponding χ_1 such that $w = W^T \chi_1$ and $\Theta_q^T \chi_1 = 0$. By taking $\tilde{\chi} = \chi_1 + \delta \chi^{pos}$, with δ sufficiently big, it follows that

$$\Theta_q^T \tilde{\chi} = 0 \tag{4.19}$$

$$R^T \tilde{\chi} \geq 0. \tag{4.20}$$

Thus, we have a point $s^+ = R^T \tilde{\chi} \geq 0$. By construction, we have that $s_i^+ > 0$ for $i \in I$, and $s_j^+ > 0$ for some $j \in \overline{I}$, (where the corresponding component of w is positive). But this contradicts the definition of I, and so the proof is complete. $\square$

Lemma 4.2.2 *There is a vector $\eta \in R^m$ such that $\eta_i = 0$ for $i \in I$ and $\eta_i > 0$ for $i \in \overline{I}$ and a vector $\zeta \in R^l$ such that*

$$R\eta + \Theta_q \zeta = 0 \tag{4.21}$$

Proof Without loss of generality, we may assume that $\Theta_q^T = [I, -N]$. Let $W^T = [A_1, A_2]$ be the corresponding partition of W^T and $\chi = [\chi_1^T, \chi_2^T]^T$ the partition of χ, where $\chi \in R^n$. Then $\chi_1 = N\chi_2$ for χ in accordance with the definition of the polytope 4.18. The fact that $\mathcal{P}_2 \equiv \{0\}$ can be now written as

$$A_1 N\chi_2 + A_2\chi_2 \geq 0 \Rightarrow A_1 N\chi_2 + A_2\chi_2 = 0 \tag{4.22}$$

and thus the system

$$A_1 N x_2 + A_2 x_2 \geq 0, \quad A_1 N x_2 + A_2 x_2 \neq 0 \tag{4.23}$$

has no solution. By Stiemke's (alternative) theorem [5, Thm.2.7.12] it follows that there is an $\eta' > 0$ such that $A_2^T\eta' + N^T A_1^T \eta' = 0$. By taking $\zeta = -A_1^T\eta'$, it follows immediately that $W\eta' + \Theta_q\zeta = 0$. But now by taking the entries of η to be the same as those of η' for the indices in $\overline{I}$ and zero otherwise, we obtain (since W is the restriction of R on the index set $\overline{I}$) that $R\eta + \Theta_q\zeta = 0$ and $\eta_i > 0$ for $i \in \overline{I}$. $\square$

Assume that $(\ddot{q}, \lambda, \nu)$ is a solution of (3.6). Then clearly $(\ddot{q}, \lambda + t\eta, \nu + t\zeta)$ is also a solution of (3.6) for any $t > 0$. This means that the bodies can assume an infinite reaction force (when taking $t \to \infty$). This behavior is likely to cause problems when introducing friction; since friction forces are proportional to the reaction forces, an indeterminacy in reaction forces can yield indeterminacy in the friction forces and, therefore, in the accelerations as well.

The previous development also makes a new point. Since there is a part of the system that cannot be disassembled (contacts for which the corresponding slacks s_i are always zero) we can treat those particular contacts as bilateral constraints. Thus, assume that W, the columns of R corresponding to $\overline{I}$, is taken into account as columns of Θ_q, and the reaction force splits as $\nu = [\nu_1, \nu_2]$, where ν_2 represents the reactions in the old bilateral constraints and ν_1 represent the reaction force in the nondisassemblable contact constraints. This new configuration is disassemblable, and the positive entries of s^{pos} are the corresponding positive slacks. Now let η^{pos} be the positive entries of η from the previous lemma corresponding to the index set $\overline{I}$. And let (λ, ν, s) with $\nu = [\nu_1, \nu_2]$ be a solution of (3.6) corresponding to the new R and Θ_q. Since the nondisassemblable constraints are treated as bilateral constraints, there is no guarantee that ν_1 is positive. By our previous developments $(\lambda, \nu_1 + t\eta^{pos}, \nu_2 + t\zeta)$ is also, for any t, a solution of the LCP (3.6) with the new data. But for a sufficiently big t we get that $\nu_1 + t\eta^{pos} > 0$, and therefore we recover a solution to the initial LCP. Hence we have proved the following result.

Theorem 4.2.3 *A frictionless contact problem can always be cast as a contact problem with disassemblable contact constraints, consistent with the bilateral constraints.*

4.2.3 Disassemblability and copositivity

Since we may assume that Θ_q has full rank it follows that $\Theta_q^T M^{-1} \Theta_q$ is invertible and we can solve for the ν's from the joint equations (the bottom rows) in (4.13). The result is a standard LCP of the form

$$s = (R^T M^{-1} R - R^T M^{-1} \Theta_q (\Theta_q^T M^{-1} \Theta_q) \Theta_q^T M^{-1} R)\lambda + R^T M^{-1} Q + \Gamma_2 -$$
$$R^T M^{-1} \Theta_q (\Theta_q^T M^{-1} \Theta_q)^{-1} (\Theta_q^T M^{-1} Q + \Gamma_1) \quad (4.24)$$
$$s^T \lambda = 0;\ s, \lambda \geq 0. \quad (4.25)$$

A *strictly copositive* matrix is a matrix M with the property $x^T M x > 0$ for any $x \geq 0, x \neq 0$ [5, Def. 3.8.1]. If the matrix of an LCP is strictly copositive then the LCP will have a solution for any free term and a special algorithm, Lemke's algorithm [5, Sec.4.4], is guaranteed to find it.

Theorem 4.2.4 *The matrix of the LCP (4.24) is strictly copositive if and only if the contact configuration is disassemblable.*

Proof The matrix in (4.24) is positive semidefinite by construction. Therefore the LCP matrix M_{LCP} is not strictly copositive if and only if there is $\lambda \geq 0, \lambda \neq 0$ such that $\lambda^T M_{LCP} \lambda = 0$. If we take $\nu = -(\Theta_q^T M^{-1} \Theta_q)^{-1} (\Theta_q^T M^{-1} R)\lambda$ and use (4.24) we can write

$$\lambda^T M_{LCP} \lambda = (\lambda\ \nu) \begin{pmatrix} R^T \\ \Theta_q^T \end{pmatrix} M^{-1} (R\ \Theta_q) \begin{pmatrix} \lambda \\ \nu \end{pmatrix}. \quad (4.26)$$

If $\lambda^T M_{LCP} \lambda = 0$ it follows that $R\lambda + \Theta_q \nu = 0$ since M^{-1} is positive definite. But by the argument in the previous section this can happen if and only if the configuration is not disassemblable, which proves the claim. □

Corollary 4.2.5 *If a multibody contact configuration with frictionless joints is disassemblable or if the nondisassemblable part is frictionless then, for a sufficiently small friction coefficient μ, the corresponding LCP has a solution that can be found by Lemke's method.*

Proof By Theorem 4.2.3 we know that the frictionless problem can be recast as an LCP problem with disassemblable contacts. From Theorem 4.2.4 we know that the LCP matrix M_{LCP} is strictly copositive. We can introduce friction in any configuration, by simply replacing R with $R + \mu F$ in (4.24) any time it appears multiplied on the right of M^{-1}. This means that $M_{LCP,\mu} = M_{LCP} + \mu M_F$. Since M_{LCP} is a strictly copositive matrix , so is $M_{LCP,\mu}$ for a sufficiently small μ which proves the claim. □

4.2.4 Aspects of nondisassemblability

One example of a typical nondisassemblable configuration is the physical realization of a translational joint such as the one shown in Figure 4.2. The slider is assumed to be in contact with the bar on four sides. Clearly, such a configuration is not disassemblable. Assume that we have friction at the contacts and that we are pulling the slider with a horizontal force Q (we neglect gravity and assume the bar is fixed). By Lemma 4.2.2 there is a set of contact forces that cancel each other.

