AN IMPLICIT TIME-STEPPING SCHEME FOR RIGID BODY DYNAMICS WITH INELASTIC COLLISIONS AND COULOMB FRICTION[0]

D.E. STEWART[1] AND J.C. TRINKLE[2]

ABSTRACT. In this paper a new time-stepping method for simulating systems of rigid bodies is given which incorporates Coulomb friction and inelastic impacts and shocks. Unlike other methods which take an instantaneous point of view, this method does not need to explicitly identify impulsive forces. Instead the treatment is similar to that of J.J. Moreau and Monteiro-Marques, except that the numerical formulation used here ensures that there is no interpenetration of rigid bodies, unlike their velocity-based formulation. Numerical results are given for the method presented here for a spinning rod impacting a table in two dimensions, and a system of four balls colliding on a table in a fully three-dimensional way. These numerical results also show the practicality of the method, and convergence of the method as the step size becomes small.

1. INTRODUCTION

This work is concerned with the numerical solution of *rigid body* dynamics. This is usually conceived as a limiting case of increasingly stiff elastic bodies. While it is possible to model the elastic bodies directly as a Signorini problem with friction[23], or approximately as a system of springs and masses[30], the level of detail in the model, and the stiffness of the corresponding differential equations, makes this approach computationally expensive. Instead, here we consider rigid systems with a finite number of degrees of freedom. This problem has been studied by a number of authors in recent years in both the robotics and mathematics communities [3, 4, 16, 26, 27, 28, 29, 31, 32, 34, 35].

Truly rigid bodies which collide must suffer impulsive forces and velocity discontinuities. Physically, velocity discontinuities are impossible, but correspond to a mathematical model which involves the rigid body limit of increasingly stiff elastic bodies where an extremely large contact force appears for a correspondingly short time interval. The limit of such forces are mathematical impulses which can be understood as Dirac-δ (generalized) functions[18], distributions[13, 18], or measures[6, 11].

The computational model developed here uses the Coulomb friction law, in spite of the paradoxes of Painlevé[33] and Delassus[9, 10]. This paper avoids these paradoxes by allowing *shocks*, or impulsive forces without collisions[3, 28, 32]. Baraff[3] computes shocks as unbounded rays generated by Lemke's algorithm for LCP's. However,

---

[1]Mathematics Dept., Texas A&M University, College Station, TX 77843-3368
[2]Computer Science Dept., Texas A&M University, College Station, TX 77843-3112

this does not always satisfy the dynamic model. The method developed here has dissipative impacts and shocks, unlike the conventional Newton and Poisson impact hypotheses which predict an increase the kinetic energy for certain complex collisions. However, the dissipative formulation of Stronge[41] is not used here.

Unlike most numerical methods for rigid body mechanics with collisions[3, 4, 16, 26, 27, 34, 35], the method presented here is a *time-stepping* algorithm. This allows the incorporation of impulses without difficulty as the method uses the *integrals* of the forces over each time-step, which are finite even if there are impulsive forces. This means that this method may violate the *"Principle of constraints"*[22] that states that impulses can only arise if there is no solution with finite forces. In fact, there is no direct way with this method to determine if impulsive forces arise, as opposed to large, finite forces. Rather, it is the behavior of the forces as the step size approaches zero that will determine if impulsive forces occur in the continuous problem. In this respect it is similar to the computational schemes of J.J. Moreau[31, 32] and Monteiro-Marques[28]. On the other hand, the method presented here produces numerical trajectories that do not violate the rigid body constraints, at least at the ends of each time-step, unlike the methods of Moreau[31, 32] and Monteiro-Marques[28] which allow the system to "drift" into inadmissible states. This is because the method presented here is a *position based* method that uses a complementarity principle for the generalized position at the end of each time-step. By contrast, the methods of Moreau[31, 32] and Monteiro-Marques[28] use a condition for the *velocity* at the end of each time interval. The method presented here is based on a mildly nonlinear complementarity problem (NCP) formulation of time-stepping. The NCP's are solved using a sequence of linear complementarity problems (LCP's) which are solved using Lemke's algorithm (see Cottle, Pang and Stone[7], Algo. 4.4.1). The direct use of complementarity principles and LCP's follows Trinkle, Pang, Sudarsky and Lo[44]. However, Trinkle *et al.*[44] takes the "instantaneous" point of view, and uses LCP's to compute forces and accelerations. The instantaneous formulation does not always have solutions. By contrast, the time-stepping formulation developed here always does.

The time-stepping version developed here uses a polyhedral approximation to the friction cone. Using the true (circular or elliptical) friction cone leads to a strongly nonlinear complementarity formulation. While such a nonlinear complementarity formulation may be useful, there is less supporting theory and the subproblems are more difficult to solve. Rather, we use polyhedral approximations to the true cone which can approximate the true friction cones as accurately as desired.

The mathematical formulation for the continuous problem used here is loosely based on that of Moreau [31, 32] and Monteiro-Marques[28] who give a theoretical formulation in terms of *measure differential inclusions*. These generalize *differential inclusions* (see the references [1, 8, 14]) which have already found application in a wide variety of mathematical models including friction problems [1, 5, 15, 38], and for which there are a variety of numerical methods [12, 19, 20, 21, 37, 38, 42, 43].

Differential inclusions are generalizations of ordinary differential equations which are expressed in the form

$$\frac{dx}{dt} \in F(t, x) \tag{1}$$

where $F(t, x)$ is a set-valued function of (t, x). That is, instead of having a formula giving the derivative at a certain point, there is a collection of possible values of dx/dt. The differential inclusions considered above have the following properties:

1. $F(t, x)$ is a compact, convex set for all (t, x).
2. The graph $\{ (x, y) \mid y \in F(t, x) \}$ of $F(t, \cdot)$ is a closed subset of $\mathbf{R}^n \times \mathbf{R}^n$ for all t.
3. The set valued function $F(\cdot, x)$ is a measurable set valued function (as described in Aubin and Frankowska[2]).

These conditions, together with a condition which prevents solutions going to infinity in finite time, guarantee the existence of solutions. The closed graph property described above means that F does not have to be smooth. In particular, differential inclusions can be applied to discontinuous ODE's, $x' = f(t, x)$ where $f(t, \cdot)$ is discontinuous, by replacing the differential equation with the inclusion (1) where $F(t, x) = \bigcap_{\delta > 0} \overline{\text{co}} \{ f(t, y) \mid \|y - x\| < \delta \}$. Note that $\overline{\text{co}} X$ is the closed convex hull of X; that is, $\overline{\text{co}} X$ is the smallest closed and convex set that contains X. Jump discontinuities are replaced by the closed convex hull of the values $f(t, y)$ where y is "close" to x.

Readers interested in the mathematical theory of differential inclusions should consult Aubin and Cellina[1]; readers interested in the behavior of solutions to differential inclusions for well structured problems should consult Filippov[15]; readers interested in numerical solution of differential inclusions should first consult the review article by Dontchev and Lempio[12]. The oldest paper on numerical solution of differential inclusions is by Taubert[42]; recent high accuracy methods can be found in Kastner-Maresch[20, 21], and in Stewart[37, 38].

2. Formulation of general rigid body dynamics problem

The formulations of Moreau[31, 32] and Monteiro-Marques[28] consider a particle of mass m in a given admissible region $C \subset \mathbf{R}^3$ with position $\mathbf{q}(t)$ and velocity $\mathbf{v}(t)$. Here we consider generalized coordinates $\mathbf{q} \in \mathbf{R}^n$, which leads to a more complex set of admissible coordinates C. It is required that $\mathbf{q}(t) \in C$ for all t. In the references [28, 31, 32] it is assumed that C has a smooth boundary and is represented by a single scalar function $f : \mathbf{R}^n \to \mathbf{R}$ by the formula

$$C = \{ \mathbf{q} \mid f(\mathbf{q}) \geq 0 \} \subset \mathbf{R}^n \tag{2}$$

where $\boldsymbol{\nabla} f(\mathbf{q}) \neq 0$ whenever $\mathbf{q}$ lies in the boundary ∂C of C. Initially, this paper will deal with a single contact and the admissible region will be defined by a single scalar function as in (2). Later, in the sections on multiple contacts (§3.4), the formulation will be extended to regions defined by a vector function $\mathbf{f} : \mathbf{R}^n \to \mathbf{R}^p$ and

$$C = \{ \mathbf{q} \mid f_i(\mathbf{q}) \geq 0, \ i = 1, \ldots, p \} \tag{3}$$

The velocity $\mathbf{v}(\cdot)$ is assumed to be a function with locally *bounded variation* (see [40]) and so is bounded, but may be discontinuous, while the position is a continuous, locally Lipschitz function satisfying

$$\mathbf{q}(t) = \mathbf{q}(0) + \int_0^t \mathbf{v}(\tau) \, d\tau. \tag{4}$$

In generalized coordinates the kinetic energy is assumed to be given by

$$T(\mathbf{q}, \mathbf{v}) = \frac{1}{2}\mathbf{v}^T \mathbf{M}(\mathbf{q})\mathbf{v} \tag{5}$$

where $\mathbf{M}(\mathbf{q})$ is the generalized mass matrix, and the potential energy is a given function $V(\mathbf{q})$. Note that $\mathbf{M}(\mathbf{q})$ may contain moments of inertia if the generalized coordinates include orientation parameters. The Lagrangian is then

$$L(\mathbf{q}, \mathbf{v}) = T(\mathbf{q}, \mathbf{v}) - V(\mathbf{q}) = \frac{1}{2}\mathbf{v}^T \mathbf{M}(\mathbf{q})\mathbf{v} - V(\mathbf{q}). \tag{6}$$

The equations of motion without unilateral constraints are

$$\frac{d}{dt}\frac{\partial L}{\partial \mathbf{v}} - \frac{\partial L}{\partial \mathbf{q}} = 0, \qquad \frac{d\mathbf{q}}{dt} = \mathbf{v}. \tag{7}$$

For frictionless bilateral constraints $f(\mathbf{q}) = 0$, there is a simple extension of these equations which incorporate contact forces in terms of Lagrange multipliers for the constraint[24]:

$$\frac{d}{dt}\frac{\partial L}{\partial \mathbf{v}} - \frac{\partial L}{\partial \mathbf{q}} = \psi(t)\boldsymbol{\nabla} f(\mathbf{q}(t)), \qquad \frac{d\mathbf{q}}{dt} = \mathbf{v}. \tag{8}$$

