

An Implicit Time-Stepping Method for Multibody Systems with Intermittent Contact

Nilanjan Chakraborty, Stephen Berard, Srinivas Akella and Jeff Trinkle

Department of Computer Science

Rensselaer Polytechnic Institute

Troy, New York 12180

Email: {chakrn2, sberard, sakella, trink}@cs.rpi.edu

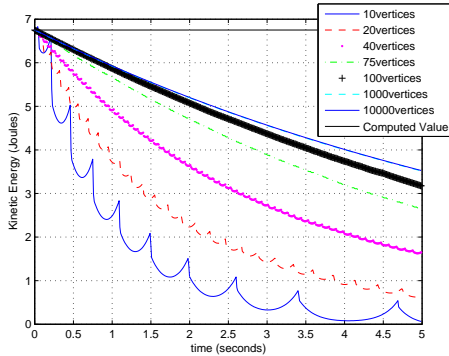
Abstract—In this paper we present an implicit time-stepping scheme for multibody systems with intermittent contact by incorporating the contact constraints as a set of complementarity and algebraic equations within the dynamics model. Two primary sources of stability and accuracy problems in prior time stepping schemes for differential complementarity models of multibody systems are the use of polyhedral representations of smooth bodies and the approximation of the distance function (arising from the decoupling of collision detection from the solution of the dynamic time-stepping subproblem). Even the simple example of a disc rolling on a table without slip encounters these problems. We assume each object to be a convex object described by an intersection of convex inequalities. We write the contact constraints as complementarity constraints between the contact force and a distance function dependent on the closest points on the objects. The closest points satisfy a set of algebraic constraints obtained from the KKT conditions of the minimum distance problem. These algebraic equations and the complementarity constraints taken together ensure satisfaction of the contact constraints. This enables us to formulate a *geometrically implicit* time-stepping scheme (*i.e.*, we do not need to approximate the distance function) as a nonlinear complementarity problem (NCP). The resulting time-stepper is therefore more accurate; further it is the first geometrically implicit time-stepper that does not rely on a closed form expression for the distance function. We demonstrate through example simulations the fidelity of this approach to analytical solutions and previously described simulation results.

I. INTRODUCTION

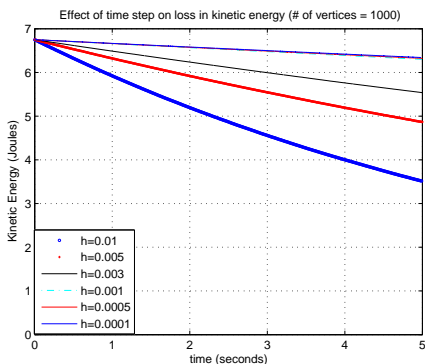
To automatically plan and execute tasks involving intermittent contact, one must be able to accurately predict the object motions in such systems. Applications include haptic interactions, collaborative human-robot manipulation, such as rearranging the furniture in a house, as well as industrial automation, such as simulation of parts feeders. Due to the intermittency of contact and the presence of stick-slip frictional behavior, dynamic models of such multibody systems are inherently (mathematically) nonsmooth, and are thus difficult to integrate accurately. In fact, commercially available software systems such as Adams, have a difficult time simulating any system with unilateral contacts. Users expect to spend considerable effort in a trial-and-error search for good simulation parameters to obtain believable, not necessarily accurate, results. Even the seemingly simple problem of a sphere rolling on a horizontal plane under only the influence of gravity is challenging for commercial simulators.

The primary sources of stability and accuracy problems are polyhedral approximations of smooth bodies, the decoupling of collision detection from the solution of the dynamic time-stepping subproblem, and approximations to the quadratic Coulomb friction model. This paper focuses on the development of *geometrically implicit* optimization-based time-stepper for dynamic simulation. More specifically, state-of-the-art time-steppers [16, 15, 9] use geometric information obtained from a collision detection algorithm at the current time, and the state of the system at the end of the time step is computed (by solving a dynamics time step subproblem) without modifying this information. Thus, state-of-the-art time-steppers can be viewed as *explicit* methods with respect to geometric information. We develop the first time-stepping method that is implicit in the geometric information (when the distance function is not available in closed form) by incorporating body geometry in the dynamic time-stepping subproblem. In other words, our formulation solves the collision detection and dynamic stepping problem in the same time-step, which allows us to satisfy contact constraints at the end of the time step. The resulting subproblem at each time-step will be a mixed nonlinear complementarity problem and we call our time-stepping scheme a *geometrically implicit* time-stepping scheme.

To illustrate the effects of geometric approximation, consider the simple planar problem of a uniform disc rolling on a horizontal support surface. For this problem, the exact solution is known, *i.e.*, the disc will roll at constant speed *ad infinitum*. However, when the disc is approximated by a uniform regular polygon, energy is lost a) due to collisions between the vertices and the support surface, b) due to contact sliding that is resisted by friction and c) due to artificial impulses generated by the approximate distance function that is to be satisfied at the end of the time-step. We simulated this example in dVC [3] using the Stewart-Trinkle time-stepping algorithm [16]. The parametric plots in Figure 1 show the reduction of kinetic energy over time caused by the accumulation of these effects. The top plot shows that increasing the number of edges, with the step-size fixed, decreases the energy loss; the energy loss approaches a limit determined by the size of the time-step. The bottom plot shows reducing energy loss with decreasing step size, with the number of vertices fixed at 1000. However, even with the decrease in time-step an energy loss limit is reached.



(a) As the number of edges of the “rolling” polygon increases, the energy loss decreases. The computed value obtained by our time-stepper using an implicit surface description of the disc is the horizontal line at the top. The time step used is 0.01 seconds.



(b) For a given number of edges, the energy loss decreases with decreasing step size, up to a limit. In this case, the limit is approximately 0.001 seconds (the plots for 0.001, 0.0005, and 0.0001 are indistinguishable).

Fig. 1. For a disc rolling on a surface, plots of the reduction of kinetic energy over time caused by approximating the disc as a uniform regular polygon.

These plots make it clear that the discretization of geometry and linearization of the distance function lead to the artifact of loss in energy in the simulations.

To address these issues and related problems that we have encountered (*e.g.*, parts feeding), we present a highly accurate geometrically implicit time-stepping method for convex objects described as an intersection of implicit surfaces. This method also takes into consideration other important nonlinear elements such as quadratic Coulomb friction. This method will provide a baseline for understanding and quantifying the errors incurred when using a geometrically explicit method and when making various linearizing approximations. Our ultimate goal is to develop techniques for automatically selecting the appropriate method for a given application, and to guide method switching, step size adjustment, and model approximations on the fly.

Our paper is organized as follows. In Section II we survey the relevant literature. Section III presents the dynamics model for multi-rigid-body systems with an elliptic dry friction law. In Section IV we develop a new formulation of the contact

constraints. In Section V, we give examples that validate and elucidate our time-stepping scheme. Finally we present our conclusions and lay out the future work.