Any positive multiple of those contact forces would also satisfy (3.6) if the motion were frictionless. It follows that the slider can acquire any acceleration a between 0 and Q/m if the slider were initially at rest, where m is the mass of the slider (for each acceleration in that range we choose the reaction force to give the total friction $Q - ma$). Therefore, nondisassemblable configurations have a very high indeterminancy. In practice, such translational joints are lubricated and a simpler model (and more accurate for small reaction forces) would be to have the friction proportional to the velocity, and thus on the free term of the LCP. Technically, the LCP to be solved is the same as the frictionless one. And since, by the previous developments, the nondisassemblably part of the configuration can be treated as a joint, one does not have to worry about determining positive forces.

A more realistic model would ensure some space between the slider and the bar. This case could still be modeled as an LCP, but simulating such a configuration would get very difficult because of the many collisions that would need to be accounted for.

Thus, a more realistic (and robust) way of modeling friction in nondisassemblable configurations is to consider viscous friction, in which case the LCP for the nondisassemblable part is the same as for frictionless joints.

In general, determining if a configuration is disassemblable means to determine whether the corresponding polytope has an interior point. But the disassemblability of most configurations can be determined by inspection, for nondisassemblability by

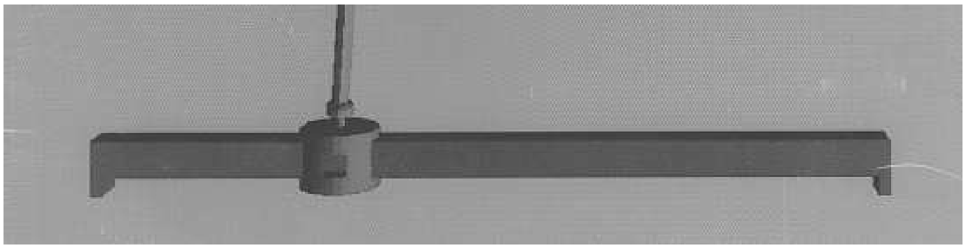

Figure 4.2: A translational joint

accident is very unusual. Therefore, most contact configurations have a solution for small values of the friction coefficient. Determining the value of μ for which we still have a strictly copositive problem is more difficult and there is no guarantee that we can determine a solution of the LCP if μ has a relatively high value. Of course, some bounds on μ can be found in terms of the operator norms of certain matrices, but at this stage, they do not seem to be very relevant. In [11] there is a bound for μ on a problem similar to (3.6). But the bound is found under the assumption that the matrix in that particular setting (a different matrix than ours) is diagonally dominant with positive entries on the diagonals. Translating the somewhat restrictive assumption to fit our setting of the problem does not seem likely to prove especially valuable. Our goal in the present paper was to develop general conditions on the existence of solutions in the presence of friction; we leave the problem of determining a good bound on the friction coefficient for future research.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

5.1 Summary of New Results

A discrete model for solving multi-rigid-body contact problems with friction based on Linear Complementarity Problems (LCP) has been presented in Chapter 2. The friction model is an approximation to Coulomb friction model and can be made as accurate as desired, by increasing the number of facets of the approximating friction pyramid. The main accomplishment is that the formulation *guarantees* that the contact problem always has a solution which will be found by Lemke's algorithm, unlike any of the previous formulations [3, 15, 16, 22]. It also has the property that the kinetic energy increase after one integration step in time cannot exceed the increase for the case where no constraint is active. Also we have formulated a model for treating impact with friction for restitution coefficients between 0 and 1. The model has a compression and a decompression phase, each needing Lemke's algorithm for finding a solution (which is guaranteed to exist).

We have discussed an explicit Euler integration strategy. If the collision resolution is guaranteed to yield a smaller kinetic energy after the collision, then, under some reasonable conditions on the external forces (at most linear increase with respect to velocity and position) the velocities given by the numerical algorithm are uniformly bounded as $h \downarrow 0$ for the Newton-Euler formulation in body coordinates. Also the velocities increase at most linearly with the time-interval considered if the external force is uniformly bounded, which is to be expected for real, continuous-time systems. An example of a collision resolution that guarantees a non-increasing kinetic energy is given in Section 2.2.

In Chapter 3 ways for setting up the contact constraint data used in (3.6), based on the contact geometry, were analyzed. In order to be able to set up the LCP (3.6) we need some kinematic data, including the contact points, of the bodies in contact. Finding that information requires solving an optimization problem (minimize the distance between two points on the bodies) or some other (even discrete) search method. We have tried to use the regularity of the contacts to find the derivatives of the contact points as a solution of a linear system (with the right hand side depending on the velocities of the bodies in contact) and use these to integrate for the contact points (but one needs to find a pair of contact points at least at the beginning, to start the simulation). If the matrix of the linear system is nonsingular, then for each relative velocity of the bodies we have a unique solution and, thus, for all nearby relative configurations there is going to be a unique pair of contact points and their position will depend smoothly on the relative position of the bodies.

At a first glance, our smooth-shaped bodies approach looks very much like an optimization problem, since we get a search direction via the right hand side of the contact evolution equations. The main difference is that in our approach the step size is specified in advance, being the integration step, plus the use of a high order integrator will result in a higher order approximation of the new contact point. If we use an optimization approach directly, we would throw away the knowledge from the previous integration steps. Furthermore, the contact evolution equations are interesting on their own.

We have also extended the analysis of the contact configuration to the case when there is no mutual contact but we intend to track the closest points on the bodies ($d > 0$) or allow a small amount of interpenetration ($d < 0$). We expect this to help enable better simulation of contacting and colliding complex-shaped bodies where frequent collision detection can be very costly. The singularity analysis we performed in Section 3.3 can help us identify the range of d for which our contact equations are valid. All the techniques presented here answer only the local problem of contact point evolution. In the case of non-convex bodies in contact, it will be necessary to use these methods in conjunction with classical collision detection algorithms.