For a *unilateral* constraint $f(\mathbf{q}) \geq 0$, if there is no contact ($f(\mathbf{q}(t)) > 0$), then the contact forces must be zero ($\psi(t) = 0$). Also, since the contact is not adhesive, the contact force cannot be directed outwards, so $\psi(t) \geq 0$. Admissibility of the solution implies that $f(\mathbf{q}(t)) \geq 0$ for all t. This gives the corresponding conditions for unilateral contact:

$$\begin{aligned}
\frac{d}{dt}\frac{\partial L}{\partial \mathbf{v}} - \frac{\partial L}{\partial \mathbf{q}} &= \psi(t)\boldsymbol{\nabla} f(\mathbf{q}(t)), \qquad \frac{d\mathbf{q}}{dt} = \mathbf{v}, \\
\psi(t), f(\mathbf{q}(t)) &\geq 0 \qquad \text{for all } t, \\
\psi(t)\, f(\mathbf{q}(t)) &= 0 \qquad \text{for all } t.
\end{aligned} \tag{9}$$

Equation (9) is a differential version of a Nonlinear Complementarity Problem (NCP) which have the form: Given $\mathbf{g}\colon \mathbf{R}^n \to \mathbf{R}^n$, find $\mathbf{z}$ such that

$$\mathbf{z} \geq 0, \qquad \mathbf{g}(\mathbf{z}) \geq 0, \qquad \mathbf{z}^T \mathbf{g}(\mathbf{z}) = 0.$$

The differential equation in (9) can be restated as

$$\begin{aligned}
\mathbf{M}(\mathbf{q})\frac{d\mathbf{v}}{dt} &= \mathbf{k}(\mathbf{q}, \mathbf{v}) + \psi \boldsymbol{\nabla} f(\mathbf{q}), \\
k_l(\mathbf{q}, \mathbf{v}) &= \frac{1}{2}\sum_{i,j}\left[\frac{\partial m_{li}}{\partial q_j} + \frac{\partial m_{lj}}{\partial q_i} - \frac{\partial m_{ij}}{\partial q_l}\right]v_i v_j - \frac{\partial V}{\partial q_l}.
\end{aligned} \tag{10}$$

In general, this system does not have solutions involving impacts unless $\psi(\cdot)$ is allowed to be a *distribution*[13, 18] or a *vector measure*[6, 11], which corresponds to allowing instantaneous impulses. Using a separate formulation for impulsive forces leads to a number of difficulties. To illustrate the difficulties of separate formulations, note that it is possible for a system to undergo an infinite number of impulses in a finite time. Indeed, this is the case for a ball bouncing on a horizontal table with a coefficient of restitution strictly between zero and one. By requiring a separate formulation at impulses, we arrive at a version of Zeno's paradox: the system cannot be solved beyond the (finite) time taken for the ball to come to rest, as an infinite number of impulsive formulations have to be solved before then. Also, these situations do not handle the case where forces are instantaneously unbounded,

but are not impulsive. These include, for example, forces having the form $f(t) = 1/\sqrt{|t|}$. Since impulsive forces arise naturally in rigid body mechanics, and then not only in the context of collisions, it appears natural to expect that such intermediate forms as these may occur in certain rigid body problems, which are not handled by either a "finite force" or "impulsive" formulations.

The approach taken here is to construct a time-stepping method. Time-stepping methods avoid these versions of Zeno's paradox by using integrals of the applied forces to compute integrals of the contact forces over small time intervals. While the bouncing ball problem does not arise with inelastic impacts, other problems in which an infinite number of impacts occur in a finite time will not prevent the time-stepping algorithm presented here from obtaining solutions.

2.1. Coulomb friction. Here Coulomb friction forces are defined in terms of a *friction cone* $FC(\mathbf{q})$ which contains the sums of the normal and frictional contact forces. This is conventionally given for a particle at a point $\mathbf{q}$ on the boundary of the admissible region $C \subset \mathbf{R}^3$ as the set

$$FC(\mathbf{q}) = \{\, \mathbf{c}_t + c_n \mathbf{n} \mid \|\mathbf{c}_t\|_2 \leq \mu c_n, \, \mathbf{c}_t \perp \mathbf{n} \,\}$$

where $\mathbf{n}$ is the normal vector of C at $\mathbf{q}$ pointing into the admissible region, $\mathbf{c}_t$ is the friction force, and $\mu \geq 0$ is the coefficient of friction. If $\mathbf{q}$ is in the interior of C, then $FC(\mathbf{q}) = \{0\}$. This simple representation does not hold for generalized coordinates, but a similar form can be used. Note that in generalized coordinates vectors must be transformed from "real" or "relative" coordinates to generalized force coordinates by a linear transformation. Thus in generalized coordinates, it can no longer be guaranteed that $\mathbf{n}$ has unit length, or that $\mathbf{c}_t$ is perpendicular to $\mathbf{n}$. Also, more general friction cones can be devised for dealing with more complex situations, such as soft finger models[17]. Further, in multiple contact situations, each contact has its own friction cone.

If $\boldsymbol{\rho}(t)$ are the total contact forces, consisting of the normal contact force $\psi \boldsymbol{\nabla} f$ plus the generalized friction force, then the basic formulation of Moreau and Monteiro-Marques states that

$$
\begin{aligned}
\mathbf{M}(\mathbf{q}) \frac{d\mathbf{v}}{dt} &= \mathbf{k}(\mathbf{q}, \mathbf{v}) + \boldsymbol{\rho}(t), \\
f(\mathbf{q}(t)) &\geq 0, \qquad \text{for all } t, \\
\boldsymbol{\rho}(t) &\in FC(\mathbf{q}(t)), \qquad \text{for all } t, \\
\boldsymbol{\rho}(t) \, f(\mathbf{q}(t)) &= 0, \qquad \text{for all } t.
\end{aligned}
\tag{11}
$$

The choice of $\boldsymbol{\rho}(t)$ from the friction cone is still not specified. Note that this is not surprising as the choice of impact models (inelastic, perfectly elastic, partially elastic) has not been considered so far. For inelastic impact models, Moreau[31, 32] and Monteiro-Marques[28] invoke a *maximal dissipation principle*, which states that the this choice must be made at each instant so as to maximize the rate of loss of the total energy, subject only to the rigid body and friction cone constraints. Note that subjecting the choice of $\boldsymbol{\rho}(t)$ to the rigid body constraints when imposing the maximum dissipation condition avoids the common claim that solutions to Coulomb friction problems "may not exist"[9, 10, 25, 27, 33].

In this paper, the explicit use of the maximal dissipation principle will not be further explored. Instead, an alternative way of guaranteeing physically meaningful solutions by means of a complementarity formulation will be developed.

3. Complementarity formulation of time-stepping algorithm

The following time-stepping formulation is a variant of the well-known implicit Euler method for ODE's. For now, we consider a problem with only one contact; a more general method for dealing with multiple contacts will be presented in §3.4. For a given time step, values for position $\mathbf{q}^l$ and velocity $\mathbf{v}^l$ are used to compute new values of $\mathbf{q}^{l+1}$ and $\mathbf{v}^{l+1}$. To start with, it will be assumed that in the neighborhood the boundary of the admissible region is well approximated by a half-space: $\mathbf{n}^T\mathbf{q} \geq \alpha_0$. This assumption will be relaxed later.

To make linear complementarity theory[7] applicable, the friction cone will be approximated by a polyhedral cone:

$$\widehat{FC}(\mathbf{q}) = \{\, c_n\mathbf{n} + \mathbf{D}\boldsymbol{\beta} \mid c_n \geq 0, \boldsymbol{\beta} \geq 0, \mathbf{e}^T\boldsymbol{\beta} \leq \mu c_n \,\} \tag{12}$$

where $e = [1, 1, \ldots, 1]^T$ in $\mathbf{R}^k$ and k is the number of edges of the polyhedral approximation. The columns of $\mathbf{D}$ are direction vectors $\mathbf{d}_j$, although these need not be of unit length for generalized coordinates, or for anisotropic friction laws. The vector $\boldsymbol{\beta}$ will be used below as a weighting vector for the direction vectors $\mathbf{d}_i$. It is assumed that for every i there is a j such that

$$\mathbf{d}_i = -\mathbf{d}_j. \tag{13}$$

It is also assumed that the vectors $\mathbf{d}_j$ span the subspace on which the friction forces act; thus if $\mathbf{v}$ is a vector in that subspace, then $\mathbf{d}_j^T\mathbf{v} \geq 0$ for all j implies $\mathbf{v} = 0$. In the case of the motion of a particle, the span of $\{\, \mathbf{d}_j \mid j = 1, \ldots k \,\}$ is the tangent plane to ∂C at $\mathbf{q}$. Figure 1 shows an eight-sided approximation to a circular friction cone. While this means that several different weighting vectors $\boldsymbol{\beta}$ could give the same friction vector $\mathbf{D}\boldsymbol{\beta}$, the complementarity conditions below will force the $\boldsymbol{\beta}$ vector to be unique provided $\mathbf{D}^T\mathbf{v}^{l+1} \neq 0$.

