DYNAMICS, FRICTION, AND COMPLEMENTARITY PROBLEMS

D.E. STEWART[1,3] AND J.C. TRINKLE[2,3]

Abstract. An overview is presented here of recent work and approaches to solving dynamic problems in rigid body mechanics with friction. It begins with the *differential inclusion* approach to Coulomb friction where the normal contact force is known. Numerical methods for these differential inclusions commonly lead to linear or nonlinear complementarity problems. Then the *measure differential inclusion* formulation of rigid body dynamics due to Moreau and others is presented. A novel algorithm for time-stepping for rigid body dynamics is given which avoids the "non-existence" issues that arise in many other formulations. The time-stepping algorithm requires the solution of an NCP at each step. Results on the convergence of the numerical solutions to solutions in the sense of measure differential inclusions are given, and numerical results for an implementation of the method are shown.

Key Words. friction, differential inclusions, rigid body dynamics, complementarity problem

1. Introduction. This paper presents an overview of some recent progress in dealing with rigid body dynamics incorporating collisions, shocks and Coulomb friction, with special emphasis on the difficulties due to friction.

Circa 1781 Coulomb introduced a model of the frictional forces at the contact between solid bodies [8, pp. 319–322]. This model stated that the frictional forces between two such bodies is μ (the coefficient of friction) times the magnitude of the normal contact force in magnitude, and directed opposite to the direction of relative motion. The Coulomb model, while widely used and accurate enough in many engineering applications, it gives rise to difficult computational and analytical problems. Even where the normal contact forces are known and smoothly varying in time (in inclined plane problems for example; see Figure 1) the discontinuous nature of Coulomb friction makes the general problem difficult.

Even in this simple setting, considering only the one-dimensional aspect of the behavior, the equations of motion lead to a radical revision of the nature of solutions to these differential equations. From simple mechanics, the differential equations are, for the forward velocity along the ramp,

$$m\frac{dv}{dt} = mg\sin\theta + F_{ext}(t) - F(t) = mg\sin\theta + F_{ext}(t) - \mu mg\cos\theta\,\mathrm{sgn}\,v \quad (1)$$

where $F_{ext}(t)$ is an external force applied to the object on the ramp. Because of the discontinuity in the friction force term, (1) can no longer be guaranteed to have solutions once $v(t) = 0$. (Even Caratheodory's existence theorem [2] fails to apply.) Yet, in physically reasonable situations, v may be zero. In the context of figure 1, if the slope θ is small enough, and the external forces $F_{ext}(t)$ are also small enough, v will reach zero in finite time. If $0 < |mg\sin\theta + F_{ext}(t)| < \mu mg\cos\theta - \epsilon$ for any $\epsilon > 0$, then the solution $v(t)$ will reach $v(t) = 0$ in finite time t_s, and $v(t) = 0$ for $t > t_s$. Yet, $v(t) \equiv 0$ is *not* a solution of the differential equation as $\mathrm{sgn}\,0 = 0$ and

$$m\frac{dv}{dt} = 0 \neq mg\sin\theta + F_{ext}(t) - 0.$$

[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

[3] This research was supported by the National Science Foundation under grant IRI-9304734, the Texas Advanced Research Program under grant 999903-078, and the Texas Engineering Experiment Station under grant 32133–98890CS.

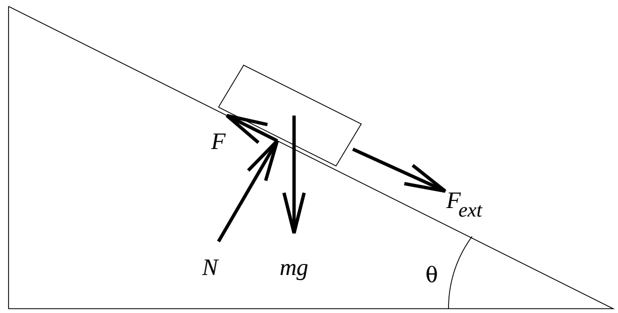

FIG. 1. *Inclined plane problem with friction*

The resolution of this difficulty is to extend the notion of differential equation to *differential inclusion*:

$$m\frac{dv}{dt} \in mg\sin\theta + F_{ext}(t) - \mu mg\cos\theta \,\mathrm{Sgn}\, v$$

where

$$\mathrm{Sgn}\, v = \begin{cases} \{+1\} & v > 0 \\ \{-1\} & v < 0 \\ [-1,+1] & v = 0. \end{cases}$$

This approach is implicitly used in most introductory treatments of friction. For example, as is described in [27, §5.4, p. 162], "If static friction is acting, the value of the friction force may vary from zero to its maximum available value, adjusting itself to the resultant force tending to cause motion."

1.1. Discontinuous ODE's and differential inclusions. In general, a discontinuous ODE

$$\frac{dx}{dt} = f(x)$$

should be replaced by the differential inclusion

$$\frac{dx}{dt} \in F(x) = \bigcap_{\delta>0} \bigcap_{\lambda_0(N)=0} \overline{\mathrm{co}} f((x+\delta B_0)\backslash N) \subset \mathbf{R}^n$$

where λ_0 is the Lebesgue measure on $\mathbf{R}^n$, B_0 is the standard unit ball in $\mathbf{R}^n$. In more intuitive language, $F(x)$ is the convex hull of values of $f(\cdot)$ nearby, ignoring the behavior on null sets. (Related versions of the Filippov regularization were developed by Krasovskij, Hermes and others. A suitable review paper for this area is Hájek [15].) The existence theory for such differential inclusions was developed by A.F. Filippov [11, 12], and these have been the subject of investigation over the past several decades. In the past two decades, there has been work on the numerical solution of these problems by a number of authors [10, 17, 23, 24, 28, 30, 33, 34, 35]. (See also the review article by Dontchev and Lempio [7].) These use a number of different techniques, but mostly concentrating on the use of suitable implicit and strongly stable Runge–Kutta schemes. The best of this work is by Kastner–Maresch [17], who is able to prove high order accuracy provided the solution is sufficiently smooth.

Stewart [28, 30], however, uses a different approach assuming that the problem itself is structured with regions where the right-hand side of the ODE is smooth, but with "manifolds of discontinuities" which can meet at boundaries. While on one of these manifolds of discontinuities, an equivalent (smooth) ODE can be devised based on the Filippov formulation. Where different manifolds of discontinuities meet, there is a combinatorial problem of deciding what manifold to move along in the next interval, and there may be several.

1.2. Connection with complementarity problems. So how are these differential inclusions related to complementarity problems? The answer lies in the solution techniques. Stewart [28, 30] explicitly uses linear complementarity problems to decide what manifolds to move along where the trajectory meets a new manifold. The use of, implicit Runge–Kutta methods, as are used by Kastner–Maresch [17] lead to *inclusions* for the result of a time-step, such as

$$x^{k+1} \in x^k + hF(x^{k+1}).$$

For many F such inclusions can be solved directly in terms of linear complementarity problems (such as for one-dimensional problems). Alternatively, by using a piecewise linear approximation to $F(\cdot)$, the problem can be "reduced" to that of solving a piecewise linear system of equations. These problems are known to be equivalent to linear complementarity problems [9], and furthermore, the methods used to solve such problems by continuation or homotopy methods [1] are exactly analogous to the pivoting algorithms of Lemke and Howson [18], Cottle and Dantzig [3], and others.