The approach can also easily be extended to the (dynamic) friction case since the kinematic description of the contact situations is identical. The only thing to be changed is the expression of the contact force which is no longer Φ_q^T in (3.6) but has a component tangential to the contact. The developments from Chapter 2 can be used both with acceleration–force based models (like in Chapters 3 and 4) or velocity–impulsed based models (like in Chapter 2).

Also a way to approach planar contacts has been developed in Chapter 3. It essentially establishes criteria under which such contacts can be treated as bilateral planar constraints. This approach has a number of shortcomings, mainly that it is difficult to extend to the friction case because of the latter's nonlinearity. Also, the analysis does not apply in the case of a singular constraint Jacobian. The primary advantage is that it allows a reduced order system and it is encouraging for many heuristic approaches (especially if we have a technique to ensure that the contact forces are continuous). Another advantage generated by Lemma 3.2.1 is that the results apply regardless of the representation of the planar contact (three point-on-plane constraints, two curve-on-plane and one point-on-plane, etc.).

In Chapter 4, the connection between the structure of contact and the properties of the corresponding (acceleration-force) LCP was analyzed. We introduced the notion of disassemblable configuration and showed that it is equivalent to an LCP with a strictly copositive contact matrix. The conclusion is that a disassemblable contact configuration is robust with respect to small variations in friction. The problem of deciding if a configuration is disassemblable is the problem of finding an interior point, and this is solvable in polynomial time. The algorithm for finding the interior point can also be modified to identify any nondisassemblable parts. In practice, one can usually tell by inspection whether or not a mechanism is disassemblable.

Moreover, we showed that nondisassemblable frictionless configurations can be reduced to disassemblable configurations. In addition, if the nondisassemblable parts are modeled with viscous (instead of dry friction) the reduction is valid for the friction case as well.

5.2 Suggestions for future research

There are several challenges that remain to be solved in connection with a velocity-impulse based model. Higher-order methods need to be developed in order to take advantage of the high regularity of the joint constraints (and local regularity of the contact constraints). It is not obvious how to extend this Euler-type method for Runge-Kutta methods for example. However, one can use extrapolation methods on the regions where the basis of the LCP does not change, for the numerical solution is locally analytic there.

Several other issues arise in the practical implementation of the numerical method. Lemke's algorithm will not only compute a solution to the LCP, but it will also compute a basis of (A.1) [5]. If a basis is known, then a solution can be computed by solving a linear system. In many simulations changes of basis are relatively rare compared to the total number of integration steps. Therefore, the same basis can be used for several steps with the cost of just a linear solve and Lemke's algorithm will be called just when the necessity of a change of basis is detected (a negative slack). This observation has been used before [10, 3] with significant improvements of the computational effort. The advantage of our setting is that it guarantees that a change of basis which yields a solution will be computed by Lemke's algorithm.

The outcome of Chapter 2 is a discrete-time numerical scheme that has most of the properties required by a discretization of a multi-body-dynamics simulation with contact and friction scheme. But one natural question that appears is what happens in the limit as $h \downarrow 0$. The fact that the scheme has uniform boundedness of the velocities leads to the equicontinuity of the position sequences and the Arzela-Ascoli theorem can be applied to obtain a limit trajectory, $q(t)$ that satisfies the constraints exactly . However, Newton's equations cannot be satisfied in the classical sense in the limit and a different setting is necessary. Preliminary results suggest that the uniform boundedness of the velocities can lead to (subsequence) convergence to a solution in a differential inclusion sense [15, 22] which should define a generalized solution to the rigid-multi-body dynamics with contact and friction.

APPENDIX

A.1 Proof of Theorem 2.1.1

We can assume without loss of generality that F^T is full row rank. If not, we can consider just a maximal set of independent rows as being F^T. The components of y corresponding to the dependent variables can be set to 0, and λ will be in the same linear space determined by (2.15).

After solving for x in (2.15), the LCP becomes

$$0 = F^T M^{-1} F y + F^T M^{-1} H \lambda + F^T M^{-1} k \tag{A.1}$$
$$s = H^T M^{-1} F y + H^T M^{-1} H \lambda + N\lambda + H^T M^{-1} k + b \tag{A.2}$$
$$s \geq 0; \quad \lambda \geq 0; \quad \lambda^T s = 0, \tag{A.3}$$

Since F is full row rank $F^T M^{-1} F$ is nonsingular. Therefore we can solve for y from the first equation to get the LCP:

$$s = (G + N)\lambda + g; \tag{A.4}$$
$$s \geq 0; \quad \lambda \geq 0; \quad \lambda^T s = 0, \tag{A.5}$$

where

$$G = H^T M^{-1} H - H^T M^{-1} F (F^T M^{-1} F)^{-1} F^T M^{-1} H,$$
$$g = -H^T M^{-1} F (F^T M^{-1} F)^{-1} F^T M^{-1} k + H^T M^{-1} k + b.$$

By construction G is the Schur complement of N in the big matrix of (2.15), if $N = 0$ [5, Def. 2.3.4]. Since the matrix in (2.15) is positive semidefinite for $N = 0$, so is G [5, Thm 4.1.5]. Therefore $G + N$ is copositive.

It is convenient to denote the above LCP as LCP(G+N,g) and to call λ its solution (once λ is found then $s = (G + N)\lambda + g$). Let now z be a solution of LCP(G+N,0). Then $z^T G z = 0$. Since G is positive semidefinite and symmetric, it follows that $Gz = 0$. Let $w = -(F^T M^{-1} F)^{-1} F^T M^{-1} H z$. Then we have from the definition of G and the last relation

$$0 = F^T M^{-1} F w + F^T M^{-1} H z \tag{A.6}$$
$$0 = H^T M^{-1} F w + H^T M^{-1} H z. \tag{A.7}$$

By multiplying the first relation by w^T and the last relation by z^T and adding them up, we obtain $(Hz + Fw)^T M^{-1} (Hz + Fw) = 0$. Since M is symmetric and positive definite, it follows that $Hz + Fw = 0$ and $z^T g = w^T F^T M^{-1} k + z^T H^T M^{-1} k + z^T b = (Fw + Hz)^T M^{-1} k + z^T b \geq 0$.

Therefore, we have proved that if z is a solution of LCP(G+N,0) then $g^T z \geq 0$. Therefore, by Corollary 4.4.12 of [5] Lemke's algorithm, with precautions taken against degeneracy will find a solution λ to the LCP(G+N,g) and, by solving for x and y in the first two rows of (2.15), a solution (x, y, λ) to the initial LCP. $\square$.