The matrix $\mathbf{D}$ for bodies moving in two dimensions can be easily constructed. If the relative velocity at the contact is given by $\boldsymbol{\gamma}^T(d\mathbf{q}/dt)$, then $\mathbf{D}$ is the $n \times 2$ matrix with columns $\pm\boldsymbol{\gamma}$. (Note that n is the dimension of the generalized coordinates vector $\mathbf{q}$.) This is just the transformation of a common unit tangent vector at contact, into the generalized velocity co-ordinates. Similarly, for two three-dimensional bodies in contact, the circular friction cone can be approximated by a polyhedral cone. For example, consider two balls in contact. There is an untransformed normal vector $\overline{\mathbf{n}} \in \mathbf{R}^3$ which is perpendicular to the tangent plane in $\mathbf{R}^3$. To obtain a vector perpendicular to $\overline{\mathbf{n}}$ in $\mathbf{R}^3$, take a cross product of $\overline{\mathbf{n}}$ with the unit vertical vector in $\mathbf{R}^3$, and normalize the result ($\overline{\mathbf{t}}_1$). To compute another vector in the tangent plane in $\mathbf{R}^3$, take the cross product $\overline{\mathbf{t}}_2 = \overline{\mathbf{t}}_1 \times \overline{\mathbf{n}} \in \mathbf{R}^3$. For an k-gon approximation, construct the vectors $\overline{\mathbf{t}}(\theta_i) = \cos\theta_i\overline{\mathbf{t}}_1 + \sin\theta_i\overline{\mathbf{t}}_2$ for $\theta_i = 2\pi i/k$, $i = 0, 1, 2, \ldots, k-1$. These vectors are then transformed $\overline{\mathbf{t}}(\theta_i)$ into generalized velocity coordinates $\mathbf{d}_i$. The details of the transformation into generalized velocity coordinates depends on the way information is stored in the generalized coordinate vector.
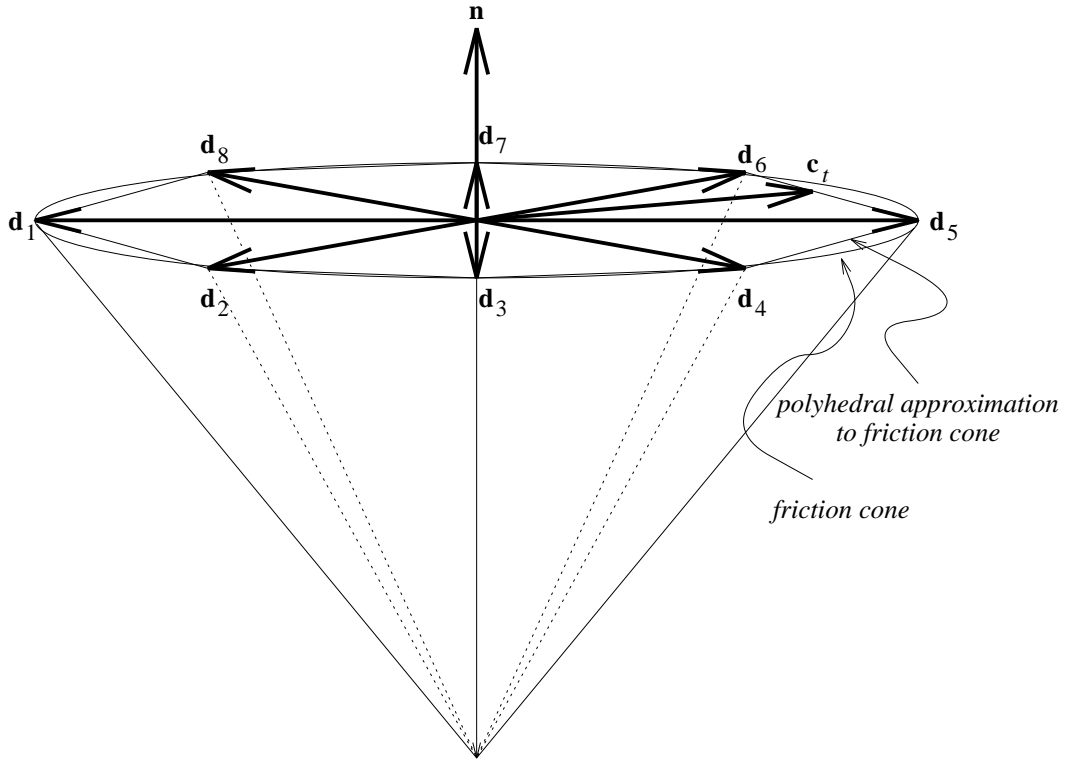
FIGURE 1. Circular friction cone and polyhedral approximation

The formulation of the time-stepping is then to compute $\mathbf{q}^{l+1}$, $\mathbf{v}^{l+1}$ and the associated variables c_n, $\boldsymbol{\beta}$ and λ satisfying

$$\mathbf{M}(\mathbf{q}^l + h\mathbf{v}^l) \cdot (\mathbf{v}^{l+1} - \mathbf{v}^l) = \mathbf{n}c_n + \mathbf{D}\boldsymbol{\beta} + h \, \mathbf{k}(\mathbf{q}^l + h\mathbf{v}^l/2, \mathbf{v}^l),$$

$$\mathbf{q}^{l+1} - \mathbf{q}^l = h\mathbf{v}^{l+1},$$

$$\lambda\mathbf{e} + \mathbf{D}^T\mathbf{v}^{l+1} \geq 0, \qquad\qquad \boldsymbol{\beta} \geq 0, \qquad\qquad (14)$$

$$\mathbf{n}^T\mathbf{q}^{l+1} \geq \alpha_0, \qquad\qquad c_n \geq 0,$$

$$\mu c_n - \mathbf{e}^T\boldsymbol{\beta} \geq 0, \qquad\qquad \lambda \geq 0,$$

with the complementarity conditions

$$[\lambda\mathbf{e} + \mathbf{D}^T\mathbf{v}^{l+1}]^T\boldsymbol{\beta} = 0,$$

$$[\mathbf{n}^T\mathbf{q}^{l+1} - \alpha_0]c_n = 0, \qquad\qquad (15)$$

$$[\mu c_n - \mathbf{e}^T\boldsymbol{\beta}]\lambda = 0.$$

The additional variable λ is in most cases an approximation to the magnitude of the relative contact velocity. In some situations, such as where there is zero relative contact velocity and the friction vector is zero, λ can take any non-negative value, and it has no physical meaning at all. However, it is necessary in this formulation to ensure that the correct components of $\boldsymbol{\beta}$ are non-zero, and that the friction vector

$\mathbf{D}\boldsymbol{\beta}$ is on the boundary of the approximate friction cone if the relative contact velocity is non-zero.

The middle inequality of (14), $\mathbf{n}^T\mathbf{q}^{l+1} \geq \alpha_0$ is a guarantee that the (linearized) rigid body constraints are satisfied. Suppose now that $\mathbf{v}^{l+1}$ is not perpendicular to the friction plane. The first inequality then implies that $\lambda \geq \max_i -\mathbf{d}_i^T\mathbf{v}^{l+1} > 0$ by (13). In this case, complementarity implies that $\mu c_n = \mathbf{e}^T\boldsymbol{\beta}$. This then means that β_i can only be positive when $\lambda + \mathbf{d}_i^T\mathbf{v}^{l+1} = 0$, or equivalently, that

$$\mathbf{d}_i^T\mathbf{v}^{l+1} = \min_j \mathbf{d}_j^T\mathbf{v}^{l+1}. \tag{16}$$

For non-zero $\mathbf{D}^T\mathbf{v}^{l+1}$ there will usually be a unique minimizing j in (16), giving exactly one non-zero component of $\boldsymbol{\beta}$, and the friction vector is along one of the $\mathbf{d}_i$ direction vectors. There may be cases where the minimizer is not unique, in which case the friction force $\mathbf{D}\boldsymbol{\beta}$ lies in a face of the approximate friction cone, and thus, there is a unique $\boldsymbol{\beta}$ vector. In the case of ordinary contact in three dimensions, the $\mathbf{d}_i$ vectors belong to a two-dimensional plane. In this case, at most two adjacent β_i components are non-zero, and the friction vector uniquely specifies $\boldsymbol{\beta}$. If $\mathbf{D}^T\mathbf{v}^{l+1} = 0$, then the $\boldsymbol{\beta}$ vector is not uniquely specified, but the vector $\mathbf{D}\boldsymbol{\beta}$ is specified through the fact that $\mathbf{D}^T\mathbf{v}^{l+1} = 0$.

Using the above results, the dissipation due to the friction forces in a single step is

$$\mathbf{v}^{l+1\ T}\mathbf{D}\boldsymbol{\beta} = -\boldsymbol{\beta}^T\mathbf{e}\lambda = -\mu\lambda c_n.$$

However,

$$\lambda = \max_i -\mathbf{d}_i^T\mathbf{v}^{l+1} \geq -\sum_i \beta_i'\mathbf{d}_i^T\mathbf{v}^{l+1}$$

for any $\boldsymbol{\beta}'$ satisfying $\beta_j' \geq 0$ for all j and $\sum_j' \beta_j = 1$. Thus the friction impulse $\mathbf{D}\boldsymbol{\beta}$ that satisfies the above complementarity problem maximizes the dissipation over all frictional contact impulses, given the normal reaction impulse c_n. Thus for isotropic friction laws and sufficiently many well-spaced $\mathbf{d}_i$ vectors, the friction impulse is close to being directly opposite to the component of the sliding velocity $\mathbf{v}^{l+1}$ in the friction plane.

Several aspects of (14–15) should also be noted.

1. It is a mixed linear complementarity system as described in Cottle, Pang and Stone[7], pp. 29–30. By solving for $\mathbf{v}^{l+1}$ and $\mathbf{q}^{l+1}$ in terms of the other quantities, a "pure" LCP will be obtained.

2. The normal contact impulse c_n, and thus $\boldsymbol{\beta}$, is only nonzero if there is contact at at the end of the interval (i.e. $\mathbf{n}^T\mathbf{q}^{l+1} = \alpha_0$). Thus there is no need to explicitly "turn on" and "turn off" the contact forces in this method.

3. The approximate Coulomb law may be violated (i.e. $\mathbf{e}^T\boldsymbol{\beta} \neq \mu c_n$) only if $\lambda = 0$, which implies that $\mathbf{D}^T\mathbf{v}^{l+1} = 0$. This corresponds to the physical situation where sliding stops during the contact period.

4. The values of the quantities $\mathbf{M}$, $\mathbf{n}$ and $\mathbf{D}$ should be obtained at some value of $\mathbf{q}$; the simplest approach is to use $\mathbf{q} = \mathbf{q}^l + h\mathbf{v}^l$. This corresponds to the explicit Euler prediction. More complex approaches can use more realistic approximations such as $\mathbf{q} = \mathbf{q}^{l+1}$ or $\mathbf{q} = (\mathbf{q}^l + \mathbf{q}^{l+1})/2$. However, these methods result in a nonlinear complementarity problem. While they may be expected to give higher accuracy, they are more complex to implement and to prove the existence of solutions.

3.1. Schur complements & LCP's.
As given, (14–15) is a mixed linear complementarity problem[7]. In order to prove existence of solutions to this problem, note that by solving for $\mathbf{q}^{l+1}$ and $\mathbf{v}^{l+1}$ in terms of the other quantities gives a Schur complement system. Let $\mathbf{k} = \mathbf{k}(\mathbf{q}^l + h\mathbf{v}^l/2)$ and $\mathbf{M} = \mathbf{M}(\mathbf{q}^l + h\mathbf{v}^l)$. The Schur complement system is:

$$\begin{bmatrix} \mathbf{D}^T\mathbf{M}^{-1}\mathbf{D} & \mathbf{D}^T\mathbf{M}^{-1}\mathbf{n} & \mathbf{e} \\ \mathbf{n}^T\mathbf{M}^{-1}\mathbf{D} & \mathbf{n}^T\mathbf{M}^{-1}\mathbf{n} & 0 \\ -\mathbf{e}^T & \mu & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ c_n \\ \lambda \end{bmatrix} + \mathbf{b} = \begin{bmatrix} \boldsymbol{\rho} \\ \sigma \\ \zeta \end{bmatrix} \geq 0,$$

$$\text{where} \qquad \mathbf{b} = \begin{bmatrix} \mathbf{D}^T(\mathbf{v}^l + h\mathbf{M}^{-1}\mathbf{k}) \\ (\mathbf{n}^T\mathbf{q}^l - \alpha_0)/h + \mathbf{n}^T(\mathbf{v}^l + h\mathbf{M}^{-1}\mathbf{k}) \\ 0 \end{bmatrix}, \qquad (17)$$

$$\text{and} \qquad \begin{bmatrix} \boldsymbol{\beta} \\ c_n \\ \lambda \end{bmatrix} \geq 0, \qquad \begin{bmatrix} \boldsymbol{\beta} \\ c_n \\ \lambda \end{bmatrix}^T \begin{bmatrix} \boldsymbol{\rho} \\ \sigma \\ \zeta \end{bmatrix} = 0.$$