II. RELATED WORK

Dynamics of multibody systems with unilateral contacts can be modeled as differential algebraic equations (DAE) [7] if the contact interactions (sliding, rolling, or separating) at each contact are known. However, in general, the contact interactions are not known *a priori*, but rather must be discovered as part of the solution process. To handle the many possibilities in a rigorous theoretical and computational framework, the model is formulated as a differential complementarity problem (DCP) [4, 17]. The differential complementarity problem is solved using a time-stepping scheme and the resultant system of equations to be solved at each step is a mixed (linear/nonlinear) complementarity problem. Let $\mathbf{u} \in \mathbb{R}^{n_1}$, $\mathbf{v} \in \mathbb{R}^{n_2}$ and let $g : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \rightarrow \mathbb{R}^{n_1}$, $f : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \rightarrow \mathbb{R}^{n_2}$ be two vector functions and the notation $\mathbf{0} \leq \mathbf{x} \perp \mathbf{y} \geq \mathbf{0}$ imply that \mathbf{x} is orthogonal to \mathbf{y} and each component of the vectors are non-negative.

Definition 1: The differential (or dynamic) complementarity problem is to find \mathbf{u} and \mathbf{v} satisfying

$$\begin{aligned} \dot{\mathbf{u}} &= g(\mathbf{u}, \mathbf{v}), & \mathbf{u}, \text{ free} \\ 0 &\leq \mathbf{v} \perp f(\mathbf{u}, \mathbf{v}) \geq 0 \end{aligned}$$

Definition 2: The mixed complementarity problem is to find \mathbf{u} and \mathbf{v} satisfying

$$\begin{aligned} g(\mathbf{u}, \mathbf{v}) &= 0, & \mathbf{u}, \text{ free} \\ 0 &\leq \mathbf{v} \perp f(\mathbf{u}, \mathbf{v}) \geq 0 \end{aligned}$$

The three primary modeling approaches for multibody systems with unilateral contacts are based on three different assumptions about the flexibility of the bodies. The assumptions from most to least realistic (and most to least computationally complex) are: 1) the bodies are fully deformable, 2) the bodies have rigid cores surrounded by compliant material, 3) the bodies are fully rigid. The first assumption leads to finite element approaches, for which one must solve very large difficult complementarity problems or variational inequalities at each time step. The second assumption leads to smaller subproblems that can be solved more easily [15, 13], but suitable values of the parameters of the compliant layer are difficult to determine. The assumption of rigid bodies leads to the smallest subproblems and avoids the latter problem of determining material deformation properties.

Independent of which underlying assumption is made, the methods developed to date have one problem in common that fundamentally limits their accuracy – they are not implicit with respect to the relevant geometric information. For example, at the current state, a collision detection routine is called to determine separation or penetration distances between the bodies, but this information is not incorporated as a function of the unknown future state at the end of the current time step. A goal of a typical time-stepping scheme is to guarantee consistency of the dynamic equations and all model constraints

at the end of each time step. However, since the geometric information at the end of the current time step is approximated from that at the start of the time step, the solution will be in error.

Early time-steppers used linear approximations of the local geometry at the current time [16, 1]. Thus each contact was treated as a point on a plane or a line on a (non-parallel) line and these entities were assumed constant for the duration of the time step. Besides being insufficiently accurate in some applications, some unanticipated side-effects arose [5].

Increased accuracy can be obtained in explicit schemes by including curvature information. This was done by Liu and Wang [9] and Pfeiffer and Glocker [14] by incorporating kinematic differential equations of rolling contact (Montana [12]). Outside the complementarity formalism, Kry and Pai [8] and Baraff [2] also make use of the contact kinematics equations in dynamic simulations of parametric and implicit surface objects respectively.

The method of Tzitzouris [19] is the only geometrically implicit method developed to date, but unfortunately it requires that the distance function between the bodies and two levels of derivatives be available in closed form. However, it is possible that this method could run successfully after replacing the closed-form distance functions with calls to collision detection algorithms and replacing derivatives with difference approximations from multiple collision detection calls, but polyhedral approximations common to most collision detection packages would generate very noisy derivatives. To our knowledge, such an implementation has not been attempted. One other problem with Tzitzouris' method is that it adapts its step size to precisely locate every collision time. While this is a good way to avoid interpenetration at the end of a time step, it has the undesirable side-effect of forcing the step size to be unreasonably small when there are many interacting bodies [10]. The method we propose does not suffer from this problem.

III. DYNAMIC MODEL FOR MULTIBODY SYSTEMS

In complementarity methods, the instantaneous equations of motion of a rigid multi-body system consist of five parts: (a) Newton-Euler equations, (b) Kinematic map relating the generalized velocities to the linear and angular velocities, (c) Equality constraints to model joints, (d) Normal contact condition to model intermittent contact, and (e) Friction law. Parts (a) and (b) form a system of ordinary differential equations, (c) is a system of (nonlinear) algebraic equations, (d) is a system of complementarity constraints, and (e) can be written as a system of complementarity constraints for any friction law that obeys the maximum work dissipation principle. In this paper we use an elliptic dry friction law [18]. Thus, the dynamic model is a *differential complementarity* problem. To solve this system of equations, we have to set up a discrete time-stepping scheme and solve a complementarity problem at each time step. We present below the continuous formulation as well as a Euler time-stepping scheme for discretizing the system. To simplify

the exposition, we ignore the presence of joints or bilateral constraints in the following discussion. However, all of the discussion below holds in the presence of bilateral constraints.

To describe the dynamic model mathematically, we first introduce some notation. Let \mathbf{q}_j be the position and orientation of body j in an inertial frame and $\boldsymbol{\nu}_j$ be the concatenated vector of linear velocities \mathbf{v} and angular velocities $\boldsymbol{\omega}$. The generalized coordinates, \mathbf{q} , and generalized velocity, $\boldsymbol{\nu}$ of the whole system are formed by concatenating \mathbf{q}_j and $\boldsymbol{\nu}_j$ respectively. Let λ_{in} be the normal contact force at the i th contact and $\boldsymbol{\lambda}_n$ be the concatenated vector of the normal contact forces. Let λ_{it} and λ_{io} be the orthogonal components of the friction force on the tangential plane at the i th contact and $\boldsymbol{\lambda}_t$, $\boldsymbol{\lambda}_o$ be the respective concatenated vectors. Let λ_{ir} be the frictional moment about the i th contact normal and $\boldsymbol{\lambda}_r$ be the concatenated vector of the frictional moments. The instantaneous dynamic model can then be written as follows.