1.3. Contact problems. Here, we wish to go beyond previously studied problems where the contacts between bodies, and the normal contact forces, are assumed known, to general rigid body problems. In these problems, the normal contact forces are unknown, and must be computed along with the friction forces. However, the total contact force must lie in a certain *cone* called the *friction cone*. For isotropic surfaces, this cone is a right circular cone in ordinary coordinates, with the central axis of the cone being the normal vector to the surfaces in contact. If the contact is sliding, the contact force should lie on the boundary of the cone in the opposite direction of the relative motion at the point of contact. For anisotropic contacts (such as occurs in ice-skating), a better model would be to choose the friction vector in the friction cone in the direction that maximizes the rate of energy loss (see, for example, Goyal [14]). In general we will assume that the friction law satisfies the the maximum work inequality, which is valid for both isotropic and anisotropic Coulomb friction.

Contact problems go beyond the theory of differential inclusions, as general rigid body motion must include collisions and other "shocks" in order to preserve the rigid body and associated no-interpenetration assumption [20]. This leads to the study of *measure differential inclusions*, which has been pioneered by J.J. Moreau [21, 22] and Monteiro–Marques [20].

It should be noted that this research area is still subject to considerable controversy and academic debate, as many fundamental aspects are not properly resolved. For example, there is a continuing debate on the relative merits of the classical theories of collisions of Newton and of Poisson. Newton's law is based on velocities, while Poisson's law is based on momenta. Both laws can lead to an increase in the total energy, and are therefore physically flawed in some situations. Only recently has there been a modification which does preserve energy due to Stronge [32]. On the other hand, the Moreau approach is based on a *maximal dissipation principle* which best

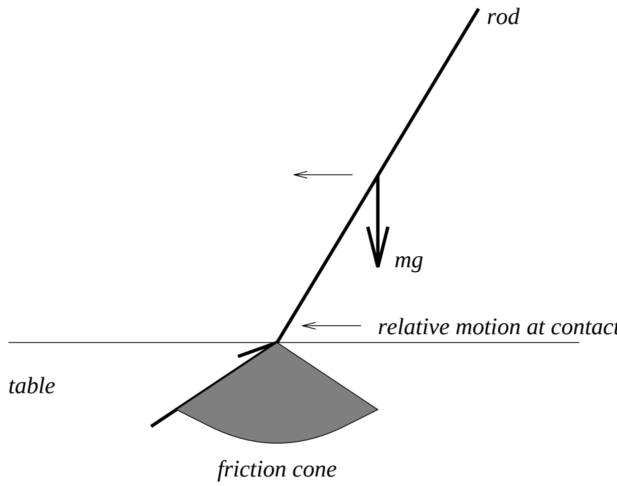

FIG. 2. *Apparent non-existence for a rod & table system*

describes purely inelastic collisions, but can be modified to deal with non-zero coefficients of restitution and still dissipate energy in all situations. Note that this maximal dissipation principle is an extension of the maximum work inequality described above.

1.4. Acceleration-based approaches and "non-existence". Indeed, the controversies in rigid body mechanics do not end there. There are a number of classical controversies about the existence of and uniqueness of solutions to rigid body problems. Some very simple problems involving just one moveable and one immoveable body give rise to apparent non-existence of solutions as are discussed in Lötstedt [19] which recapitulates some of the examples and questions of Delassus [5, 6] and Painlevé [25] last century.

While no claims are made here about removing the possibility of multiple solutions, the question of existence can be answered affirmatively, provided the reader is willing to accept impulsive forces and inelastic collisions. The problem with the examples mentioned above, is that the analysis is done "instant by instant". That is, the positions, orientations, velocities and angular velocities of all bodies are assumed fixed for an instant in time, and the forces are calculated in order to respect the friction law and the rigid body and non-interpenetration constraints. The problem with the rod and table example, is that if the coefficient of friction μ is too large, then a contact force on the boundary of the friction cone gives a torque that drives the end of the rod *into* the table, violating the non-interpenetration constraints. (See Figure 2.)