3.2. Existence of solutions.
The LCP (17) has solutions, as will be shown here. The matrix

$$\mathbf{N} = \begin{bmatrix} \mathbf{D}^T\mathbf{M}^{-1}\mathbf{D} & \mathbf{D}^T\mathbf{M}^{-1}\mathbf{n} & \mathbf{e} \\ \mathbf{n}^T\mathbf{M}^{-1}\mathbf{D} & \mathbf{n}^T\mathbf{M}^{-1}\mathbf{n} & 0 \\ -\mathbf{e}^T & \mu & 0 \end{bmatrix}$$

is *copositive* (see Cottle, Pang and Stone[7], Defn. 3.8.1), since for any vector $\mathbf{z} = \begin{bmatrix} \boldsymbol{\beta}^T & c_n & \lambda \end{bmatrix}^T \geq 0$,

$$\begin{bmatrix} \boldsymbol{\beta} \\ c_n \\ \lambda \end{bmatrix}^T \begin{bmatrix} \mathbf{D}^T\mathbf{M}^{-1}\mathbf{D} & \mathbf{D}^T\mathbf{M}^{-1}\mathbf{n} & \mathbf{e} \\ \mathbf{n}^T\mathbf{M}^{-1}\mathbf{D} & \mathbf{n}^T\mathbf{M}^{-1}\mathbf{n} & 0 \\ -\mathbf{e}^T & \mu & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ c_n \\ \lambda \end{bmatrix}$$

$$= (c_n\mathbf{n} + \mathbf{D}\boldsymbol{\beta})^T\mathbf{M}^{-1}(c_n\mathbf{n} + \mathbf{D}\boldsymbol{\beta}) + \mu c_n\lambda \geq 0 \quad (18)$$

as $\mathbf{M}^{-1}$ is positive definite and $\mu \geq 0$. Note, however, that $\mathbf{N}$ is not *copositive plus* (Cottle, Pang and Stone[7], Def. 3.8.1), since $\mathbf{z}^T\mathbf{N}\mathbf{z} = 0$ implies that $\mathbf{D}\boldsymbol{\beta} = 0$, $c_n = 0$, but not $\lambda = 0$, and $(\mathbf{N} + \mathbf{N}^T)\mathbf{z}$ can still be non-zero if $\lambda > 0$. Nevertheless, solutions do exist (by Cottle, Pang and Stone[7], Thm. 4.4.12) since the only solutions of LCP($\mathbf{N}, 0$) have the form $\mathbf{z} = \begin{bmatrix} 0 & 0 & \lambda \end{bmatrix}$, and

$$\mathbf{z}^T\mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ \lambda \end{bmatrix}^T \begin{bmatrix} \mathbf{D}^T(\mathbf{v}^l + h\mathbf{M}^{-1}\mathbf{k}) \\ (\mathbf{n}^T\mathbf{q}^l - \alpha_0)/h + \mathbf{n}^T(\mathbf{v}^l + h\mathbf{n}^T\mathbf{M}^{-1}\mathbf{k}) \\ 0 \end{bmatrix} = 0 \not< 0. \qquad (19)$$

Thus (by Cottle, Pang and Stone[7], Thm. 4.4.12) not only do solutions exist, but also Lemke's algorithm (described in Cottle, Pang and Stone[7], Algo. 4.4.1) can compute a solution, provided precautions are taken against cycling due to degeneracy. (See Cottle, Pang and Stone[7], §4.9.) While solutions are guaranteed to exist, uniqueness is not guaranteed, although it is expected for most problems.

3.3. Dissipativity of the method.
A limited dissipativity result holds for this method where the system is linearized with $\mathbf{M}$ and $\mathbf{k}$ assumed constant, and the admissible region (the set of allowable positions) is a half-plane. In the general

nonlinear case, numerical methods cannot usually be guaranteed to be dissipative. Consider the following method for a particle moving in a potential field $V(\mathbf{q})$:

$$m(\mathbf{v}^{l+1} - \mathbf{v}^l) = -h\boldsymbol{\nabla}V(\mathbf{q}^{l+1}), \qquad \mathbf{q}^{l+1} = \mathbf{q}^l + \frac{h}{2}(\mathbf{v}^l + \mathbf{v}^{l+1}). \tag{20}$$

This method is dissipative if V is convex, but anti-dissipative if V is concave. For general V the method may or may not be dissipative depending on the step taken. Even symplectic methods[36] for smooth Hamiltonian systems do not exactly conserve the true energy, but can only conserve an approximate Hamiltonian[36].

Dissipativity of (14,15) will now be shown under the assumptions that $\mathbf{M}$, $\mathbf{k}$ are constant and the admissible region is a half-plane. Note that for non-constant $\mathbf{M}$, $\mathbf{k}$ and $\mathbf{n}$ this method is not guaranteed to be dissipative. Indeed, if there are no contacts, constant $\mathbf{M}$ and $\mathbf{n}$, but non-constant $\mathbf{k}$ arising from a conservative system, the method (14,15) reduces to an explicit Euler method for a conservative system. As noted above, such methods are not guaranteed to be exactly dissipative. However, a later paper, which deals with the convergence theory of this method, will show that the *limit* of the numerical solutions is dissipative.

From the linearization of (14,15), note that

$$(\mathbf{v}^{l+1})^T\mathbf{M}(\mathbf{v}^{l+1} - \mathbf{v}^l)$$
$$= \frac{1}{2}((\mathbf{v}^{l+1} + \mathbf{v}^l) + (\mathbf{v}^{l+1} - \mathbf{v}^l))^T\mathbf{M}(\mathbf{v}^{l+1} - \mathbf{v}^l)$$
$$= \frac{1}{2}(\mathbf{v}^{l+1} + \mathbf{v}^l)^T\mathbf{M}(\mathbf{v}^{l+1} - \mathbf{v}^l) + \frac{1}{2}(\mathbf{v}^{l+1} - \mathbf{v}^l)^T\mathbf{M}(\mathbf{v}^{l+1} - \mathbf{v}^l) \tag{21}$$
$$= \frac{1}{2}[(\mathbf{v}^{l+1})^T\mathbf{M}\mathbf{v}^{l+1} - \mathbf{v}^{l\,T}\mathbf{M}\mathbf{v}^l] + \frac{1}{2}(\mathbf{v}^{l+1} - \mathbf{v}^l)^T\mathbf{M}(\mathbf{v}^{l+1} - \mathbf{v}^l)$$

while

$$(\mathbf{v}^{l+1})^T\mathbf{M}(\mathbf{v}^{l+1} - \mathbf{v}^l) = (\mathbf{v}^{l+1})^T(c_n\mathbf{n} + \mathbf{D}\boldsymbol{\beta} + h\mathbf{k})$$
$$= (1/h)(\mathbf{q}^{l+1} - \mathbf{q}^l)^Tc_n\mathbf{n} + (\mathbf{v}^{l+1})^T\mathbf{D}\boldsymbol{\beta} + (\mathbf{q}^{l+1} - \mathbf{q}^l)^T\mathbf{k}. \tag{22}$$

Since $\mathbf{n}^T\mathbf{q}^l \geq \alpha_0$, it follows that $(\mathbf{q}^{l+1} - \mathbf{q}^l)^T\mathbf{n}c_n \leq ((\mathbf{q}^{l+1})^T\mathbf{n} - \alpha_0)c_n = 0$ by complementarity. Thus

$$(\mathbf{v}^{l+1})^T\mathbf{M}(\mathbf{v}^{l+1} - \mathbf{v}^l) \leq \boldsymbol{\beta}^T\mathbf{D}^T\mathbf{v}^{l+1} + (\mathbf{q}^{l+1} - \mathbf{q}^l)^T\mathbf{k}$$
$$= -\lambda\boldsymbol{\beta}^T\mathbf{e} + (\mathbf{q}^{l+1} - \mathbf{q}^l)^T\mathbf{k} = -\lambda\mu c_n + (\mathbf{q}^{l+1} - \mathbf{q}^l)^T\mathbf{k}. \tag{23}$$

This gives

$$\frac{1}{2}(\mathbf{v}^{l+1})^T\mathbf{M}\mathbf{v}^{l+1} + \mathbf{k}^T\mathbf{q}^{l+1} \leq \frac{1}{2}\mathbf{v}^{l\,T}\mathbf{M}\mathbf{v}^l + \mathbf{k}^T\mathbf{q}^l - \lambda\mu c_n - \frac{1}{2}(\mathbf{v}^{l+1} - \mathbf{v}^l)^T\mathbf{M}(\mathbf{v}^{l+1} - \mathbf{v}^l)$$
$$\leq \frac{1}{2}\mathbf{v}^{l\,T}\mathbf{M}\mathbf{v}^l + \mathbf{k}^T\mathbf{q}^l \tag{24}$$

and the method is dissipative. The dominant term during a period of sliding along a boundary is $-\lambda\mu c_n$. For circularly isotropic friction applied to a particle, λ is approximately the magnitude of the relative velocity of the particle, and μc_n is the frictional impulse applied to the particle over the time step. This corresponds to the physically correct rate of dissipation of energy. The quadratic term $-(\mathbf{v}^{l+1} - \mathbf{v}^l)^T\mathbf{M}(\mathbf{v}^{l+1} - \mathbf{v}^l)$ has different character in different parts of the motion. During non-contact periods, it is numerical dissipation. That is, during non-contact periods, this energy loss is due solely to the numerical scheme, and does not correspond to any physical process. Note that in this case, $\|\mathbf{v}^{l+1} - \mathbf{v}^l\|$ is of the same

order as the step size h, and so the energy dissipation is of order h^2 for one time step, and over a given finite interval is of order h. In an impact, it gives (approximately) the energy loss due to the inelastic collision.

3.4. **Multiple contacts.** Multiple contacts can be easily incorporated into this framework. The additional data needed to form the LCP for a given time step consist of the following quantities for each contact. The vector $\mathbf{n}^{(j)}$, is the relative normal vector for the jth contact, transformed to generalized coordinates; and $\alpha_0^{(j)}$ the corresponding scalar for locating the boundary of the jth half-plane. The matrix $\mathbf{D}^{(j)}$, is formed from the vectors in the plane of relative motion for the jth contact, whose convex hull approximates the friction cone, transformed to generalized coordinates. The scalar $\mu^{(j)} \geq 0$, is the coefficient of friction for the jth contact.