Newton-Euler equations of motion:

$$\mathbf{M}(\mathbf{q})\dot{\boldsymbol{\nu}} = \mathbf{W}_n\boldsymbol{\lambda}_n + \mathbf{W}_t\boldsymbol{\lambda}_t + \mathbf{W}_o\boldsymbol{\lambda}_o + \mathbf{W}_r\boldsymbol{\lambda}_r + \boldsymbol{\lambda}_{app} + \boldsymbol{\lambda}_{vp} \quad (1)$$

where $\mathbf{M}(\mathbf{q})$ is the inertia tensor, $\boldsymbol{\lambda}_{app}$ is the vector of external forces, $\boldsymbol{\lambda}_{vp}$ is the vector of Coriolis and centripetal forces, \mathbf{W}_n , \mathbf{W}_t , \mathbf{W}_o , and \mathbf{W}_r are dependent on \mathbf{q} and map the normal contact forces, frictional contact forces, and frictional moments to the body reference frame.

Kinematic Map:

$$\dot{\mathbf{q}} = \mathbf{G}(\mathbf{q})\boldsymbol{\nu} \quad (2)$$

where \mathbf{G} is the matrix mapping the generalized velocity of the body to the time derivative of the position and orientation. The Jacobian \mathbf{G} may be a non-square matrix (e.g., using a unit quaternion to represent orientation) but $\mathbf{G}^T\mathbf{G} = \mathbf{I}$.

Nonpenetration Constraints: The normal contact constraint for the i th contact is

$$0 \leq \lambda_{in} \perp \psi_{in}(\mathbf{q}, t) \geq 0 \quad (3)$$

where ψ_{in} is a signed distance function or *gap function* for the i th contact with the property $\psi_{in}(\mathbf{q}, t) > 0$ for separation, $\psi_{in}(\mathbf{q}, t) = 0$ for touching, and $\psi_{in}(\mathbf{q}, t) < 0$ for interpenetration. The above gap function is defined in the configuration space of the system. Note that there is usually no closed form expression for $\psi_{in}(\mathbf{q}, t)$.

Friction Model:

$$(\lambda_{it}, \lambda_{io}, \lambda_{ir}) \in \operatorname{argmax}\{-v_{it}\lambda'_{it} + v_{io}\lambda'_{io} + v_{ir}\lambda'_{ir}\} : (\lambda'_{it}, \lambda'_{io}, \lambda'_{ir}) \in \mathbf{F}_i(\lambda_{in}, \mu_i)$$

where \mathbf{v}_i is the relative velocity at contact i and the friction ellipsoid is defined by $\mathbf{F}_i(\lambda_{in}, \mu_i) = \{(\lambda_{it}, \lambda_{io}, \lambda_{ir}) : \left(\frac{\lambda_{it}}{e_{it}}\right)^2 + \left(\frac{\lambda_{io}}{e_{io}}\right)^2 + \left(\frac{\lambda_{ir}}{e_{ir}}\right)^2 \leq \mu_i^2 \lambda_{in}^2\}$ where e_{it} , e_{io} and e_{ir} are given positive constants defining the friction ellipsoid and μ_i is the coefficient of friction at the i th contact.

We use a velocity-level formulation and an Euler time-stepping scheme to discretize the above system of equations. Let t_ℓ denote the current time, and h be the time step.

Use the superscripts ℓ and $\ell + 1$ to denote quantities at the beginning and end of the ℓ th time step respectively. Using $\dot{\mathbf{v}} \approx (\mathbf{v}^{\ell+1} - \mathbf{v}^\ell)/h$ and $\dot{\mathbf{q}} \approx (\mathbf{q}^{\ell+1} - \mathbf{q}^\ell)/h$, and writing in terms of the impulses we get the following discrete time system.

$$\begin{aligned} \mathbf{M}\mathbf{v}^{\ell+1} &= \mathbf{M}\mathbf{v}^\ell + h(\mathbf{W}_n\boldsymbol{\lambda}_n^{\ell+1} + \mathbf{W}_t\boldsymbol{\lambda}_t^{\ell+1} + \mathbf{W}_o\boldsymbol{\lambda}_o^{\ell+1} \\ &\quad + \mathbf{W}_r\boldsymbol{\lambda}_r^{\ell+1} + \boldsymbol{\lambda}_{\text{app}}^\ell + \boldsymbol{\lambda}_{\text{vp}}) \\ \mathbf{q}^{\ell+1} &= \mathbf{q}^\ell + h\mathbf{G}\mathbf{v}^{\ell+1} \\ 0 &\leq h\boldsymbol{\lambda}_n^{\ell+1} \perp \psi_n(\mathbf{q}^{\ell+1}) \geq 0 \\ h(\lambda_{it}^{\ell+1}, \lambda_{io}^{\ell+1}, \lambda_{ir}^{\ell+1}) &\in \text{argmax}\{-(v_{it}^{\ell+1})\lambda'_{it} + (v_{io}^{\ell+1})\lambda'_{io} \\ &\quad + (v_{ir}^{\ell+1})\lambda'_{ir}\} \\ &: h(\lambda_{it}^{\ell+1}, \lambda_{io}^{\ell+1}, \lambda_{ir}^{\ell+1}) \in \mathbf{F}_i(h\lambda_{in}, \mu_i) \end{aligned} \quad (4)$$

The argmax formulation of the friction law has a useful alternative formulation obtained from the Fritz-John optimality conditions [18]:

$$\begin{aligned} \mathbf{E}_t^2 \mathbf{U} \mathbf{p}_n \circ \mathbf{W}_t^T \mathbf{v}^{\ell+1} + \mathbf{p}_t \circ \boldsymbol{\sigma} &= 0 \\ \mathbf{E}_o^2 \mathbf{U} \mathbf{p}_n \circ \mathbf{W}_o^T \mathbf{v}^{\ell+1} + \mathbf{p}_o \circ \boldsymbol{\sigma} &= 0 \\ \mathbf{E}_r^2 \mathbf{U} \mathbf{p}_n \circ \mathbf{W}_r^T \mathbf{v}^{\ell+1} + \mathbf{p}_r \circ \boldsymbol{\sigma} &= 0 \\ (\mathbf{U} \mathbf{p}_n) \circ (\mathbf{U} \mathbf{p}_n) - (\mathbf{E}_t^2)^{-1} (\mathbf{p}_t \circ \mathbf{p}_t) - (\mathbf{E}_o^2)^{-1} (\mathbf{p}_o \circ \mathbf{p}_o) \\ &\quad - (\mathbf{E}_r^2)^{-1} (\mathbf{p}_r \circ \mathbf{p}_r) \geq 0 \end{aligned} \quad (5)$$

where the impulse $\mathbf{p}_{(\cdot)} = h\boldsymbol{\lambda}_{(\cdot)}$, the matrices \mathbf{E}_t , \mathbf{E}_o , \mathbf{E}_r , and \mathbf{U} are diagonal with i th diagonal element equal to e_i , e_o , e_r , and μ_i respectively, $\boldsymbol{\sigma}$ is a concatenated vector of the Lagrange multipliers arising from the conversion from the argmax formulation and σ_i is equal to the magnitude of the slip velocity at contact i , and \circ connotes the Hadamard product.