To explain in more detail, let m is the mass of the rod, l the half length of the rod, J the moment of inertia about the center of mass of the rod, which is assumed to lie at the center of the rod. The coordinates used are the (x, y) coordinates of the center of mass, and θ, the angle of the rod counter-clockwise from the horizontal. The vertical displacement of the left-hand tip of the rod is $y_c = y - l\sin\theta$, which must be non-negative to avoid penetration with a horizontal table at $y = 0$. Let N be the normal contact force, and F the (horizontal) frictional force; $|F| \leq \mu N$. Then the condition that $d^2y_c/dt^2 \geq 0$ is complementary to $N \geq 0$ leads (assuming that $dx_c/dt < 0$ so that $F = \mu N$) to the one-dimensional linear complementarity problem

(LCP), with $F = \mu N$:

$$d^2 y_c/dt^2 = \left[1/m + l^2 \cos^2\theta/J - \mu l^2 \sin\theta\cos\theta/J\right] N + (l \sin\theta\,(d\theta/dt)^2 - g) \geq 0,$$
$$N \geq 0, \qquad d^2 y_c/dt^2 \cdot N = 0.$$

If $1/m + l^2\cos^2\theta/J - \mu l^2 \sin\theta\cos\theta/J < 0$ and $l\sin\theta\,(d\theta/dt)^2 - g < 0$ then there is no solution to this LCP.

On the other hand, the above analysis ignores the possibility of impulsive forces, under which the assumption of fixed velocities even "for an instant" is wrong. This is an example of *jamming* where the contact velocity is driven to zero in an instant. If the velocity jumps to zero instantaneously, there is no longer any need to force the contact force (or impulse) to lie on the boundary of the friction cone. Then the normal component of the contact force or impulse can be made sufficiently large to ensure no interpenetration, while the frictional force or impulse is dissipative.

Specifically, with an impulse, the requirement that $F = \mu N$ arising from $dx_c/dt < 0$, need not hold. If the velocity dx_c/dt immediately after the impulse is zero, then it is only required that $|F| \leq \mu N$, and the coefficient of N in the resulting complementarity problem for the impulses is $1/m + l^2\cos^2\theta/J$ which is always positive.

The rod and table example illustrates the need to use impulses directly in the formulation of the dynamics of rigid body systems.

1.5. Multiple solutions. Philosophically there is a reason to treat multiple solutions less harshly than non-existence of solutions. Non-existence of solutions implies that the model is inherently wrong, since the physical system does not cease to exist when a situation where non-existence arises in the model. Multiplicity of solutions indicates that there is insufficient information in the model to *uniquely* predict the outcome of a situation (rather than claiming *no outcome*). And in these dynamics problems, the missing information most likely lies in the microscopic details of the contact between the bodies. The assumption of rigidity is of course not completely true: there will be some elastic give in the materials although the deformation will be measured in microns rather than millimeters or meters for most rigid materials (provided the elastic limits are not exceeded).

An example of multiplicity of solutions that can arise in rigid body mechanics is the rod example, but with $1/m+l^2\cos^2\theta/J-\mu l^2\sin\theta\cos\theta/J < 0$ and $l\sin\theta\,(d\theta/dt)^2 - g > 0$. In this case, there are two solutions for N: one with $N = 0$ and one with $N > 0$. There is also a third impulsive solution where the the velocity of the contact point is driven to zero. The second solution with $N > 0$ but finite (corresponding to smooth continued sliding) is actually very unstable. In particular, if the rigid table-top body is approximated by a stiff but elastic spring, then continued sliding corresponds to an unstable solution of the associated ordinary differential equations. Further, the growth rate for perturbations about the continued sliding solution increases without bound as the stiffness of the spring increases to infinity.