A.2 Proof of Theorem 2.4.1

Lemma A.2.1 *Let $\{z_n\}_{n=0}^{N}$, $\{w_n\}_{n=0}^{N}$, be two finite nonnegative sequences and $\{h_n\}_{n=0}^{N-1}$ a positive sequence such that $\sum_{j=0}^{N-1} h_j = T$. Let c_i, $i = 1..5$ be positive constants. Assume that the sequences satisfy the following inequalities:*

$$\begin{aligned} z_{n+1}^2 &\leq z_n^2 + c_1 h_n z_n^2 + c_2 h_n w_n + c_3 h_n z_n + c_4 h_n && \text{(A.8)} \\ w_{n+1} &\leq w_n + c_5 h_n z_n && \text{(A.9)} \end{aligned}$$

for $0 \leq n < N-1$. Let $c = \max 2c_i$, $i = 1..5$. Then $z_i^2 \leq (z_0^2 + w_0)e^{2cT} + 2cTe^{2cT}$, $w_i \leq (z_0^2 + w_0)e^{2cT} + 2cTe^{2cT}$, $\forall i \leq N$.

Proof Using the inequality $2z_n \leq z_n^2 + 1$ we can write

$$\begin{aligned} z_{n+1}^2 &\leq z_n^2 + c_1 h_n z_n^2 + c_2 h_n w_n + c_3 h_n z_n + c_4 h_n && \text{(A.10)} \\ &\leq z_n^2 + c_1 h_n z_n^2 + c_2 h_n w_n + c_3 h_n (z_n^2/2 + 1/2) + c_4 h_n && \text{(A.11)} \\ &\leq z_n^2 + c h_n z_n^2 + c h_n w_n + c h_n. && \text{(A.12)} \end{aligned}$$

We also have that

$$\begin{aligned} w_{n+1} &\leq w_n + c_5 h_n z_n \leq w_n + c_5 h_n (z_n^2/2 + 1/2) && \text{(A.13)} \\ &\leq w_n + c h_n z_n^2 + c h_n w_n + c h_n. && \text{(A.14)} \end{aligned}$$

Let $r_n = z_n^2 + w_n$. By adding (A.12) with (A.14) we obtain

$$r_{n+1} \leq r_n + 2ch_n r_n + 2ch_n \leq r_n e^{2ch_n} + 2ch_n. \tag{A.15}$$

Let $u_n = \sum_{i=0}^{n} h_i$. We will prove that $r_n \leq (r_0 + 2cu_{n-1})e^{2cu_{n-1}}$ by induction. It is obvious for $n = 1$. Assume it is true for n. Then

$$\begin{aligned} r_{n+1} &\leq r_n e^{2ch_n} + 2ch_n \leq (r_0 + 2cu_{n-1})e^{2cu_{n-1}} e^{2ch_n} + 2ch_n && \text{(A.16)} \\ &\leq (r_0 + 2cu_{n-1})e^{2cu_n} + 2ch_n e^{2cu_n} = (r_0 + 2cu_n)e^{2cu_n} && \text{(A.17)} \end{aligned}$$

Using the fact that $u_i \leq T$ for $0 \leq i \leq N-1$, we can write $z_i^2 \leq r_i \leq (z_0^2 + w_0 + 2cT)e^{2cT}$, $w_i \leq r_i \leq (z_0^2 + w_0 + 2cT)e^{2cT}$, for $0 \leq i \leq N$, which proves the claim. $\square$.

Proof of Theorem 2.4.1 a) By Theorem 2.1.3 we have

$$v^{(l+1)^T} M v^{(l+1)} \leq (v^{(l)} + h_l M^{-1} k)^T M (v^{(l)} + h_l M^{-1} k).$$

Let $\gamma^{(l)} = M^{1/2}v^{(l)}$ and $\chi^{(l)} = M^{-1/2}k(q^{(l)}, v^{(l)})$. Then the last relation can be rewritten as

$$\|\gamma^{(l+1)}\|^2 \leq \|\gamma^{(l)} + h_l\chi^{(l)}\|^2 = \|\gamma^{(l)}\|^2 + 2{\gamma^{(l)}}^T\chi + h_l^2\|\chi^{(l)}\|^2. \tag{A.18}$$

By the definition of the Coriolis force we have $v^T f_c = 0$. Therefore

$$\begin{aligned} {\gamma^{(l)}}^T\chi &= {v^{(l)}}^T(f_c(v^{(l)}) + k_1(v^{(l)}, q^{(l)}) + k_2(v^{(l)}) + k_3(q^{(l)})) && \text{(A.19)}\\ &= {v^{(l)}}^T(k_1(v^{(l)}, q^{(l)}) + k_2(v^{(l)}) + k_3(q^{(l)})) && \text{(A.20)}\\ &= {\gamma^{(l)}}^T(\chi_1(v^{(l)}, q^{(l)}) + \chi_2(v^{(l)}) + \chi_3(q^{(l)})), && \text{(A.21)} \end{aligned}$$

where $\chi_i = M^{-1/2}k_i, i = 1, 2, 3$. By using the inequalities in the hypothesis, we get

$$\begin{gathered} {\gamma^{(l)}}^T(\chi_1(v^{(l)}, q^{(l)}) + \chi_2(v^{(l)}) + \chi_3(q^{(l)})) \leq \\ \|\gamma^{(l)}\|(c_1 + c_2\|\gamma^{(l)}\| + c_3\|q^{(l)}\|), \qquad \text{(A.22)} \\ \|\chi^{(l)}\| \leq (c_1 + c_2\|\gamma^{(l)}\| + c_3\|q^{(l)}\| + c_4\|\gamma^{(l)}\|^2), \qquad \text{(A.23)} \end{gathered}$$

where the constants $c_i, i = 1..4$ depend only on the constants $d_i, i = 1..4$ and the mass matrix. The term $c_4\|\gamma^{(l)}\|^2$ is a bound on the Coriolis force. Therefore, according to (A.18) we can write

$$\begin{aligned} \|\gamma^{(l+1)}\|^2 &\leq \|\gamma^{(l)}\|^2 + 2\|\gamma^{(l)}\|(c_1 + c_2\|\gamma^{(l)}\| + c_3\|q^{(l)}\|) \\ &\qquad + h_l^2\|\chi^{(l)}\|^2. \end{aligned} \tag{A.24}$$

Denoting $z_l = \|\gamma^{(l)}\|$ and $w_l = \|q^{(l)}\|$ we get

$$z_{l+1}^2 \leq z_l^2 + c_1h_lz_l + c_2h_lz_l^2 + c_3h_lw_lz_l + h_l^2\|\chi^{(l)}\|^2. \tag{A.25}$$

Also, $q^{(l+1)} = q^{(l)} + h_lv^{(l)} = q^{(l)} + h_lM^{-1/2}\gamma^{(l)}$ and therefore

$$w_{l+1} \leq w_l + h_lc_5z_l, \tag{A.26}$$

where c_5 depends only on the mass matrix and $w_l = \|q^{(l)}\|$. Let $c = \max 2c_i,\ i = 1..5$. We choose h such that

$$\begin{gathered} (c_1 + c_2(\|\gamma_0\|^2 + \|q_0\| + 2cT)^{\frac{1}{2}}e^{cT} \\ +(c_3 + c_4)(\|\gamma_0\|^2 + \|q_0\| + 2cT)e^{2cT})^2 \leq \tfrac{c}{2h}. \end{gathered} \tag{A.27}$$