The above subproblem at a time step is either an LCP or an NCP depending on the time evaluation of $\mathbf{W}_{(\cdot)}$ the approximation used for $\psi_n(\mathbf{q}^{\ell+1})$, and the representation of the friction model. If $\mathbf{W}_{(\cdot)}$ are evaluated at ℓ , and we use a first order Taylor series expansion for $\psi_n(\mathbf{q})$ and a linearized representation of the friction ellipsoid, we have an LCP. However, the approximations involved introduce numerical artifacts as discussed in Section I. Moreover, the linear approximation of the friction ellipsoid also leads to certain artifacts. In contrast, if we evaluate $\mathbf{W}_{(\cdot)}$ at $\ell + 1$, use a quadratic friction law (Equation (5)), and use $\psi_n(\mathbf{q}^{\ell+1})$, we have an NCP. We call this formulation a *geometrically implicit* formulation because it ensures that the contact conditions are satisfied at the end of the time step. However, evaluating $\psi_n(\mathbf{q}^{\ell+1})$ is possible only if we have a closed form expression for the distance function, which we do not have in general. Instead, we propose to define the gap function in terms of the closest points between the two objects and provide a set of algebraic equations for finding these closest points during the time step. The next section discusses this approach in detail and proves that the conditions will enforce satisfaction of contact constraints at the end of the time step.

IV. CONTACT CONSTRAINT

In this section we rewrite the contact condition (Equation 3) as a complementarity condition in the work space, combine it with an optimization problem to find the closest points, and prove that the resultant system of equations ensures that the contact constraints are satisfied. Let us consider the i th contact. For ease of exposition, we assume here that each object is a convex object described by a single implicit surface. A more general formulation where each object is described by an intersection of implicit surfaces is given in Appendix A. Let the two objects be defined by convex functions $f(\xi_1) \leq 0$ and $g(\xi_2) \leq 0$ respectively, where ξ_1 and ξ_2 are the coordinates of points in the two objects. Let \mathbf{a}_1 and \mathbf{a}_2 be the closest points on the two objects. The equation of an implicit surface has the property that for any point \mathbf{x} , the point lies inside the object for $f(\mathbf{x}) < 0$, on the object surface for $f(\mathbf{x}) = 0$, and outside the object for $f(\mathbf{x}) > 0$. Thus, we can define the gap function in work space as either $f(\mathbf{a}_2)$ or $g(\mathbf{a}_1)$ and write the complementarity conditions as either one of the following two conditions:

$$\begin{aligned} 0 &\leq \lambda_{in} \perp f(\mathbf{a}_2) \geq 0 \\ 0 &\leq \lambda_{in} \perp g(\mathbf{a}_1) \geq 0 \end{aligned} \quad (6)$$

where \mathbf{a}_1 and \mathbf{a}_2 are given by

$$\text{argmin} \{ \|\xi_1 - \xi_2\|^2 : f(\xi_1) \leq 0, g(\xi_2) \leq 0 \} \quad (7)$$

It can be shown easily from the Karush-Kuhn-Tucker (KKT) conditions of Equation 7 that \mathbf{a}_1 and \mathbf{a}_2 are the solutions of the following system of algebraic equations.

$$\begin{aligned} \mathbf{a}_1 - \mathbf{a}_2 &= -l_1 \nabla f(\mathbf{a}_1) = l_2 \nabla g(\mathbf{a}_2) \\ f(\mathbf{a}_1) &= 0, \quad g(\mathbf{a}_2) = 0 \end{aligned} \quad (8)$$

where l_1 and l_2 are the Lagrange multipliers. The geometric meaning of the first two equations is that the normals to the two surfaces at their closest points are aligned with the line joining the closest points. The solution to Equation 8 gives the closest point when the two objects are separate. However, when $\mathbf{a}_1 = \mathbf{a}_2$, the solution is either the touching point of the two surfaces or a point lying on the intersection curve of the two surfaces. Thus, as written, Equations 8 and 6 do not guarantee non-penetration. However, note that the distinction

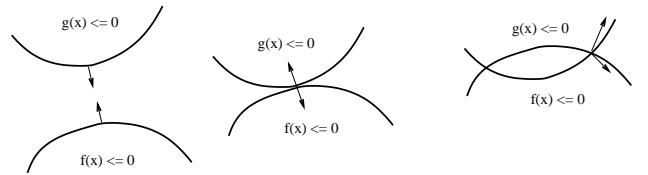


Fig. 2. Three Contact cases: (left) Objects are separate (middle) Objects are touching (right) Objects are intersecting.

between touching points and intersecting points is that the normals to the two surfaces at the touching points are aligned while it is not so for intersection points. When $\mathbf{a}_1 = \mathbf{a}_2$, we lose the geometric information that the normals at the two points are aligned if we write our equations in the form above.

Rewriting the above equations in terms of the unit vectors allows us to avoid this problem.

$$\begin{aligned} \mathbf{a}_1 - \mathbf{a}_2 &= -\|\mathbf{a}_1 - \mathbf{a}_2\| \frac{\nabla f(\mathbf{a}_1)}{\|\nabla f(\mathbf{a}_1)\|} \\ \frac{\nabla f(\mathbf{a}_1)}{\|\nabla f(\mathbf{a}_1)\|} &= -\frac{\nabla g(\mathbf{a}_2)}{\|\nabla g(\mathbf{a}_2)\|} \\ f(\mathbf{a}_1) &= 0, \quad g(\mathbf{a}_2) = 0 \end{aligned} \quad (9)$$

Proposition: Equation 6 and 9 together represent the contact constraints, i.e., the two objects will satisfy the contact constraints at the end of each time step if and only if Equation 6 and 9 hold together.

Proof: As discussed above.

Note that since the first two vector equations are equating unit vectors, there are only two independent equations for each, and the above system has 6 independent equations in 6 variables. We can now formulate the mixed NCP for the geometrically-implicit time-stepper. The vector of unknowns \mathbf{z} can be partitioned into $\mathbf{z} = [\mathbf{u}, \mathbf{v}]$ where $\mathbf{u} = [\boldsymbol{\nu}, \mathbf{a}_1, \mathbf{a}_2, \mathbf{p}_t, \mathbf{p}_o, \mathbf{p}_r]$ and $\mathbf{v} = [\mathbf{p}_n, \boldsymbol{\sigma}]$. The equality constraints in the mixed NCP are:

$$\begin{aligned} 0 &= -\mathbf{M}\boldsymbol{\nu}^{\ell+1} + \mathbf{M}\boldsymbol{\nu}^\ell + \mathbf{W}_n^{\ell+1}\mathbf{p}_n^{\ell+1} + \mathbf{W}_t^{\ell+1}\mathbf{p}_t^{\ell+1} \\ &\quad + \mathbf{W}_o^{\ell+1}\mathbf{p}_o^{\ell+1} + \mathbf{W}_r^{\ell+1}\mathbf{p}_r^{\ell+1} + \mathbf{p}_{\text{app}}^\ell + \mathbf{p}_{\text{vp}}^\ell \\ 0 &= (\mathbf{a}_1^{\ell+1} - \mathbf{a}_2^{\ell+1}) + \|\mathbf{a}_1^{\ell+1} - \mathbf{a}_2^{\ell+1}\| \frac{\nabla f(\mathbf{a}_1^{\ell+1})}{\|\nabla f(\mathbf{a}_1^{\ell+1})\|} \\ 0 &= \frac{\nabla f(\mathbf{a}_1^{\ell+1})}{\|\nabla f(\mathbf{a}_1^{\ell+1})\|} + \frac{\nabla g(\mathbf{a}_2^{\ell+1})}{\|\nabla g(\mathbf{a}_2^{\ell+1})\|} \\ 0 &= f(\mathbf{a}_1^{\ell+1}) \\ 0 &= g(\mathbf{a}_2^{\ell+1}) \\ 0 &= \mathbf{E}_t^2 \mathbf{U} \mathbf{p}_n^{\ell+1} \circ (\mathbf{W}_t^T)^{\ell+1} \boldsymbol{\nu}^{\ell+1} + \mathbf{p}_t^{\ell+1} \circ \boldsymbol{\sigma}^{\ell+1} \\ 0 &= \mathbf{E}_o^2 \mathbf{U} \mathbf{p}_n^{\ell+1} \circ (\mathbf{W}_o^T)^{\ell+1} \boldsymbol{\nu}^{\ell+1} + \mathbf{p}_o^{\ell+1} \circ \boldsymbol{\sigma}^{\ell+1} \\ 0 &= \mathbf{E}_r^2 \mathbf{U} \mathbf{p}_n^{\ell+1} \circ (\mathbf{W}_r^T)^{\ell+1} \boldsymbol{\nu}^{\ell+1} + \mathbf{p}_r^{\ell+1} \circ \boldsymbol{\sigma}^{\ell+1} \end{aligned} \quad (10)$$

The complementarity constraints on \mathbf{v} are:

$$0 \leq \begin{bmatrix} \mathbf{p}_n^{\ell+1} \\ \boldsymbol{\sigma}^{\ell+1} \end{bmatrix} \perp \begin{bmatrix} f(\mathbf{a}_2^{\ell+1}) \\ \boldsymbol{\zeta} \end{bmatrix} \geq 0 \quad (11)$$

where

$$\begin{aligned} \boldsymbol{\zeta} &= \mathbf{U} \mathbf{p}_n^{\ell+1} \circ \mathbf{U} \mathbf{p}_n^{\ell+1} - (\mathbf{E}_t^2)^{-1} (\mathbf{p}_t^{\ell+1} \circ \mathbf{p}_t^{\ell+1}) \\ &\quad - (\mathbf{E}_o^2)^{-1} (\mathbf{p}_o^{\ell+1} \circ \mathbf{p}_o^{\ell+1}) - (\mathbf{E}_r^2)^{-1} (\mathbf{p}_r^{\ell+1} \circ \mathbf{p}_r^{\ell+1}) \end{aligned}$$

In the above formulation, we see $\mathbf{u} \in \mathbb{R}^{6n_b+9n_c}$, $\mathbf{v} \in \mathbb{R}^{2n_c}$, the vector function of equality constraints maps $[\mathbf{u}, \mathbf{v}]$ to $\mathbb{R}^{6n_b+9n_c}$ and the vector function of complementarity constraints maps $[\mathbf{u}, \mathbf{v}]$ to \mathbb{R}^{2n_c} where n_b and n_c are the number of bodies and number of contacts respectively. If using convex bodies only, the number of contacts can be determined directly from the number of bodies, $n_c = \sum_{i=1}^{n_b-1} i$.

V. ILLUSTRATIVE EXAMPLES

In this section we present three examples to validate our technique against known analytical results and previous approaches. The first example is the same example of a disc

rolling without slip on a plane that we studied in Section I. The second example, taken from [18], consists of a sphere spinning on a horizontal surface that comes to rest due to torsional friction. The time taken by the sphere to come to a complete stop is known analytically and we demonstrate that the results of our simulation agree with the analytical predictions. The final example consists of a small ball moving in contact with two larger fixed balls. We include it here to compare our solutions with those based on earlier approaches [18, 9]. All of our numerical results were obtained by PATH [6], a free solver that is one of the most robust complementarity problem solvers available.

A. Example 1: Disc on a Plane

In this example we revisit the unit disc example from Section I. For illustrative purposes, we explain the formulation of the full dynamic model in detail. The normal axis of the contact frame \hat{n} always points in the inertial y -axis direction and tangential axis \hat{t} always coincides with the x -direction. The mass matrix, \mathbf{M} is constant and the only force acting on the body is due to gravity. The equation of the disc is given by $f_1(x, y) = (x - q_x)^2 + (y - q_y)^2 - 1$, where \mathbf{q} is the location of the center of the disc in the inertial frame. Let \mathbf{a}_1 be the closest point on body 1 (the disc) to the x -axis. Similarly, let \mathbf{a}_2 be the closest point on the x -axis to body 1 ($a_{2y} = 0$ and can be removed from the system of unknowns). Given this information, the matrices for this system can be shown to be: $\mathbf{M} = \text{diag}(m, m, 0.5m)$ $\mathbf{p}_{\text{app}} = [0, -9.81 \cdot m \cdot h, 0]^T$.

$$\begin{aligned} \mathbf{W}_n &= \begin{bmatrix} \hat{n} \\ \mathbf{r}^{a_1} \otimes \hat{n} \end{bmatrix} & \mathbf{W}_t &= \begin{bmatrix} \hat{t} \\ \mathbf{r}^{a_1} \otimes \hat{t} \end{bmatrix} \\ \nabla_{a_1} f_1(a_1^{\ell+1}) &= \begin{bmatrix} 2(a_{1x}^{\ell+1} - q_x) \\ 2(a_{1y}^{\ell+1} - q_y) \end{bmatrix} \end{aligned}$$

where \mathbf{r}^{a_1} is the vector from the center of gravity of the disc to \mathbf{a}_1 and \otimes denotes the 2D analog of the cross product.