There are also new variables for each contact: $c_n^{(j)}$ (the normal contact impulse for contact j), $\beta^{(j)}$ (the coefficients for the frictional impulse for contact j), and $\lambda^{(j)}$. With these data and variables the formulation of the time-stepping method follows. Note that p is the number of contacts.

$$\mathbf{M}(\mathbf{q}^l + h\mathbf{v}^l) \cdot (\mathbf{v}^{l+1} - \mathbf{v}^l) = \sum_{j=1}^{p}(\mathbf{n}^{(j)}c_n^{(j)} + \mathbf{D}^{(j)}\beta^{(j)}) + h\,\mathbf{k},(\mathbf{q}^l + h\mathbf{v}^l/2, \mathbf{v}^l)$$

$$\mathbf{q}^{l+1} - \mathbf{q}^l = h\mathbf{v}^{l+1},$$

$$\begin{aligned}
\lambda^{(j)}\mathbf{e}^{(j)} + \mathbf{D}^{(j)\,T}\mathbf{v}^{l+1} \geq 0, && \beta^{(j)} \geq 0, && j = 1, \ldots, p, \\
\mathbf{n}^{(j)\,T}\mathbf{q}^{l+1} \geq \alpha_0^{(j)}, && c_n^{(j)} \geq 0, && j = 1, \ldots, p, \\
\mu^{(j)}c_n^{(j)} - \mathbf{e}^{(j)\,T}\beta^{(j)} \geq 0, && \lambda^{(j)} \geq 0, && j = 1, \ldots, p,
\end{aligned} \tag{25}$$

with the complementarity conditions

$$\begin{aligned}
[\lambda^{(j)}e^{(j)} + \mathbf{D}^{(j)\,T}\mathbf{v}^{l+1}]^T\beta^{(j)} = 0, && j = 1, \ldots, p, \\
[\mathbf{n}^{(j)\,T}\mathbf{q}^{l+1} - \alpha_0^{(j)}]c_n^{(j)} = 0, && j = 1, \ldots, p, \\
[\mu^{(j)}c_n^{(j)} - e^{(j)\,T}\beta^{(j)}]\lambda^{(j)} = 0, && j = 1, \ldots, p.
\end{aligned} \tag{26}$$

Note that $e^{(j)}$ is the column vector of ones of the appropriate size. Writing $\widetilde{\mathbf{n}} = [\mathbf{n}^{(1)}, \ldots, \mathbf{n}^{(p)}]$, $\widetilde{\mathbf{c}}_n = [c_n^{(1)}, \ldots, c_n^{(p)}]^T$, $\widetilde{\beta} = [\beta^{(1)\,T}, \ldots, \beta^{(p)\,T}]^T$, $\widetilde{\mathbf{D}} = [\mathbf{D}^{(1)}, \ldots, \mathbf{D}^{(p)}]$, $\widetilde{\alpha}_0 = [\alpha_0^{(1)}, \ldots, \alpha_0^{(p)}]$, $\widetilde{\mu} = \mathrm{diag}(\mu^{(1)}, \ldots, \mu^{(p)})$, $\widetilde{\lambda} = [\lambda^{(1)}, \ldots, \lambda^{(p)}]^T$, and $\widetilde{\mathbf{E}} = \mathrm{diag}(\mathbf{e}^{(1)}, \ldots, \mathbf{e}^{(p)})$, (25–26) can be written as

$$\mathbf{M}(\mathbf{q}^l + h\mathbf{v}^l) \cdot (\mathbf{v}^{l+1} - \mathbf{v}^l) = \widetilde{\mathbf{n}}\widetilde{\mathbf{c}}_n + \widetilde{\mathbf{D}}\widetilde{\beta} + h\,\mathbf{k}(\mathbf{q}^l + h\mathbf{v}^l/2, \mathbf{v}^l),$$

$$\mathbf{q}^{l+1} - \mathbf{q}^l = h\mathbf{v}^{l+1},$$

$$\begin{aligned}
\widetilde{\mathbf{E}}\widetilde{\lambda} + \widetilde{\mathbf{D}}^T\mathbf{v}^{l+1} \geq 0, && \widetilde{\beta} \geq 0, \\
\widetilde{\mathbf{n}}^T\mathbf{q}^{l+1} \geq \widetilde{\alpha}_0, && \widetilde{\mathbf{c}}_n \geq 0, \\
\widetilde{\mu}\widetilde{\mathbf{c}}_n - \widetilde{\mathbf{E}}^T\widetilde{\beta} \geq 0, && \widetilde{\lambda} \geq 0,
\end{aligned} \tag{27}$$

with the complementarity conditions

$$[\widetilde{\mathbf{E}}\widetilde{\boldsymbol{\lambda}} + \widetilde{\mathbf{D}}^T\mathbf{v}^{l+1}]^T\widetilde{\boldsymbol{\beta}} = 0,$$
$$[\widetilde{\mathbf{n}}^T\mathbf{q}^{l+1} - \widetilde{\boldsymbol{\alpha}}_0]^T\widetilde{\mathbf{c}}_n = 0, \tag{28}$$
$$[\widetilde{\mu}\widetilde{\mathbf{c}}_n - \widetilde{\mathbf{E}}^T\widetilde{\boldsymbol{\beta}}]^T\widetilde{\boldsymbol{\lambda}} = 0.$$

The associated LCP is:

$$\begin{bmatrix} \widetilde{\mathbf{D}}^T\mathbf{M}^{-1}\widetilde{\mathbf{D}} & \widetilde{\mathbf{D}}^T\mathbf{M}^{-1}\widetilde{\mathbf{n}} & \widetilde{\mathbf{E}} \\ \widetilde{\mathbf{n}}^T\mathbf{M}^{-1}\widetilde{\mathbf{D}} & \widetilde{\mathbf{n}}^T\mathbf{M}^{-1}\widetilde{\mathbf{n}} & 0 \\ -\widetilde{\mathbf{E}}^T & \widetilde{\mu} & 0 \end{bmatrix} \begin{bmatrix} \widetilde{\boldsymbol{\beta}} \\ \widetilde{\mathbf{c}}_n \\ \widetilde{\boldsymbol{\lambda}} \end{bmatrix} + \mathbf{b} = \begin{bmatrix} \widetilde{\boldsymbol{\rho}} \\ \widetilde{\boldsymbol{\sigma}} \\ \widetilde{\boldsymbol{\zeta}} \end{bmatrix} \geq 0,$$

$$\text{where} \qquad \mathbf{b} = \begin{bmatrix} \widetilde{\mathbf{D}}^T(\mathbf{v}^l + h\mathbf{M}^{-1}\mathbf{k}(\mathbf{q}^l + h\mathbf{v}^l/2)) \\ (\widetilde{\mathbf{n}}^T\mathbf{q}^l - \widetilde{\boldsymbol{\alpha}}_0)/h + \widetilde{\mathbf{n}}^T(\mathbf{v}^l + h\mathbf{M}^{-1}\mathbf{k}) \\ 0 \end{bmatrix}, \tag{29}$$

$$\text{and} \qquad \begin{bmatrix} \widetilde{\boldsymbol{\beta}} \\ \widetilde{\mathbf{c}}_n \\ \widetilde{\boldsymbol{\lambda}} \end{bmatrix} \geq 0, \qquad \begin{bmatrix} \widetilde{\boldsymbol{\beta}} \\ \widetilde{\mathbf{c}}_n \\ \widetilde{\boldsymbol{\lambda}} \end{bmatrix}^T \begin{bmatrix} \widetilde{\boldsymbol{\rho}} \\ \widetilde{\boldsymbol{\sigma}} \\ \widetilde{\boldsymbol{\zeta}} \end{bmatrix} = 0.$$

The arguments of §§3.2–3.3 can be applied to (27–28) and to (29) to show that solutions of (27–28) exist and can be computed by Lemke's algorithm, under the assumption that the $\mathbf{n}^{(i)}$ vectors are linearly independent, and that the only vector in both the span of the columns of $\widetilde{\mathbf{n}}$ and of $\widetilde{\mathbf{D}}$ is the zero vector. Furthermore, the above formulation gives a dissipative method provided $\mathbf{M}$, $\mathbf{k}$, $\widetilde{\mathbf{n}}$, $\widetilde{\mathbf{D}}$ and $\widetilde{\boldsymbol{\alpha}}_0$ are constant.

3.5. Other friction.

In many pieces of machinery, there is friction that is associated with contacts that are commonly not assumed to be unilateral. For example, the friction arising in a gear box is not due to unilateral contact in models unless the details of the contact between the gear teeth and the contact in the bearings is explicitly modeled. In many cases, this is an excessively detailed level of modeling.

In such a situation, the friction forces lie within a given convex subset of the plane, which is again approximated by polyhedral sets that are the convex hulls of the sets $\{\mathbf{d}_i \mid i = 1,\ldots,m\}$ which form the columns of the matrix $\mathbf{D} = [\mathbf{d}_1,\ldots,\mathbf{d}_m]$. Then the frictional impulse is assumed to have the form $\mathbf{D}\boldsymbol{\beta}$ where $\boldsymbol{\beta} \geq 0$, and $\mathbf{e}^T\boldsymbol{\beta} \leq h$. The latter inequality holds because the friction force is bounded as a function of time, since the normal contact forces involved are also bounded. The corresponding time-stepping scheme has the form

$$\mathbf{M} \cdot (\mathbf{v}^{l+1} - \mathbf{v}^l) = \mathbf{D}\boldsymbol{\beta} + h\,\mathbf{k},$$
$$\mathbf{q}^{l+1} - \mathbf{q}^l = h\mathbf{v}^{l+1},$$
$$\lambda\mathbf{e} + \mathbf{D}^T\mathbf{v}^{l+1} \geq 0, \qquad \boldsymbol{\beta} \geq 0, \tag{30}$$
$$-\mathbf{e}^T\boldsymbol{\beta} + h \geq 0, \qquad \lambda \geq 0,$$

with the complementarity conditions

$$[\lambda\mathbf{e} + \mathbf{D}^T\mathbf{v}^{l+1}]^T\boldsymbol{\beta} = 0, \qquad [-\mathbf{e}^T\boldsymbol{\beta} + h]\lambda = 0. \tag{31}$$

The corresponding LCP is

$$\begin{bmatrix} \mathbf{D}^T\mathbf{M}^{-1}\mathbf{D} & \mathbf{e} \\ -\mathbf{e}^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \lambda \end{bmatrix} + \mathbf{b} = \begin{bmatrix} \boldsymbol{\rho} \\ \zeta \end{bmatrix} \geq 0,$$

$$\text{where} \qquad \mathbf{b} = \begin{bmatrix} \mathbf{D}^T(\mathbf{v}^l + h\mathbf{M}^{-1}\mathbf{k}) \\ h \end{bmatrix}, \qquad \begin{bmatrix} \boldsymbol{\beta} \\ \lambda \end{bmatrix} \geq 0, \qquad \begin{bmatrix} \boldsymbol{\beta} \\ \lambda \end{bmatrix}^T \begin{bmatrix} \boldsymbol{\rho} \\ \zeta \end{bmatrix} = 0. \quad (32)$$

This LCP has solutions which are computable by Lemke's algorithm since the matrix $\mathbf{N} = \begin{bmatrix} \mathbf{D}^T\mathbf{M}^{-1}\mathbf{D} & \mathbf{e} \\ -\mathbf{e}^T & 0 \end{bmatrix}$ is copositive plus, and the feasible set for the LCP is non-empty (take $\boldsymbol{\beta} = 0$ and $\lambda \geq 0$ sufficiently large). Alternatively, if $\mathbf{z}$ solves LCP($\mathbf{N}, 0$), then $\mathbf{D}\boldsymbol{\beta} = 0$, and

$$\mathbf{z}^T\mathbf{b} = \boldsymbol{\beta}^T\mathbf{D}^T(\mathbf{v}^l + h\mathbf{M}^{-1}\mathbf{k}) + h\lambda = h\lambda \geq 0,$$

and so, by standard LCP theory (Cottle, Pang and Stone[7], Thm. 4.4.12) Lemke's algorithm applied to (32) terminates at a solution.