We will prove by induction that

$$z_{l+1}^2 \leq z_l^2 + c_1h_lz_l + c_2h_lz_l^2 + c_3h_lw_lz_l + \frac{c}{2}h_l. \tag{A.28}$$

for $0 \leq l \leq N$. From (A.25) it follows that it is sufficient to prove by induction that

$$h_l \|\chi^{(l)}\|^2 \leq \frac{c}{2}. \tag{A.29}$$

For $l = 0$ this follows from (A.23) and (A.27) with

$$\|\chi^0\| \leq (c_1 + c_2\|\gamma^{(0)}\| + c_3\|q^{(0)}\| + c_4\|\gamma^{(0)}\|^2) \leq \frac{c}{2h_l}$$

Assume that (A.29), and hence (A.28), holds for some $0 \leq n < N-1$. From (A.26) it follows that the sequences w_l, z_l, $0 \leq l \leq n-1$ satisfy the conditions from Lemma A.2.1. Since $\sum_{l=0}^{n-1} h_l \leq T$, we have

$$\begin{aligned} z_l^2 &\leq (z_0^2 + w_0)e^{2cT} + 2cTe^{2cT}, \qquad 0 \leq l \leq n-1, \\ w_l &\leq (z_0^2 + w_0)e^{2cT} + 2cTe^{2cT}, \qquad 0 \leq l \leq n-1. \end{aligned}$$

Using (A.27) and (A.23) we deduce that $h_n\|\chi^{(n)}\|^2 \leq \frac{c}{2}$ which completes the induction.

Therefore w_l, z_l satisfy (A.28) and (A.26) and the assumptions of Lemma A.2.1, for $0 \leq l \leq N$, so that the conclusion follows.

b) A similar sequence of inequalities shows that

$$\|\gamma^{(l+1)}\|^2 \leq \|\gamma^{(l)}\|^2 + 2h_l c_1\|\gamma^{(l)}\| + h_l\|\chi^{(l)}\|^2 \tag{A.30}$$

Here c_1 is a constant dependent on d_1 and the mass matrix, M. Following the hypothesis on the external force, we have that

$$\|\chi^{(l)}\| \leq c_1 + c_4\|\gamma^{(l)}\|^2 \tag{A.31}$$

where c_4 depends only on the mass matrix. Let $c = 2\max\{c_1, c_4\}$. We choose h such that

$$h(c_1 + c_4(\sqrt{\|\gamma_0\|^2 + 1} + cT)^2)^2 \leq \frac{c}{2}. \tag{A.32}$$

Let $z_l = \|\gamma^{(l)}\|$. We will prove by induction that

$$\sqrt{z_n^2 + 1} \leq \sqrt{z_0^2 + 1} + c\sum_{l=0}^{n-1} h_l \tag{A.33}$$

for $0 \leq n \leq N$. For $n = 0$ the claim is obvious. Assume it is true for $n - 1 \leq N - 1$. Then by virtue of (A.30) we can write

$$z_n^2 \leq z_{n-1}^2 + 2h_{n-1}c_1 z_{n-1} + h_l^2\|\chi^{(l)}\|^2. \tag{A.34}$$

From (A.31) and the induction hypothesis it follows that:

$$h_{n-1}\|\chi^{(n-1)}\|^2 \le h(c_1 + c_4 z_{n-1}^2)^2$$
$$\le h(c_1 + c_4(\sqrt{\|\gamma_0\|^2 + 1} + cT)^2)^2 \le \frac{c}{2}.$$

Therefore, from (A.34) we deduce that

$$z_n^2 \le z_{n-1}^2 + 2h_{n-1}c_1 z_{n-1} + h_{n-1}\frac{c}{2}.$$

Adding 1 to both sides, applying Cauchy's inequality and the induction hypothesis, we get

$$z_n^2 + 1 \le z_{n-1}^2 + 1 + 2h_{n-1}c\sqrt{z_{n-1}^2 + 1} \le (\sqrt{z_{n-1}^2 + 1} + h_{n-1}c)^2$$
$$\le (\sqrt{z_0^2 + 1} + c\sum_{l=1}^{n-2} h_l + ch_{n-1})^2$$

which proves that (A.33) is true for n. The Theorem is proved by taking $n = N$ in (A.33) . $\square$

A.3 The algebra of relative velocities

We briefly describe the algebra of relative velocities from [17] that we will use in order to describe the equations of the contact.

The relative position of two frames, say a, b is described by the position vector p_{ab} and by the relative rotation matrix R_{ab}. The absolute rotation matrices satisfy $R_b = R_a R_{ab}$. All the transformations can be defined in terms of a matrix $g_{ab} \in I\!R^{4\times 4}$ defined as

$$g_{ab} = \begin{bmatrix} R_{ab} & p_{ab} \\ 0 & 1 \end{bmatrix} \tag{A.35}$$

With this definition, the spatial and body velocities are defined to be $\hat{V}_{ab}^s, \hat{V}_{ab}^b \in I\!R^{4\times 4}$, $\hat{V}_{ab}^s = \dot{g_{ab}} g_{ab}^{-1}$ and $\hat{V}_{ab}^b = g_{ab}^{-1}\dot{g_{ab}}$. We will be mostly interested in the latter case, for which we have the following explicit representation

$$\hat{V}_{ab}^b = \begin{bmatrix} R_{ab}^T \dot{R}_{ab} & R_{ab}^T \dot{p}_{ab} \\ 0 & 0 \end{bmatrix} \tag{A.36}$$

On this basis we recover the familiar form of the relative velocity

$$V_{ab}^b = \begin{bmatrix} v_{ab}^b \\ \omega_{ab}^b \end{bmatrix} = \begin{bmatrix} R_{ab}^T \dot{p}_{ab} \\ (R_{ab}^T \dot{R}_{ab})^\vee \end{bmatrix} \tag{A.37}$$

One of the most important relations used from here on will be the one relating velocities of several frames. For this purpose, we associate with each g_{ab} type matrix the following matrix:

$$Ad_{g_{ab}} = \begin{bmatrix} R_{ab} & \hat{p}_{ab} R_{ab} \\ 0 & R_{ab} \end{bmatrix} \tag{A.38}$$

Finally, with this notation we have the transformation relation

$$V^b_{ac} = Ad_{g^{-1}_{bc}} V^b_{ab} + V^b_{bc}. \tag{A.39}$$

which we use extensively in the paper.