There are 9 unknowns for this system: $\mathbf{z} = [\boldsymbol{\nu}, \mathbf{a}_1, a_{2x}, p_n, p_t, \boldsymbol{\sigma}]$. We can now formulate the entire system of equations for this simple model:

$$0 = -\mathbf{M}\boldsymbol{\nu}^{\ell+1} + \mathbf{M}\boldsymbol{\nu}^\ell + \mathbf{W}_n^{\ell+1}\mathbf{p}_n^{\ell+1} + \mathbf{W}_t^{\ell+1}\mathbf{p}_t^{\ell+1} + \mathbf{p}_{\text{app}} \quad (12)$$

$$0 = \mathbf{a}_2^{\ell+1} - \mathbf{a}_1^{\ell+1} + \|\mathbf{a}_2^{\ell+1} - \mathbf{a}_1^{\ell+1}\| \hat{n} \quad (13)$$

$$0 = \frac{\nabla_{a_1} f_1(\mathbf{a}_1^{\ell+1})}{\|\nabla_{a_1} f_1(\mathbf{a}_1^{\ell+1})\|} + \hat{n} \quad (14)$$

$$0 = f_1(\mathbf{a}_1^{\ell+1}) \quad (15)$$

$$0 \leq f_1(\mathbf{a}_2^{\ell+1}) \quad (16)$$

$$0 = \mu p_n^{\ell+1} (\mathbf{W}_t^T)^{\ell+1} \boldsymbol{\nu}^{\ell+1} + \sigma^{\ell+1} p_t^{\ell+1} \quad (17)$$

$$0 \leq \mu^2 p_n^{\ell+1} p_n^{\ell+1} - p_t^{\ell+1} p_t^{\ell+1} \quad (18)$$

where equations 13 and 14 each provide one independent equation.

The initial configuration of the disc is $\mathbf{q} = [0, 1, 0]$, initial velocity is $\boldsymbol{\nu} = [-3, 0, 3]$, mass is $m = 1$, and $\mu = 0.4$.

Figure 1(a) shows the kinetic energy of the disc for our implicit representation along with the Stewart-Trinkle LCP implementation using various levels of discretization as it rolls along the horizontal surface. When using an implicit curve representation to model the disc and our formulation we get no energy loss (within the numerical tolerance of 10^{-6} used for our simulations) as seen by the horizontal line. When using the LCP formulation we have energy loss as discussed earlier.

B. Example 2: Sphere on a Plane

Here we consider a sphere spinning on a plane about the normal of the plane. The initial configuration of the sphere is $\mathbf{q} = [0, 0, 1, 1, 0, 0, 0]^T$ where the first three elements are the position of the center of the sphere and the last 4 are the unit quaternion representing the orientation. The initial generalized velocity is $\boldsymbol{\nu} = [0, 0, 0, 0, 0, 1.962]^T$, the first three elements are the linear terms, the last three elements the angular. The friction parameters used were $e_t = 1$, $e_o = 1$, $e_r = 0.4$ and $\mu = 0.2$. A step size of $h = 0.07$ seconds was chosen.

From the initial conditions given, the sphere should rotate in place with a constant deceleration due to the torsional friction. Figure 3 shows a plot of the velocities for the sphere given by the time-stepping formulation. The analytical solution for this problem predicts that all velocities except ω_z should be zero, and w_z should be decreasing linearly to 0 with a slope of -1.962, reaching 0 at exactly $t = 1$ seconds. The plot shows that we agree with the analytical solution except that we reach zero velocity at $t = 1.05$, since we are using a fixed time step and the time at which spinning stops is in the middle of the time step. The friction forces (Figure 4) also follow the analytical solution. The tangential component of friction force is 0. The tangential moment does not drop to 0 at 1.05 s, since we are using an impulse-based time-stepping formulation with a fixed time step and there is a torsional moment between 0.98 to 1 second which contributes to the impulse. Our results match those of the Tzitzouris-Pang and Stewart methods presented in [18].

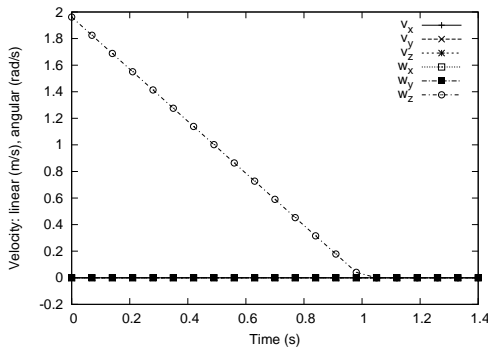


Fig. 3. Linear and angular velocities for Example 2. All velocities except w_z are zero throughout the simulation.

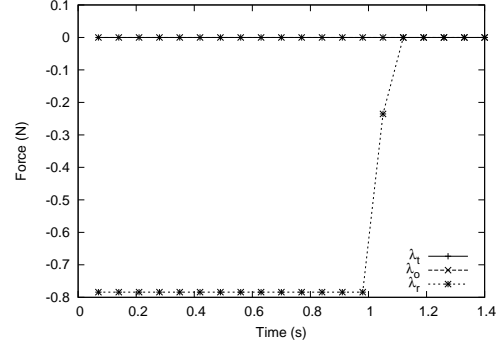


Fig. 4. Forces for Example 2. The tangential forces are both 0 for the entire simulation, and the torsional force transitions to zero when the sphere switches from a sliding contact to sticking.

C. Example 3: Sphere on Two Spheres

This example consists of a small ball rolling and sliding on two larger balls, and is chosen to compare our model with those presented in [18] and [9]. Figure 5 shows a small unit sphere in simultaneous contact with two larger fixed spheres. The sphere of radius 10 units is located at $(0, 0, 0)$ in the inertial frame and the sphere of radius 9 units is located at $(0, 11.4, 0)$. There is also a constant force of $\lambda_{app} = [1.0, 2.6, -9.81, 0, 0, 0]^T$ applied to the small sphere. With this force, the sphere initially has one of its contacts rolling while the other contact is simultaneously sliding, the rolling contact transitions to sliding, and both contacts eventually separate. It is important to emphasize that all these transitions are captured using a fixed time step implementation.

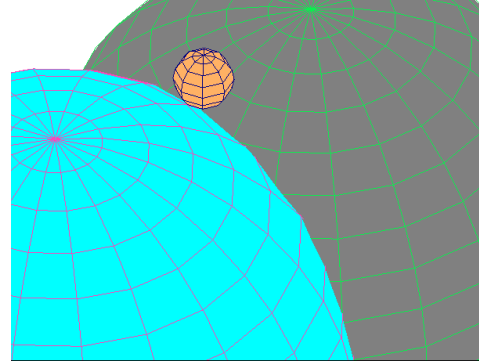


Fig. 5. A small sphere in contact with two large spheres.

The initial configuration of the small sphere is $\mathbf{q} = [0, 6.62105263157895, 8.78417110772903, 1, 0, 0, 0]^T$. The initial velocity is $\boldsymbol{\nu} = [0, 0, 0, 0, 0, 0]$. The friction parameters are: $e_t = 1$, $e_o = 1$, $e_r = 0.3$, and $\mu = 0.2$. There were a total of 28 unknowns in our NCP formulation. We used a step size $h = 0.01$ (Tzitzouris-Pang use $h = 0.1$).

The generalized velocity of the sphere is shown in Figure 6. The smooth velocity profile agrees well with the nonlinear Tzitzouris-Pang formulation [18]. The Liu-Wang formulation [9] experienced non-smooth velocity jumps when the small sphere separated from the larger fixed spheres,

which they attributed to an explicit time-stepping scheme. In the LCP Stewart-Trinkle implementation, the velocity profiles were very non-smooth. These results further confirm our belief that both linearization and explicit time-stepping lead to inaccuracies.