2. The measure differential formulation. The differential equation approach to writing down "equations of motion" needs to use set-valued right-hand sides to represent friction, and measures to represent impulsive forces; then it is a *measure differential inclusion*. For rigid body mechanics, the measure differential inclusions have the form

$$M(q)\frac{dv}{dt} + k(q,v) \in FC(q),$$

where $v(\cdot)$ is a function with bounded variation so that dv/dt is actually a *measure*. The matrix $M(q)$ is symmetric positive definite and is called the *mass matrix*, $k_l(q,v) = \frac{1}{2}\sum_{i,j}\left[(\partial m_{li}/\partial q_j) + (\partial m_{lj}/\partial q_i) - (\partial m_{ij}/\partial q_l)\right] v_i v_j - (\partial V/\partial q_l)$, and $FC(q)$ is the *friction cone*. In what follows, c_n is the normal contact force, and c_f is the frictional contact force. Then for a particle, the friction cone is commonly given by

$$FC(q) = \{\, c_n n(q) + c_f \mid c_f \perp n(q), \quad \|c_f\| \leq \mu c_n \,\}$$

for q on the boundary of the admissible region, with $n(q)$ being the *inward* pointing normal of the boundary; if q is not on the boundary, then $FC(q) = \{0\}$. For more general systems, $FC(q)$ is less easily described as it must be transformed from the physical (x,y,z) coordinates to generalized coordinates (which may include angles, orientations and related quantities). However, $FC(q)$ is still a closed convex cone. Furthermore, the graph of $FC(\cdot)$ is closed.

The interpretation of the measure differential inclusion is as follows: for any continuous $\phi \geq 0$ on $\mathbf{R}$ not everywhere zero,

$$\frac{\int \phi(t)\, M(q)\, dv(t) + \int \phi(t) k(q(t), v(t))\, dt}{\int \phi(t)\, dt} \in \overline{\text{co}} \bigcup_{\tau:\phi(\tau)\neq 0} FC(q(\tau)). \tag{2}$$

An alternative solution concept is given in Monteiro-Marques [20] which uses the singular decomposition of the measure dv and Radon–Nikodym derivatives. Provided the cone $FC(q)$ is *pointed* for all q (that is, $FC(q)$ does not contain a vector subspace), (2) is equivalent to the Monteiro-Marques definition. The equivalence of the two solution concepts has recently been shown in Stewart [29].

Additional requirements need to be imposed to ensure that the friction vector is opposite to the direction of motion, at least in the plane of contact, and that the impact is inelastic or elastic with a given coefficient of restitution. In this paper, only inelastic collisions are considered. Inelastic collisions are formulated by Moreau [22] and Monteiro-Marques [20] in terms of a *maximal dissipation* principle, where the choice of vector in the friction cone is made to maximize the rate of loss of energy (or amount of energy if there is an impulsive contact force).

3. Complementarity formulation for time-stepping. In this section a complementarity formulation of the time-stepping algorithm is given which can be solved by means of Lemke's algorithm for linear complementarity problems (LCP's). (See, e.g., [4].) This formulation is due to Stewart and Trinkle [31].

The representation of general rigid body systems will continue to use generalized coordinates in order to develop the most general framework for algorithm development. For example, for a rod in two-dimensions, the appropriate coordinate vector is $[x, y, \theta]^T$ where x is the horizontal displacement of the center of mass, y the vertical displacement of the center of mass, and θ is the angle of the rod from horizontal in the counter clockwise direction. The rod is assumed to have rounded ends with radius of curvature ρ.

3.1. Representation of friction cones. The friction cone $FC(q)$ is a closed convex cone which contains the inward normal $n(q)$. The friction cone is represented by an approximate friction cone $\widehat{FC}(q)$ which is the convex cone generated by the set $\{\, n(q) + \mu d_i(q) \mid i = 1, \ldots, m \,\}$. It is assumed that for each i there is a j such that $d_j(q) = -d_i(q)$, that $n(q)$ is not in the span of $\{\, d_i(q) \mid i = 1, \ldots, m \,\}$, and that no $d_i(q)$ lies in the convex hull of $d_j(q)$ for $j \neq i$.