At times, we will need the following velocity: $V^b_{l_o l_f}$, where l_o and l_f are frames which are fixed with respect to o and f respectively. Thus $V^b_{ol_o} = V^b_{fl_f} = 0$. Hence, from A.39

$$V^b_{l_o l_f} = Ad_{g^{-1}_{fl_f}} V^b_{l_o f} = Ad_{g^{-1}_{fl_f}} V^b_{of} \tag{A.40}$$

A.4 The proof of Theorem 3.3.1

All the calculations are done adapting the treatment from [13] for this case. Since all calculations involve body velocities, we drop the superscript b. We also need the following

Lemma A.4.1 . Induced velocity of the contact frame
The (body) velocity of the contact frame C relative to the reference frame O of the body is given by $V^b_{oc} = (v_{oc}, w_{oc})$ where

$$v_{oc} = \begin{bmatrix} M\dot{\alpha} \\ 0 \end{bmatrix} \tag{A.41}$$

$$\hat{\omega}_{oc} = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix} = \begin{bmatrix} \left| \begin{matrix} 0 & -TM\dot{\alpha} \\ TM\dot{\alpha} & 0 \end{matrix} \right| & KM\dot{\alpha} \\ -(KM\dot{\alpha})^T & 0 \end{bmatrix} \tag{A.42}$$

Proof See [13]. □

Proof of Theorem 3.3.1. Since the frame $L_f(t)$ is fixed relative to the frame F, the velocity of $L_f(t)$ relative to F is given by $V_{fl_f} = V_{ol_o} = 0$. Therefore using the velocity relationships (A.39),we get:

$$V_{fc_f} = V_{l_f c_f} \tag{A.43}$$

$$V_{oc_o} = V_{l_o c_o}. \tag{A.44}$$

Also, at moment t, $L_o(t) = C_o$ and $L_f(t) = C_f$. Therefore $p_{l_oc_o} = p_{l_fc_f} = 0$ and $R_{l_oc_o} = Rl_fc_f = I$ at moment t. Also

$$p_{c_oc_f} = \begin{pmatrix} 0 \\ 0 \\ d \end{pmatrix}; R_{c_oc_f} = \begin{bmatrix} R_\psi & 0 \\ 0 & -1 \end{bmatrix} \tag{A.45}$$

The latter relation holds because the z directions of the two contact frames coincide, and ψ is the amount of rotation measured around this axis.

Now we will compute the velocity of C_f relative to L_o via two intermediate frames, namely L_f and C_o. Following (A.39) we have

$$V_{l_ol_f} + V_{l_fc_f} = V_{l_oc_f} = Ad_{g^{-1}_{c_oc_f}} V_{l_oc_o} + V_{c_oc_f} =$$
$$\begin{bmatrix} R^T_{c_oc_f} & -(R^T_{c_oc_f} p_{c_oc_f})\hat{} R^T_{c_oc_f} \\ 0 & R_{c_oc_f} \end{bmatrix} V_{oc_o} + Vc_oc_f. \tag{A.46}$$

By definition, $V_{l_ol_f} = (v_x, v_y, v_z, \omega_x, \omega_y, \omega_z)$ and $V_{oc_o}, V_{l_fc_f} = V_{fc_f}$ are given by Lemma A.4.1. Thus we can use the last relations to get the derivatives of the coordinates of the contact point.

$$\begin{pmatrix} M_f\dot{\alpha}_f \\ 0 \end{pmatrix} + \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = \tag{A.47}$$

$$\begin{pmatrix} R_\psi M_o\dot{\alpha}_o \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \dot{d} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ d \end{pmatrix} \times \begin{pmatrix} R_\psi & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} -(K_oM_o\dot{\alpha}_o)_y \\ (K_oM_o\dot{\alpha}_o)_x \\ T_fM_f\dot{\alpha}_f \end{pmatrix} = \tag{A.48}$$

$$\begin{pmatrix} R_\psi M_o\dot{\alpha}_o \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \dot{d} \end{pmatrix} + \begin{pmatrix} dR_\psi K_oM_o\dot{\alpha}_o \\ 0 \end{pmatrix} \tag{A.49}$$

$$\begin{pmatrix} K_fM_f\dot{\alpha}_f \\ T_fM_f\dot{\alpha}_f \end{pmatrix} + \begin{pmatrix} \omega_y \\ -\omega_x \\ \omega_z \end{pmatrix} = \tag{A.50}$$

$$\begin{pmatrix} 0 \\ 0 \\ \dot{\psi} \end{pmatrix} - \begin{pmatrix} R_\psi K_oM_o\dot{\alpha}_o \\ T_oM_o\dot{\alpha}_o \end{pmatrix}. \tag{A.51}$$

In describing the equations relating the angular velocities, we have multiplied the first line with -1 and switched it with the second line. After some algebra we get the required form for the contact coordinate derivatives. $\square$.

A.5 Proof of the Theorem 3.3.3

Lemma A.5.1 . Induced velocity of the contact frame

The (body) velocity of the contact frame C relative to the reference frame O of the *body is given by* $V^b_{oc} = (v_{oc}, w_{oc})$ *where*

$$v_{oc} = \begin{bmatrix} \dot{l}_o \\ 0 \\ 0 \end{bmatrix} \tag{A.52}$$

$$\omega_{oc} = \begin{bmatrix} \tau_o \dot{l}_o + \dot{\phi}_o \\ k_o \sin\phi_o \dot{l}_o \\ k_o \cos\phi_o \dot{l}_o \end{bmatrix}, \tag{A.53}$$

and where $l(t)$ *is the variation of the intrinsic parameter of the curve in time and* k_o, τ_o *are the curvature and respectively the torsion of the curve on body* O.

Proof Let $\vec{t}, \vec{n}, \vec{b}$ be the tangent, the normal and the binormal of the curve. Let $R_T = [\vec{t}, \vec{n}, \vec{b}]$ and

$$R_{\phi_o} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi_o & -\sin\phi_o \\ 0 & \sin\phi_o & \cos\phi_o \end{bmatrix} \tag{A.54}$$

Then clearly $R_{oc} = R_T R_{\phi_o}$. Let also

$$\mathcal{K} = \begin{bmatrix} 0 & k_o & 0 \\ -k_o & 0 & \tau_o \\ 0 & -\tau_o & 0 \end{bmatrix} \tag{A.55}$$

where k_o, τ_o are the curvature and respectively the torsion of the curve. Frenet's relations state that $\dot{R_T} = -R_T \mathcal{K} \dot{l}_o$. Therefore

$$R^T_{oc}\dot{R_{oc}} = R^T_{\phi_o} R^T_T \dot{R_T} R_{\phi_o} + R^T_{\phi_o} R^T_T R_T \dot{R_{\phi_o}} = \tag{A.56}$$

$$-\dot{l}_o R^T_{\phi_o} R^T_T R_T \mathcal{K} R_{\phi_o} + R^T_{\phi_o} R^T_T R_T \dot{R_{\phi_o}} = \tag{A.57}$$

$$-\dot{l}_o R^T_{\phi_o} \mathcal{K} R_{\phi_o} + R^T_{\phi_o} \dot{R_{\phi_o}} \tag{A.58}$$

It is easy to see that

$$R^T_{\phi_o} \dot{R_{\phi_o}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -\dot{\phi}_o \\ 0 & \dot{\phi}_o & 0 \end{bmatrix}. \tag{A.59}$$

Since $\hat{\omega_{oc}} = R^T_{oc} \dot{R}_{oc}$ we get the claimed relation for ω_{oc} after replacing all the terms in A.58.