A "mixed" method for a combination of unilateral constraints with friction, and friction not associated with unilateral constraints, can be developed based on the above formulations.

3.6. Nonlinear versions. One problem with the method as it is given, is that for real problems the admissible regions are not half-spaces, but are more general sets, usually with smooth boundaries. Thus the vector $\mathbf{n}$, and α_0, and the matrices $\mathbf{D}$, depend crucially on the geometry of the problem, and vary from point to point. One particular consequence of this is the failure of dissipativity for the methods (14–15) and (25–26). This can happen where, due to the variation in $\mathbf{n}(\mathbf{q})$ and $\alpha_0(\mathbf{q})$, while $\mathbf{q}^{l+1}$ may be admissible for the linearization based at $\mathbf{q}^l$, this does not mean that it would be admissible for the linearization at $\mathbf{q}^{l+1}$. Because admissibility for $\mathbf{q}^{l+2}$ for the linearization at $\mathbf{q}^{l+1}$ is required, sufficient impulse to achieve this is applied at the step computing $\mathbf{q}^{l+2}$. This reaction impulse for such a step can be far in excess of that for the real system, and this method has been observed by the author to produce extremely large velocities that are entirely unphysical.

To circumvent this problem, a nonlinear method should be employed. In particular, the complementarity condition

$$\mathbf{n}^T\mathbf{q}^{l+1} - \alpha_0 \geq 0, \qquad c_n \geq 0, \qquad (\mathbf{n}^T\mathbf{q}^{l+1} - \alpha_0)c_n = 0$$

should be replaced by the following nonlinear complementarity condition between

$$f(\mathbf{q}^{l+1}) \geq 0, \qquad c_n \geq 0, \qquad f(\mathbf{q}^{l+1})c_n = 0,$$

and $\mathbf{n} = \mathbf{n}(\mathbf{q}^{l+1}) = \boldsymbol{\nabla}f(\mathbf{q}^{l+1})$.

Such a nonlinear complementarity problem can often be solved by a sequence of linear complementarity problems of the form (14,15,17), or (25,26,29) for multiple contacts. Indeed, a fixed point iteration can be used to solve these nonlinear complementarity problems, as follows. At the first stage, an estimate for $\widehat{\mathbf{q}}^{l+1}$ is computed (such as $\widehat{\mathbf{q}}^{l+1} = \mathbf{q}^l + h\mathbf{v}^l$). Then the LCP (17) (or (29)) is solved using $\mathbf{D} = \mathbf{D}(\widehat{\mathbf{q}}^{l+1})$ and $\mathbf{n} = \mathbf{n}(\widehat{\mathbf{q}}^{l+1})$. The resulting estimate for $\mathbf{q}^{l+1}$ is used as $\widehat{\mathbf{q}}^{l+1}$ for the next iteration. Provided the LCP (17) (or (29) for multiple contacts) has unique solutions and the solution operator is locally Lipschitz, then for sufficiently small h, the method is convergent. In practice the iteration converges very quickly.

4. Numerical implementation and results

The numerical method has been developed based on an implementation of Lemke's algorithm for solving LCP's. The version of Lemke's algorithm used is based on an explicit tableau, together with lexicographic degeneracy resolution. Rather than explicitly include all potentially active constraints, only those that were active at the previous step, plus the violated constraints at $\mathbf{q}^l + h\mathbf{v}^l$, are initially assumed to be active. Once the solution of the LCP has been found, the constraints are checked at the new value $\mathbf{q}^{l+1}$ for feasibility. If any other constraints are found to be violated, they are added to the set of potentially active constraints. All of the methods have been implemented in C using the Meschach library[39] to provide the linear algebra and basic data structures.

The first test problem used was a simple rod-and-table problem where a spinning rod falls onto a table in two dimensions. The rod in Figure 3 is not an ideal one-dimensional rod, but is in fact a two-dimensional object with straight parallel sides and rounded (semi-circular) ends. It impacts a fixed horizontal table inelastically where it eventually comes to rest. The generalized co-ordinate vector is $\mathbf{q} = [x, y, \theta]^T$, where (x, y) are the co-ordinates of the center of mass (y vertical), and θ is the angle relative to the horizontal of the rod. The two constraint functions (one for each end point) are $f_{\pm}(\mathbf{q}) = y \pm (l/2) \sin \theta \geq \rho$ where ρ is the radius of the end of the rod. The corresponding normal vectors are $\mathbf{n}_{\pm}(\mathbf{q}) = [0, 1, \pm(l/2)\cos\theta]^T$, and the four friction vectors are $\mathbf{d}_i = [\pm 1, 0, \pm(l/2)\sin\theta]^T$. The mass matrix $\mathbf{M}(\mathbf{q})$ is $\mathrm{diag}(m, m, J)$ where m is the mass of the rod, and J is its moment of inertia. Gravitation is the only external force applied to the rod.

The relevant specifications of the rod are as follows: length (excluding the ends) 0.5 m; mass 1 kg; half-width of rod (which is also the radius of the ends) 0.05 m; moment of inertia $0.002 \, \mathrm{kg \, m}^2$; coefficient of friction between the rod and the table is 0.6. The initial angle of the rod is $30°$ to the horizontal, with zero initial translational velocity, but with an initial rotational velocity of $4 \, \mathrm{rad \, s}^{-1}$. At the first contact, the standard methods would claim that there is no solution. The method presented here has no difficulty with this situation. The numbers in Figure 2 are the times for the corresponding configuration.

To demonstrate convergence of the algorithm, graphs of the numerical results for different values of step size h are shown in Figures 3, 4 and 5. Note the absence of numerical chattering in the solutions. However, there is a spike in the angular velocity for the first contact. This is because of the position-based time-stepping used: after the first step in which contact is made, $f(\mathbf{q}) = 0$; on the following step, the velocity is made tangential to the contact surface, with $\mathbf{v}^T \nabla f(\mathbf{q}) = 0$.

In these simulations, the rod begins by spinning and falling, and then collides with the table at time $t \approx 0.383$. After the impact, the rod slides while spinning. The angle and velocity values at the first impact would result in failure for conventional "instantaneous" methods. Later, the left-hand end of the rod eventually hits the table at time $t \approx 0.548$, and the rod then remains horizontal and slides left for a short time (0.02 seconds). The rod is then at rest for the remainder of the integration. Note that fairly fine step sizes need to be used to properly resolve all the details of the behavior. Step sizes of about 1/25th of a second are not sufficient.

Other tests were carried out using for a system of balls and a horizontal table which gave an opportunity for carrying out fully three dimensional tests on the
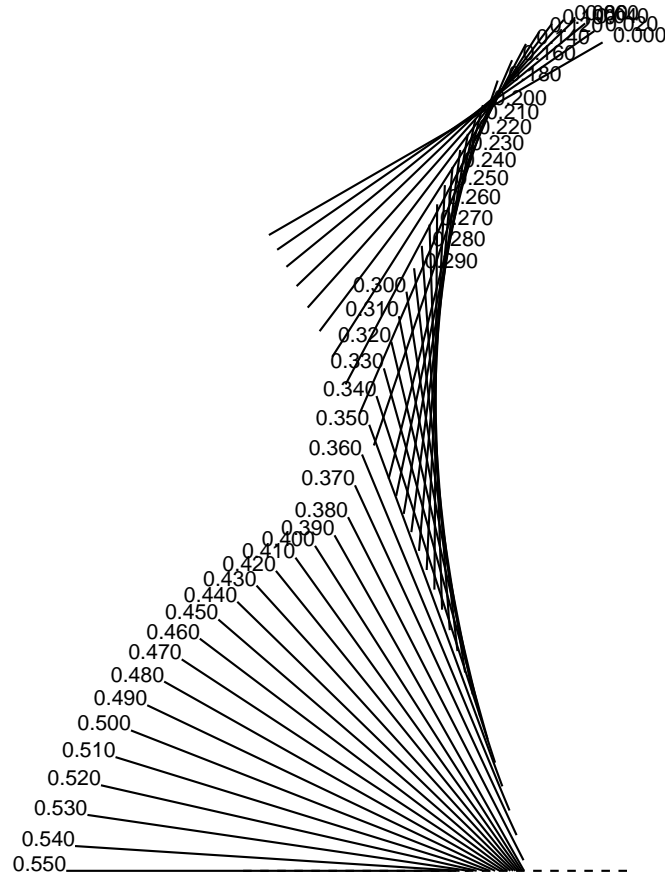
FIGURE 2. Falling and spinning rod, $h = 0.0025$

system. The first test of the balls involved a pair of balls of equal size and weight where one ball is thrown to either collide with the second stationary ball, or to land on the table and begin rolling before then colliding with the second stationary ball. In these problems, there is a potentially infinite number of direction vectors defining the friction cone for each contact; here we chose to use eight direction vectors to give an octagonal approximation to the friction cone.