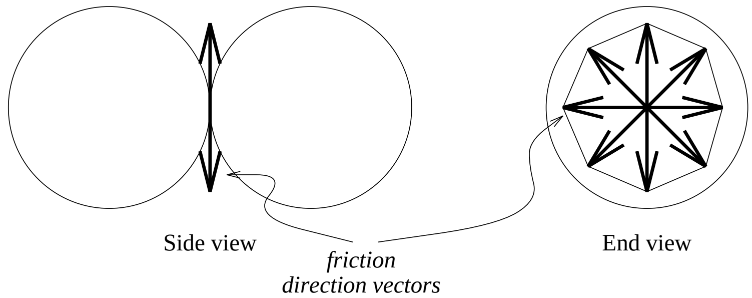

FIG. 3. *Approximation to circular friction cone*

This representation gives a great deal of flexibility; isotropic or anisotropic friction for contact in two or three dimensions, or even for soft finger models in which the frictional contact can support a frictional torque about the normal direction [16], can be represented. In generalized coordinates, the $d_i(q)$ vectors may incorporate torques, as well as as ordinary forces.

The friction direction vectors for an octagonal approximation to a circular friction cone are shown in Figure 3. Note that the frictional torques are not shown for clarity.

3.2. The formulation and the LCP. The task now is to compute q^{l+1} and v^{l+1} from q^l and v^l via the following LCP *for one contact*:

$$\begin{array}{rcl} M(v^{l+1}-v^l) & = & nc_n + D\beta + hk(q^l, v^l), \\ q^{l+1}-q^l & = & hv^{l+1}, \\ \lambda e + D^T v^{l+1} \geq 0 & \perp & \beta \geq 0, \\ n^T q^{l+1} \geq \alpha_0 & \perp & c_n \geq 0, \\ \mu c_n - e^T \beta \geq 0 & \perp & \lambda \geq 0. \end{array} \tag{3}$$

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. Its purpose is to ensure that the correct components of β are non-zero and that the method is dissipative.

The middle inequality of (3), $n^T q^{l+1} \geq \alpha_0$ is a guarantee that the (linearized) rigid body constraints are satisfied. The final inequality ensures that $D\beta$ lies in the approximate friction cone, and the corresponding complementarity condition ensures that if there is relative motion at the point of contact at the end of the time interval, that $D\beta$ lies on the boundary of the friction cone.

Multiple contacts can be dealt with using a suitable generalization of (3) as discussed in [31].

3.3. Solvability of the LCP. To show the solvability of the above mixed LCP, q^{l+1} and v^{l+1} are written in terms of c_n and β to give the Schur complement system

$$\begin{array}{c} \begin{bmatrix} D^T M^{-1} D & D^T M^{-1} n & e \\ n^T M^{-1} D & n^T M^{-1} n & 0 \\ -e^T & \mu & 0 \end{bmatrix} \begin{bmatrix} \beta \\ c_n \\ \lambda \end{bmatrix} + b = \begin{bmatrix} \rho \\ \sigma \\ \zeta \end{bmatrix} \geq 0 \quad \perp \quad \begin{bmatrix} \beta \\ c_n \\ \lambda \end{bmatrix} \geq 0 \\ \text{where} \quad b = \begin{bmatrix} D^T(v^l + hM^{-1}k) \\ (n^T q^l - \alpha_0)/h + n^T(v^l + hM^{-1}k) \\ 0 \end{bmatrix}. \end{array} \tag{4}$$

Writing this LCP as LCP(N, b) it should be noted that N is *copositive* (see Cottle, Pang and Stone [4, Defn. 3.8.1]), but not *copositive plus* ([4, Def. 3.8.1]). Nevertheless, solutions do exist ([4, Thm. 4.4.12]) since the only solutions of LCP$(N, 0)$ have the form $z^T = [0,\ 0,\ \lambda]$, and

$$z^T b = \begin{bmatrix} 0 \\ 0 \\ \lambda \end{bmatrix}^T \begin{bmatrix} D^T(v^l + hM^{-1}k) \\ (n^T q^l - \alpha_0)/h + n^T(v^l + hM^{-1}k) \\ 0 \end{bmatrix} = 0 \not< 0. \tag{5}$$

In fact, not only do solutions exist, but Lemke's algorithm, (described in [4, Algo. 4.4.1]) can compute solutions, provided precautions are taken against cycling due to degeneracy. (See [4, §4.9] for details on anti-cycling methods for LCP's.)