Things are easier for v_{oc}. We have that $p_{oc}(t)$ is the parameterization of the curve with respect to O. Therefore $\dot{p}_{oc}(t) = \dot{l}_o \vec{t}$ and

$$v_{oc} = R^T_{oc} \dot{p}_{oc} = \dot{l}_o R^T_{\phi_o} R^T_T \vec{t} \tag{A.60}$$

$$\dot{l}_o R^T_{\phi_o} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \dot{l}_o \\ 0 \\ 0 \end{bmatrix} \tag{A.61}$$

which completes the proof. $\square$

Proof of Theorem 3.3.3 All the calculations are done again adapting the treatment from [13] to this case. Because we are dealing only with body velocities, we drop the superscript b once more. Since the frame $L_f(t)$ is fixed relative to the frame F, the velocity of $L_f(t)$ relative to F is given by $V_{fl_f} = V_{ol_o} = 0$. Therefore using the velocity relationships (A.39), we get

$$V_{fc_f} = V_{l_fc_f} \tag{A.62}$$

$$V_{oc_o} = V_{l_oc_o}. \tag{A.63}$$

The difference from the surface-surface contact is that, at moment t, $C_o = L_o R_{\phi_o}$ and $C_f = L_f R_{\phi_f}$, with the matrices R_ϕ being defined by A.54. Therefore $p_{l_oc_o} = p_{l_fc_f} = 0$ and $R_{l_oc_o} = R_{\phi_o}$ and $Rl_fc_f = R_{\phi_f}$ at moment t. Also

$$p_{c_oc_f} = \begin{pmatrix} 0 \\ 0 \\ d \end{pmatrix}; R_{c_oc_f} = \begin{bmatrix} R_\psi & 0 \\ 0 & -1 \end{bmatrix} \tag{A.64}$$

The latter relation holds because the z directions of the two contact frames coincide, and ψ is the amount of rotation measured around this axis.

Now we compute the velocity of C_f relative to L_o via two intermediate frames, namely L_f and C_o. Following (A.39) we have

$$Ad_{g^{-1}_{l_fc_f}} V_{l_ol_f} + V_{l_fc_f} = V_{l_oc_f} = Ad_{g^{-1}_{c_oc_f}} V_{l_oc_o} + V_{c_oc_f} \tag{A.65}$$

One difference from the surface-surface case is that since L_f and C_f do not coincide any longer but are mutually rotated by an angle ϕ_f, $Ad_{g^{-1}_{l_fc_f}}$ is not the identity. After replacing the expressions for the adjoints and using (A.62) we obtain

$$\begin{bmatrix} R^T_{l_fc_f} & -R^T_{l_fc_f}\hat{p}_{l_fc_f} \\ 0 & R^T_{l_fc_f} \end{bmatrix} V_{l_ol_f} + \begin{pmatrix} \tilde{v}_x \\ \tilde{v}_y \\ \tilde{v}_z \end{pmatrix} = \tag{A.66}$$

$$\begin{pmatrix} R_\psi \begin{pmatrix} \dot{l}_o \\ 0 \end{pmatrix} \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \dot{d} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ d \end{pmatrix} \times \begin{bmatrix} R_\psi & 0 \\ 0 & -1 \end{bmatrix} \begin{pmatrix} \tau_o \dot{l}_o + \dot{\phi}_o \\ k_o \sin\phi_o \dot{l}_o \\ k_o \cos\phi_o \dot{l}_o \end{pmatrix} \tag{A.67}$$

or

$$\begin{pmatrix} \dot{l}_f \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \tilde{v}_x \\ \tilde{v}_y \\ \tilde{v}_z \end{pmatrix} = \begin{pmatrix} R_\psi \begin{pmatrix} \dot{l}_o \\ 0 \end{pmatrix} \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \dot{d} \end{pmatrix} - \begin{pmatrix} R_\psi \begin{pmatrix} -dk_o \sin\phi_o \dot{l}_o \\ d\tau_o \dot{l}_o + d\dot{\phi}_0 \end{pmatrix} \\ 0 \end{pmatrix}. \tag{A.68}$$

The equations involving the rotational velocities (the last three rows) are done by multiplying the first row by -1 and permuting it with the second row.

$$\begin{pmatrix} k_f \sin\phi_f \dot{l}_f \\ -\tau_f \dot{l}_f - \dot{\phi}_f \\ k_f \cos\phi_f \dot{l}_f \end{pmatrix} + \begin{pmatrix} \tilde{\omega}_y \\ -\tilde{\omega}_x \\ \tilde{\omega}_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \dot{\psi} \end{pmatrix} - \begin{pmatrix} R_\psi \begin{pmatrix} k_o \sin\phi_o \dot{l}_o \\ -\tau_o \dot{l}_o - \dot{\phi}_o \end{pmatrix} \\ -k_o \cos\phi_o \dot{l}_o \end{pmatrix}. \tag{A.69}$$

After separating the first two rows from the last, both for translational and rotational velocities, we get the conclusion. □

A.6 Proof of Theorem 3.3.6

Proof We equal the left-hand sides of the computations from the curve-curve section with the right-hand sides of the computations from the surface-surface section and we get for the translational velocities:

$$\begin{pmatrix} \dot{l}_f \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \tilde{v}_x \\ \tilde{v}_y \\ \tilde{v}_z \end{pmatrix} = \begin{pmatrix} R_\psi M_o \dot{\alpha}_o \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \dot{d} \end{pmatrix} + \begin{pmatrix} d R_\psi K_o M_o \dot{\alpha}_o \\ 0 \end{pmatrix} \tag{A.70}$$

and for the rotational velocities

$$\begin{pmatrix} k_f \sin\phi_f \dot{l}_f \\ -\tau_f \dot{l}_f - \dot{\phi}_f \\ k_f \cos\phi_f \dot{l}_f \end{pmatrix} + \begin{pmatrix} \tilde{\omega}_y \\ -\tilde{\omega}_x \\ \tilde{\omega}_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \dot{\psi} \end{pmatrix} - \begin{pmatrix} R_\psi K_o M_o \dot{\alpha}_o \\ T_o M_o \dot{\alpha}_o \end{pmatrix}. \tag{A.71}$$

After separating the first two rows from the last for each equation, we get the conclusion. □

A.7 The explicit value of $V^b_{l_o l_f}$

By (A.40) we have $V^b_{l_o l_f} = Ad_{g^{-1}_{fl_f}} V^b_{of} = Ad^{-1}_{g_{fl_f}} V^b_{of}$. V^b_{of} is a velocity that can be computed from (A.37) and it depends on the state of the bodies in contact and not on their particular geometry. Thus the only thing left that is geometry-dependent is $Ad_{g_{fl_f}}$. Most of the notations used here come from the introductory section of the appendix.