The second test scenario is a set of three balls in a line along with the x-axis on the table, and another ball in the air thrown towards three balls. Initially there is a small gap of about 10^{-5}m between the initially stationary balls. The ball in the air first hits the table at $t \approx 0.42$, and begins rolling, and then collides with the line of balls at $t \approx 0.59$. The three initially stationary balls are brought into contact within the first time step of the collision. A frictional impulse applied at the first impact point lifts the thrown ball a small distance, due to the frictional impulse. During most of this period, all balls remain in contact. Most of the horizontal impulse is transmitted to the other balls. The corresponding downward impulse on the first stationary ball induces a vertical reaction force to prevent it from penetrating the table. Nevertheless, there is an impulsive frictional torque on
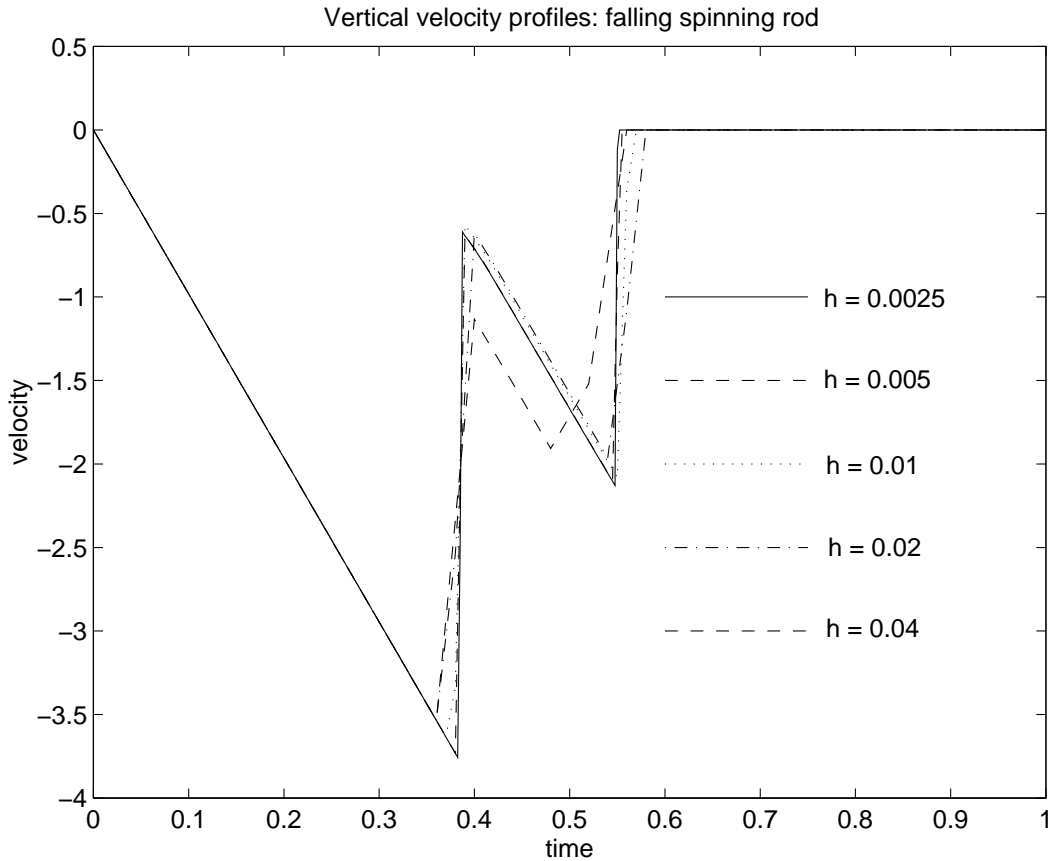
FIGURE 3. Vertical velocity of rod for different values of h

the first ball which makes it rotate contrary to the horizontal motion. There is also a frictional impulse in the negative x-direction due to the sliding and the downward frictional impulse. Because of the backward rotation on first stationary ball and the horizontal impulse, there is a frictional impulse on the second stationary ball to make it jump up, just as the thrown ball did. This forward rotation, in turn, induces an additional upward frictional impulse on the second stationary ball, as well as a downward frictional impulse on the third stationary ball. This downward impulse results in an upward reaction impulse from the table, and because of the rotations induced by the impulse torques, there is also a frictional impulse in the negative x-direction. This complex system of impulses is illustrated in Figure 6. The horizontal impulse makes the balls move horizontally, sliding at first, but then rolling. After a short period, the balls separate, rolling in different directions. The trajectories of the balls are shown in elevation and plan views in Figures 7 and 8.

The specific data for this problem is as follows: all balls are 1kg in mass, have a radius of 0.1m, with moments of inertia $4 \times 10^{-3} = (2/5)mr^2$; the coefficients of friction for all contacts is 0.4. The initial positions of the balls (given as x, y, z triples, with z being vertically up) are $[0, 0, 1]$, $[1, 0, 0.1]$, $[1.2 + 10^{-5}, 0, 0.1]$ and
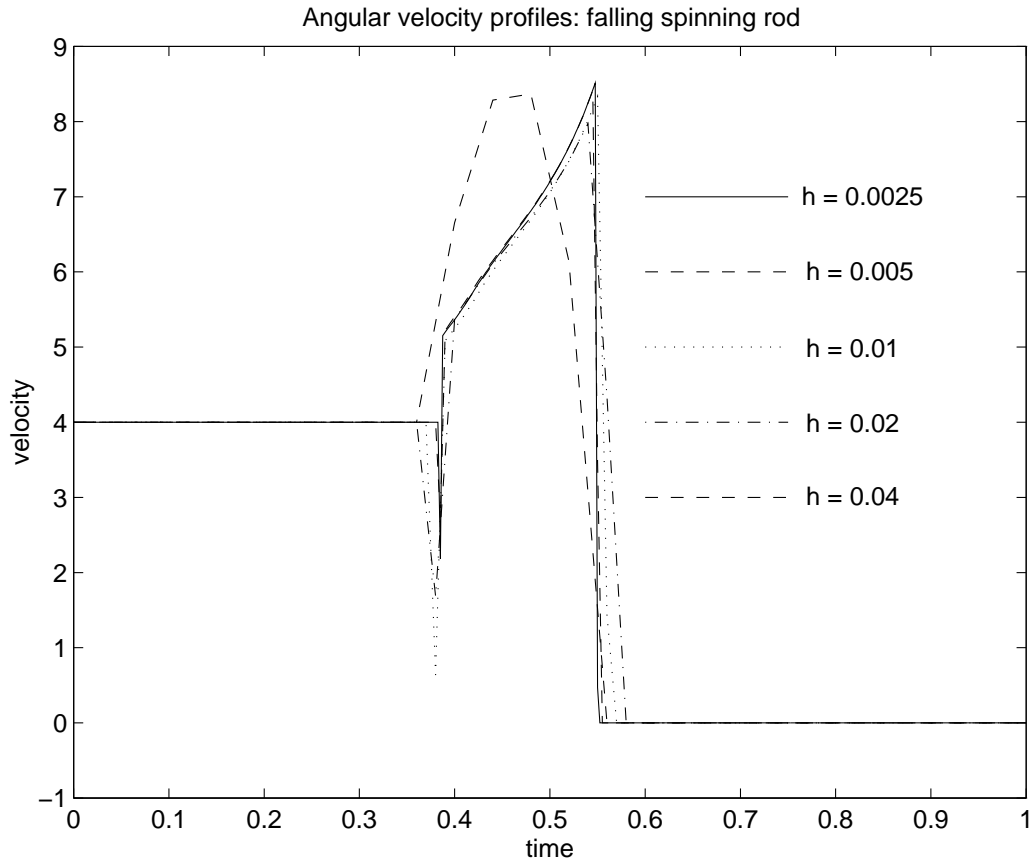
Angular velocity profiles: falling spinning rod

FIGURE 4. Angular velocity of rod for different values of h

$[1.4 + 2 \times 10^{-5}, 0, 0.1]$, and the initial velocities are $[1.5, 0.1, 0]$ for the first ball and zero for the others.

Table 1 gives error estimates and variation estimates for the numerically generated solutions for step sizes $h = 0.0025$ to $h = 0.02$ using the solution for $h = 0.00125$ as "exact". It also shows the variations of the numerically computed velocities, $V\mathbf{v}^h$, for different values of h. The error measures used are $\int_0^1 \|\mathbf{v}^h(t) - \mathbf{v}(t)\| \, dt$ for the velocities, and the supremum norm for the position errors. The ∞-norm is the norm used on $\mathbf{R}^n$ (here $n = 24$). The size of the LCP's solved for this problem were sometimes as large as 70×70 for 7 contacts. This is because there are 8 direction vectors for each frictional forces, one variable for each normal force, and a λ variable for each contact. It should be noted that the largest errors in the velocities were in the angular velocity components rather then the translation velocities; the angular velocity errors were about an order of magnitude larger. Table 1 shows that the errors are roughly of order h for small values of h.

REFERENCES

1. J.-P. Aubin and A. Cellina, *Differential Inclusions*, Springer-Verlag, Berlin, New York, 1984.
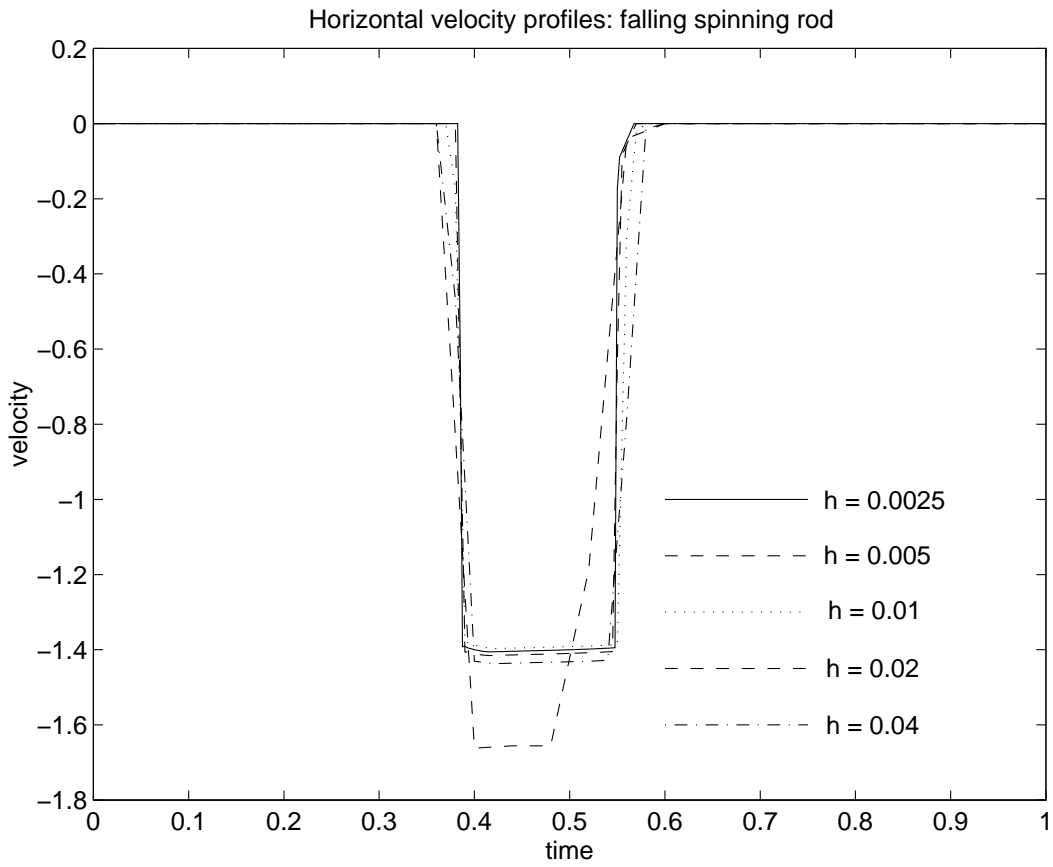
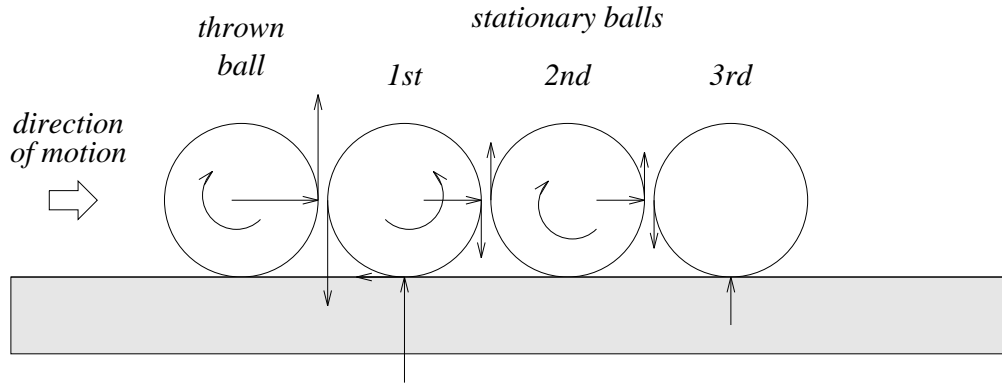FIGURE 5. Horizontal velocity of rod for different values of h