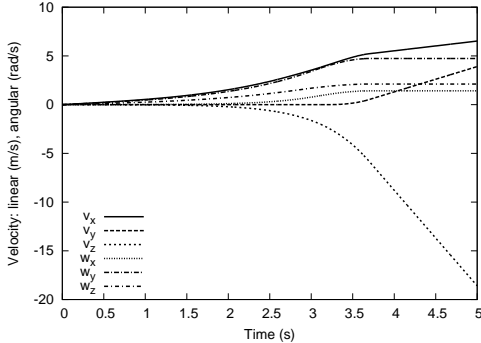


Fig. 6. Velocities of small moving sphere.

The fidelity of our method is further emphasized by Figures 7 and 8 that show the forces and sliding speed magnitudes at the two contacts. Contact 1 starts as a sliding contact and we see the sliding speed increases as the normal force decreases. Also, the magnitude of the friction force is equal to $\mu\lambda_{1n}$, consistent with our friction law for a sliding contact. At approximately 3.2 seconds, the small sphere separates from the large sphere at this contact, and all forces acting at contact 1 and the sliding speed drop to zero. Contact 2 on the other hand starts out as a rolling contact until approximately $t = 3$ seconds when it transitions to sliding. During the rolling phase the frictional magnitude is bounded by $\mu\lambda_{2n}$ as required by the friction law, and the sliding speed is 0. At the transition to sliding, the magnitude of the friction force becomes equal to $\mu\lambda_{2n}$ and the sliding speed begins to increase. Finally, at approximately $t = 3.6$ seconds, the contact breaks and all forces at this contact and the sliding speed drop to zero.

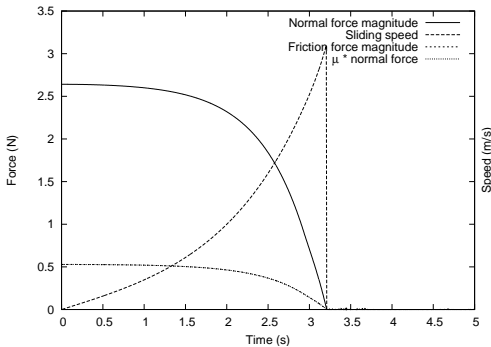


Fig. 7. Force and sliding speed at contact 1. Contact 1 is always sliding until separation, hence the μ normal force curve and friction magnitude curve overlap for the duration. The value of $\mu = 0.2$

Unlike the other approaches, we modeled the spheres as special cases of an ellipsoid:

$$f(x, y, z) = \left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 + \left(\frac{z}{c}\right)^2 - 1$$

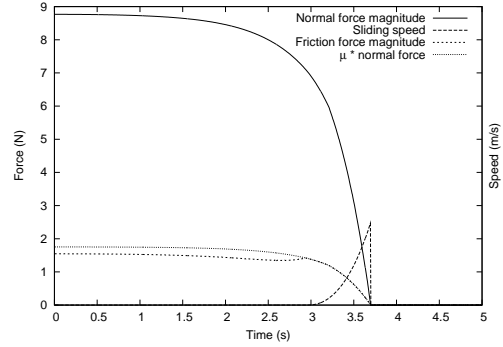


Fig. 8. Force and sliding speed at contact 2. The value of $\mu = 0.2$

where we set $a = b = c = 1$ for a unit sphere centered at the origin. This flexibility allows us to test various shapes by altering the 3 parameters, while the more difficult dynamical equations remain unaltered. For example, changing the a, b, c parameters would allow us to model ellipsoids rolling on ellipsoids; the Tzitzouris-Pang formulation cannot be applied since there is no analytical gap function.

VI. CONCLUSION

We presented the first geometrically implicit time-stepping scheme for objects described as intersections of convex inequalities. This approach overcomes stability and accuracy problems associated with polygonal approximations of smooth objects and approximation of the distance function for two objects in intermittent contact. We developed a new formulation for the contact constraints in the work space which enabled us to formulate a geometrically implicit time-stepping scheme as an NCP. We demonstrated through example simulations the fidelity of this approach to analytical solutions and previously described simulation results.

We see several directions for future work. We would like to address the question of existence and uniqueness of solutions of the NCP we formulate. We will perform more extensive numerical experimentation, and compare these solutions with solutions obtained when the closest distance is computed through a function call. We plan to precisely quantify the tradeoffs between the computation speed and physical accuracy of simulations for different object representations (*e.g.*, polyhedral, implicit, spline), friction approximations, and the choice between geometrically explicit or implicit methods. Although we have restricted our discussion to convex objects, we believe that this framework can be extended to non-convex objects described as unions of convex objects as well as parametric surfaces. Finally, we want to incorporate different impact laws to simulate a broader class of problems.

ACKNOWLEDGMENT

We wish to thank Todd Munson for his insightful discussions on NCPs. This work was supported by the National Science Foundation under grants IIS-0093233(CAREER), 0139701 (DMS-FRG), 0413227 (IIS-RCV), and 0420703