A.7.1 The surface on surface contact

We have $p_{fl_f} = c_f(\alpha_f)$, the local coordinates of the contact point with respect to the body F. We also have that $R_{fl_f} = [\frac{r_{u_f}}{\|r_{u_f}\|} \frac{r_{v_f}}{\|r_{v_f}\|} N]$, where $\alpha_f = (u_f, v_f), c = c(\alpha_f)$ is the parameterization of the bodies and N is the external normal, chosen such that the determinant of R_{fl_f} is 1. R_{fl_f} is the natural frame on any smooth surface. Now $Ad_{g_{fl_f}}$ is found by (A.38) and subsequently $V^b_{l_o l_f}$.

A.7.2 The curve on curve contact

For this type of contact $p_{fl_f} = c_f(l_f)$, the local coordinates of the contact point with respect to the body F, where l_f is the natural curve parameter (the length of the curve with respect to some fixed initial point). $H_{fl_f} = [\vec{t}\vec{n}\vec{b}]$, the Frenet frame or Frenet trihedron [6, p.19] of the curve in local coordinates, thus with respect to body F. $Ad_{g_{fl_f}}$ is now found by (A.38) after replacing the corresponding expressions of p_{fl_f} and H_{fl_f}.

For applying Theorem 3.3.1 we also need $Ad_{g^{-1}_{l_f c_f}} = Ad^{-1}_{g_{l_f c_f}}$. Since the points l_f and c_f coincide, we have $p_{l_f c_f} = 0$. The frames L_f and C_f are rotated by the angle ϕ_f, thus $H_{l_f c_f} = R_{\phi_f}$ with the latter being given by (A.54) with ϕ_o replaced by ϕ_f. $Ad_{l_f c_f}$ can now be computed from $p_{l_f c_f}$ and $H_{l_f c_f}$ by using (A.38).

A.7.3 The curve on surface contact

Since the curve is on body F, the approach is identical to the one from the previous case.

REFERENCES

[1] Anitescu, M., Cremer, J. and Potra, F.A., 'Formulating 3D contact dynamics problems', *Mechanics of Structures and Machines* **24(4)**, 1996, 405-437.

[2] Anitescu, M., Lesaja, G. and Potra, F.A., 'Equivalence between different formulations of the linear complementarity problem', *Applied Mathematics and Optimization*, to appear.

[3] Baraff, D., ' Issues in computing contact forces for non-penetrating rigid bodies', *Algorithmica* **10**, 1993, 292-352.

[4] Cai, C. and Roth, B., ' On the spatial motion of rigid bodies with point contact', *Proc. 1987 IEEE Conf. on Robotics and Automation*, 1987, 686-695.

[5] Cottle, R.W., Pang, J.-S. and Stone, R.E., *The Linear Complementarity Problem*, Academic Press, Boston, 1992.

[6] Do Carmo, M.P., *Differential Geometry of Curves and Surfaces*, Prentice-Hall,Englewood Cliffs,NJ,1976.

[7] Featherstone, R., *Robot Dynamics Algorithms*, Kluwer Academic Publishers, Boston, 1987.

[8] Glocker, Ch. and Pfeiffer, F., 'Multiple impacts with friction in rigid multi-body systems', *Nonlinear Dynamics* **7**, 1995, 471-497.

[9] Haug, E.J., *Computer Aided Kinematics and Dynamics of Mechanical Systems*, Allyn and Bacon, Boston, 1989.

[10] Haug, E.J., Wu, S.C. and Yang, S.M., 'Dynamic mechanical systems with Coulomb friction, stiction, impact and constraint addition-deletion I: Theory', *Mechanisms and Machine Theory* **21(5)**, 1986, 407-416.

[11] Liu, Z., ' Convergence analysis of projected Jacobi method for contact force computation with friction', *Proceedings of the First ACM Workshop on Simulation and Interaction in Virtual Environments*, Iowa City, 1995, July, 1995.

[12] Lötstedt, P., 'Mechanical systems of rigid bodies subject to unilateral constraints', *SIAM Journal of Applied Mathematics* **42(2)**, 1982, 281-296.

[13] Montana, D.J., 'The kinematics of contact and grasp', *International Journal Of Robotics Research* **7(3)**, 1988, 17-32.

[14] Moreau, J.J., 'Standard inelastic shocks and the dynamics of unilateral constraints', in *Unilateral Problems in Structural Analysis*, G. Del Piero and F. Macieri (eds.), CISM Courses and Lectures no. 288, Springer - Verlag, Wien - New York, 1983.

[15] Moreau, J.J., 'Unilateral constraints and dry friction in finite freedom dynamics', in *Nonsmooth mechanics and applications*, J.J. Moreau and P.D. Panagiotopoulos (eds.), CISM Courses and Lectures no. 302, Springer - Verlag, Wien - New York, 1988.

[16] Moreau, J.J. and Jean, Michel, 'Numerical treatment of contact and friction: the contact dynamics method', in *Proceedings of the Third Biennial Joint Conference on Engineering Systems and Analysis*, Montpellier, France, July 1-4, 1996, to appear.

[17] Murray R.M., Li, Z. and Sastry, S. S., *A Mathematical Introduction to Robotic Manipulation*, CRC Press, Boca Raton, FL, 1993.

[18] Murray, R.M. and Sastry, S.S., 'Grasping and manipulation using multifingered robot hands', *Robotics: Proceedings of Symposia in Applied Mathematics, Volume 41*, R.W.Brockett (ed.), American Mathematical Society, 1990, 91–128.

[19] Painleve, P., 'Sur les lois du frotement de glissement', *Comptes Rendus de l'Academie des Sciences* **121**, Paris, 1895, 112-115.

[20] Pang, J.-S. and Trinkle, J.C., 'Complementarity Formulations and Existence of Solutions of Multi-Rigid-Body With Coulomb Friction', *Mathematical Programming*, to appear.

[21] Schimmels, J.M . and Peshkin, M.A., 'Admittance Matrix Design for Force–Guided Assembly', *IEEE Transactions on Robotics and Automation*, **8(2)**, 1992, 213-227,1992.

[22] Stewart, D.E. and Trinkle, J.C., 'An implicit time-stepping scheme for rigid-body dynamics with inelastic collisions and Coulomb friction', *International Journal for Numerical Methods in Engineering* **39**, 1996, 2673-2691.

[23] Trinkle, J.C.,' On the Stability and Instantaneous Velocity of Grasped Frictionless Objects', *IEEE Transactions on Robotics and Automation* **8(5)**, 1992, 560-572.

[24] Trinkle, J.C., Pang, J.S., Sudarsky, S. and Lo, G., 'On dynamic multi-rigid-body contact problems with Coulomb friction', *Zeithschrift fur Angewandte Mathematik und Mechanik*, to appear.

[25] Wright, S.J., 'A path–following interior–point algorithm for linear and quadratic problems', *Preprint P401–1293*, Argonne National Laboratory, Mathematics and Computer Science Division, 1993.