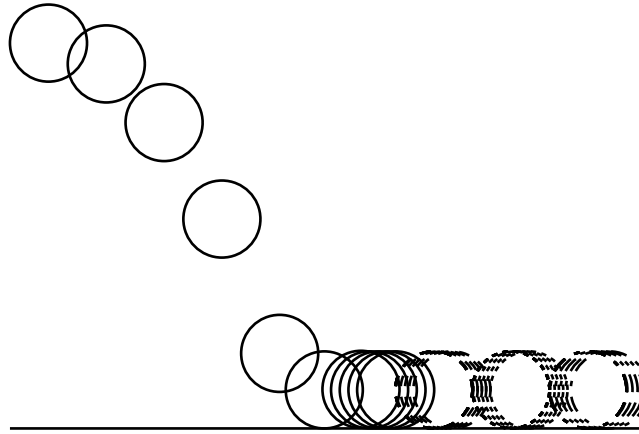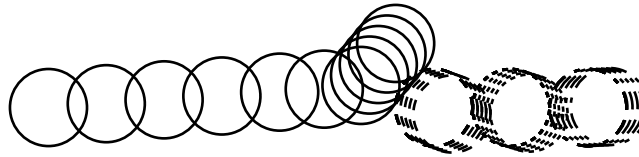FIGURE 6. Impulses at collision of balls

FIGURE 7. Elevation view of balls

FIGURE 8. Plan view of balls

Step size h	Velocity error	Position error	Velocity variation
0.02	0.5050	0.2505	19.4046
0.01	0.3523	0.2015	19.1728
0.005	0.1657	0.0838	19.1702
0.0025	0.0700	0.0298	19.0862
0.00125	—	—	19.0690

TABLE 1. Errors and variations for numerical solutions of four balls problems

2. J.-P. Aubin and H. Frankowska, *Set-Valued Analysis*, Progress in Systems and Control, #2. Birkhäuser, Boston, Basel, Berlin, 1990.
3. D. Baraff, 'Issues in computing contact forces for non-penetrating rigid bodies', *Algorithmica*, **10**, 292–352, (1993).
4. D. Baraff, 'Fast contact force computation for non-penetrating rigid bodies', In *Proceedings, SIGGRAPH '94*, pages 23–34. ACM, ACM, 1994.
5. F.H. Clarke, *Nonsmooth Analysis and Optimization*, SIAM Publ., Philadelphia, PA, 1990, Originally published by the Canadian Math. Soc., 1983.
6. D.L. Cohn, *Measure theory*, Birkhäuser, Boston, 1980.
7. R.W. Cottle, J.-S. Pang, and R.E. Stone, *The Linear Complementarity Problem*, Academic Press, Boston, San Diego, New York, 1992, Series on Computer Science and Scientific Computing.
8. K. Deimling, *Multivalued Differential Equations*, Walter de Gruyter, Berlin, New York, 1992, Series on Nonlinear Analysis and Applications #1.
9. E. Delassus, 'Considérations sur le frottement de glissement', *Nouv. Ann. de Math. (4ème série)*, **20**, 485–496, (1920).

10. E. Delassus, 'Sur les lois du frottement de glissement', *Bull. Soc. Math. France*, **51**, 22–33, (1923).
11. N. Dinculeanu, *Vector Measures*, Pergamon Press, London, New York, 1967.
12. A. Dontchev and F. Lempio, 'Difference methods for differential inclusions: A survey', *SIAM Rev.*, **34**, 263–294, (1992).
13. N. Dunford and J.T. Schwartz, *Linear Operators, Part I: General theory*, Wiley Interscience, New York, 1957.
14. A.F. Filippov, 'Differential equations with discontinuous right-hand side', *Amer. Math. Soc. Tranlations*, **42**, 199–231, (1964), Original in Russian in *Math. Sbornik*, **5**, pp. 99–127, (1960).
15. A.F. Filippov, *Differential Equations with Discontinuous Right-Hand Side*, Kluver Academic, 1988.
16. E.J. Haug, S.C. Wu, and S.M. Yang, 'Dynamics of mechanical systems with Coulomb friction, stiction, impact and constraint addition, deletion I, II and III', *Mechanism and Machine Theory*, **21**, 401–425, (1986).
17. R.D. Howe and M.R. Cutkosky, 'Practical force-motion models for sliding manipulation', *Internat. J. Robotics Research*, (1995), In press.
18. R.P. Kanwal, *Generalized functions: theory and techniques*, Mathematics in Science and Engineering #171. Academic Press, New York, London, 1983.
19. A. Kastner-Maresch, *Diskretisierungsverfahren zur Lösung von Differentialinklusionen*, PhD thesis, Universität Bayreuth, 1990.
20. A. Kastner-Maresch, 'Implicit Runge-Kutta methods for differential inclusions', *Numer. Funct. Anal. Optim.*, **11**, 937–958, (1990).
21. A. Kastner-Maresch, 'The implicit midpoint rule applied to discontinuous differential equations', *Computing*, **49**, 49–62, (1992).
22. C.W. Kilmister and J.E. Reeve, *Rational Mechanics*, American Elsevier, New York, NY, 1966.
23. A. Klarbring, 'Derivation and analysis of rate boundary-value problems of frictional contact', *Euro. J. Mech. A (Solids)*, **9**, 53–85, (1990).
24. C. Lanczos, *The Variational Principles of Mechanics*, Dover, Mineola, NY, 1986, Originally published by the University of Toronto Press, 1970.
25. P. Lötstedt, 'Coulomb friction in two-dimensional rigid-body systems', *Z. Angewandte Math. Mech.*, **61**, 605–615, (1981).
26. P. Lötstedt, 'Mechanical systems of rigid bodies subject to unilateral constraints', *SIAM J. Appl. Math.*, **42**(2), 281–296, (1982).
27. P. Lötstedt, 'Time-dependent contact problems in rigid-body mechanics', *Math. Prog. Study*, **17**, 103–110, (1982).
28. M.D.P. Monteiro Marques, *Differential Inclusions in Nonsmooth Mechanical Problems: Shocks and Dry Friction*, volume 9 of *Progress in Nonlinear Differential Equations and Their Applications*, Birkhäuser Verlag, Basel, Boston, Berlin, 1993.
29. M.T. Mason and Y. Wang, 'On the inconsistency of rigid-body frictional planar mechanics', In *Proceedings, IEEE Int. Conf. Robotics Automation*, pages 524–528, 1988.
30. J.K. Mills and C.V. Nguyen, 'Robotic manipulator collisions: modeling and simulation', *J. Dynamic Systems, Measurement Control*, **114**, 650–659, (1992).
31. J.J. Moreau, 'Frottement, adhésion, lubrification', *Comptes Rendus, Série II*, **302**, 799–801, (1986).
32. J.J. Moreau, 'Unilateral contact and dry friction in finite freedom dynamics', In J.J. Moreau and P.D. Panagiotopoulos, editors, *Nonsmooth Mechanics and Applications*, pages 1–82. Springer–Verlag, Vienna, New York, 1988, International Centre for Mechanical Sciences, Courses and Lectures #302.
33. P. Painlevé, 'Sur le lois du frottement de glissemment', *Comptes Rendus Acad. Sci. Paris*, **121**, 112–115, (1895), Following articles under the same title appeared in this journal, vol. 141, pp. 401–405 and 546–552 (1905).
34. J.-S. Pang and J.C. Trinkle, 'Complementarity formulations and existence of solutions of dynamic multi-rigid-body contact problems with Coulomb friction', Technical Report TAMU-CS TR94-057, Texas A&M University, 1994.
35. J.-S. Pang, J.C. Trinkle, and G. Lo, 'A complementarity approach to a quasistatic multi-rigid–body contact problem', *Comp. Optimization Appl.*, (1995), To appear.
36. J.M. Sanz-Serna, 'Symplectic integrators for Hamiltonian problems: an overview', In A. Iserles, editor, *Acta Numerica 1992*, pages 273–279. Cambridge Uni. Press, 1992.

37. D.E. Stewart, 'A high accuracy method for solving ODEs with discontinuous right-hand side', *Numer. Math.*, **58**, 299–328, (1990).

38. D.E. Stewart, 'Numerical methods for friction problems with multiple contacts', *J. Austral. Math. Soc., Ser. B*, (1995), Accepted May, 1994.

39. D.E. Stewart and Z. Leyk, *Meschach: Matrix Computations in C*, Australian National University, Canberra, 1994, Proceedings of the CMA, #32.

40. K.R. Stromberg, *An Introduction to Classical Real Analysis*, Wadsworth, Belmont, CA, 1981.

41. W.J. Stronge, 'Rigid body collisions with friction', *Proc. Royal Society of London*, **A431**, 168–181, (1990).

42. K. Taubert, 'Differenzverfahren für Schwingungen mit trockener und zäher Reibung und für Reglungsysteme', *Numer. Math.*, **26**, 379–395, (1976).

43. K. Taubert, 'Converging multistep methods for initial value problems involving multivalued maps', *Computing*, **27**, 123–136, (1981).

44. J.C. Trinkle, J.-S. Pang, S. Sudarsky, and G. Lo, 'On dynamic multi-rigid-body contact problems with Coulomb friction', submitted to *Zeitschrift fur Angewandte Math. Mech.*, 1995.

MATHEMATICS DEPT., TEXAS A&M UNIVERSITY, COLLEGE STATION, TX 77843-3368