(MRI). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

- [1] M. Anitescu, J. F. Cremer, and F. A. Potra. Formulating 3D contact dynamics problems. *Mechanics of structures and machines*, 24(4):405, 1996.
- [2] D. Baraff. Curved surfaces and coherence for non-penetrating rigid body simulation. *Computer Graphics*, 24(4):19–28, August 1990.
- [3] S. Berard, J. Trinkle, B. Nguyen, B. Roghani, V. Kumar, and J. Fink. daVinci Code: A multi-model simulation and analysis tool for multi-body systems. In *IEEE International Conference on Robotics and Automation*, pages 2588–2593, Rome, Italy, Apr. 2007.
- [4] R. W. Cottle, J. Pang, and R. E. Stone. *The Linear Complementarity Problem*. Academic Press, 1992.
- [5] K. Egan, S. Berard, and J. Trinkle. Modeling nonconvex constraints using linear complementarity. Technical Report 03-13, Department of Computer Science, Rensselaer Polytechnic Institute, 2003.
- [6] M. C. Ferris and T. S. Munson. Complementarity problems in GAMS and the PATH solver. *Journal of Economic Dynamics and Control*, 24(2):165–188, Feb. 2000.
- [7] E. J. Haug, S. C. Wu, and S. M. Yang. Dynamics of mechanical systems with Coulomb friction, stiction, impact and constraint addition-deletion theory. *Mechanism and Machine Theory*, 21:365–446, 1986.
- [8] P. G. Kry and D. K. Pai. Continuous contact simulation for smooth surfaces. *ACM Transactions on Graphics*, 22(1):106–129, Jan. 2003.
- [9] T. Liu and M. Y. Wang. Computation of three-dimensional rigid-body dynamics with multiple unilateral contacts using time-stepping and Gauss-Seidel methods. *IEEE Transactions on Automation Science and Engineering*, 2(1):19–31, January 2005.
- [10] B. Mirtich. *Impulse-based Dynamics Simulation of Rigid Body Systems*. PhD thesis, University of California, Berkeley, 1996.
- [11] B. Mirtich, Y. Zhuang, K. Goldberg, J. Craig, R. Zanutta, B. Carlisle, and J. Canny. Estimating pose statistics for robotic part feeders. In *IEEE International Conference on Robotics and Automation*, pages 1140–1146, Minneapolis, MN, Apr. 1996.
- [12] D. J. Montana. The kinematics of contact and grasp. *International Journal of Robotics Research*, 7(3):17–32, June 1998.
- [13] M. Pauly, D. K. Pai, and L. J. Guibas. Quasi-rigid objects in contact. In *ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, August 2004.
- [14] F. Pfeiffer and C. Glocker. *Multibody Dynamics with Unilateral Constraints*. John Wiley, New York, 1996.
- [15] P. Song, J.-S. Pang, and V. Kumar. A semi-implicit time-stepping model for frictional compliant contact problems. *International Journal for Numerical Methods in Engineering*, 60:2231–2261, 2004.
- [16] D. Stewart and J. Trinkle. An implicit time-stepping scheme for rigid body dynamics with inelastic collisions and Coulomb friction. *International Journal of Numerical Methods in Engineering*, 39:2673–2691, 1996.
- [17] J. Trinkle, J. Pang, S. Sudarsky, and G. Lo. On dynamic multi-rigid-body contact problems with Coulomb friction. *Zeitschrift für Angewandte Mathematik und Mechanik*, 77(4):267–279, 1997.
- [18] J. C. Trinkle, J. Tzitzouris, and J. S. Pang. Dynamic multi-rigid-body systems with concurrent distributed contacts: Theory and examples. *Philosophical Transactions on Mathematical, Physical, and Engineering Sciences, Series A*, 359(1789):2575–2593, Dec. 2001.
- [19] J. E. Tzitzouris. *Numerical Resolution of Frictional Multi-Rigid-Body Systems via Fully Implicit Time-Stepping and Nonlinear Complementarity*. PhD thesis, Mathematics Department, Johns Hopkins University, 2001.

APPENDIX

CONTACT CONDITIONS FOR OBJECTS REPRESENTED AS INTERSECTIONS OF CONVEX INEQUALITIES

We present here the contact conditions for the general case where each convex object is defined as an intersection of convex inequalities. Let $f_j(\xi_1) \leq 0, j = 1, \dots, m, g_j(\xi_2) \leq$

$0, j = m + 1, \dots, n$, be convex functions representing the two convex objects. Since the closest point is outside the object if it is outside at least one of the intersecting surfaces, the complementarity conditions for nonpenetration can be written as either one of the following two sets of conditions:

$$\begin{aligned} 0 &\leq \lambda_{in} \perp \max\{f_j(\mathbf{a}_2)\} \geq 0 \quad j = 1, \dots, m \\ 0 &\leq \lambda_{in} \perp \max\{g_j(\mathbf{a}_1)\} \geq 0 \quad j = m + 1, \dots, n \end{aligned} \quad (19)$$

where \mathbf{a}_1 and \mathbf{a}_2 are the closest points on the two bodies and are given by the KKT conditions

$$\begin{aligned} \mathbf{a}_1 - \mathbf{a}_2 &= - \sum_{i=1}^m l_i \nabla f_i(\mathbf{a}_1) = \sum_{j=m+1}^n l_j \nabla g_j(\mathbf{a}_2) \\ f_i(\mathbf{a}_1) + s_i &= 0 \\ g_j(\mathbf{a}_2) + s_j &= 0 \\ l_i s_i = l_j s_j &= 0 \\ s_i \geq 0, \quad s_j &\geq 0 \end{aligned} \quad (20)$$

where s_i, s_j are the slack variables. At the optimal solution only some of the constraints are active. Thus the optimality conditions can be written as the following set of nonlinear equations:

$$\begin{aligned} \mathbf{a}_1 - \mathbf{a}_2 &= - \sum_{i \in \mathbb{I} \cap \{i\}} l_i \nabla f_i(\mathbf{a}_1) = \sum_{j \in \mathbb{I} \cap \{j\}} l_j \nabla g_j(\mathbf{a}_2) \\ f_k(\mathbf{a}_1) &= 0 \quad k \in \mathbb{I} \cap \{i\} \\ g_k(\mathbf{a}_2) &= 0 \quad k \in \mathbb{I} \cap \{j\} \end{aligned} \quad (21)$$

where \mathbb{I} is the index set of active constraints. Equations 19 and 21 together represent the contact constraints as long as $\mathbf{a}_1 \neq \mathbf{a}_2$. Using arguments similar to the single surface case in Section IV we can see that it is not possible to distinguish between touching points and intersecting points using the above formulation. In this case also, we can rewrite Equation 21 suitably by equating unit vectors to eliminate the intersection point solutions. Without loss of generality, we can set one of the Lagrange multipliers to be 1 and scale the others and rewrite Equation 21 as

$$\begin{aligned} \mathbf{a}_1 - \mathbf{a}_2 &= \|\mathbf{a}_1 - \mathbf{a}_2\| \left(\frac{\nabla f_{k_1}(\mathbf{a}_1)}{\|\nabla f_{k_1}(\mathbf{a}_1)\|} \right. \\ &\quad \left. + \sum_{k \in \{\mathbb{I} \setminus k_1\} \cap \{i\}} l_k \frac{\nabla f_k(\mathbf{a}_1)}{\|\nabla f_k(\mathbf{a}_1)\|} \right) \\ &\quad + \frac{\nabla f_{k_1}(\mathbf{a}_1)}{\|\nabla f_{k_1}(\mathbf{a}_1)\|} + \sum_{k \in \{\mathbb{I} \setminus k_1\} \cap \{i\}} l_k \frac{\nabla f_k(\mathbf{a}_1)}{\|\nabla f_k(\mathbf{a}_1)\|} \\ &= \frac{\nabla g_{k_2}(\mathbf{a}_2)}{\|\nabla g_{k_2}(\mathbf{a}_2)\|} + \sum_{k \in \{\mathbb{I} \setminus k_2\} \cap \{j\}} l_k \frac{\nabla g_k(\mathbf{a}_2)}{\|\nabla g_k(\mathbf{a}_2)\|} \\ f_k(\mathbf{a}_1) &= 0 \quad k \in \mathbb{I} \cap \{i\} \\ g_k(\mathbf{a}_2) &= 0 \quad k \in \mathbb{I} \cap \{j\} \end{aligned} \quad (22)$$

Proposition: Equation 19 and 22 together represent the non-penetration constraints, i.e., the two objects will satisfy the contact constraints at the end of each time step if and only if Equation 19 and 22 hold together.