While solutions are guaranteed to exist, uniqueness is not guaranteed by these results, although for most problems it is expected to be unique.

3.4. The nonlinear versions. In practice, the complementarity problems that need to be solved are nonlinear, since the LCP formulation (3) uses only a linearization of the feasible set, and various quantities in the LCP (such as the mass matrix M, the direction vectors D, and the normal vector(s) n) all depend on q. While it is possible to ignore this dependence and evaluate the quantities at q^l rather than q^{l+1} (or some combination of these), it is important to perform the evaluation at q^{l+1} for n, α_0 and M.

The reason that n and α_0 must be evaluated at q, is that it is important for q^{l+1} to be exactly feasible. Without exact feasibility, the normal contact forces will force $q^h(\cdot)$ to be nearly feasible, but at a cost of dissipativity. Indeed, some numerical solutions computed by the author without forcing exact feasibility "blew up" with large increases in the kinetic energy that are quite unphysical. To ensure exact feasibility, $n = n(q^{l+1})$ and $\alpha_0 = n^T q^{l+1} - f(q^{l+1})$ where the feasible set is given by $\{\, q \mid f(q) \geq 0 \,\}$.

If the mass matrix used in (3) is $M = M(q^{l+1})$, then the local boundedness of the velocities can be proven. For computation, this does not seem to be crucial, but it is useful for the theory.

The existence of solutions to the nonlinear version can be shown using Brouwer's fixed point theorem. (For a description of Brouwer's fixed point theorem and related results, such as the Kakutani fixed point theorem, see [13].)

These nonlinear versions can be solved quickly and accurately in practice by a simple iteration, where q^{l+1} is approximated by $\widehat{q}^{l+1} = q^l + hv^l$, and the LCP (3) solved with M, n and α_0 evaluated at $\widehat{q}^{l+1}$. The solution of the LCP then gives a new approximation $\widehat{q}^{l+1}$, which can be used to re-evaluate M, n and α_0, which are used for the next iteration. This iteration is used until a small error tolerance is achieved. In practice, this iteration is quite effective, with convergence being roughly geometric with a factor of 10^{-2} or better, depending on the step size.

3.5. Dissipativity. Dissipativity for the case of constant M and constant k can be shown directly from the LCP (3) giving

$$\begin{aligned} \tfrac{1}{2}(v^{l+1})^T M v^{l+1} + k^T q^{l+1} &\leq \tfrac{1}{2} v^{l\,T} M v^l + k^T q^l - \lambda\mu c_n - \tfrac{1}{2}(v^{l+1} - v^l)^T M (v^{l+1} - v^l) \\ &\leq \tfrac{1}{2} v^{l\,T} M v^l + k^T q^l. \end{aligned}$$

Dissipativity for the general nonlinear case cannot be guaranteed in general; this is the case even for general smooth Hamiltonian systems. While there has been considerable recent work on *symplectic* integrators (see, e.g., the survey by Sanz–Serna [26]), these do not in general conserve energy, but can almost conserve an approximate energy function. Guaranteeing dissipativity for the general case is not appropriate for the numerical approximations, but rather dissipativity can be guaranteed for the limit solutions through the convergence theory.

4. Convergence and dynamics issues. With the results of the previous section, it is clear that the method can produce numerical trajectories. While the formulation appears correct for a single time-step, convergence of the numerical trajectories to true trajectories of the measure differential inclusion.

To discuss convergence issues, consider the step size $h \downarrow 0$. Let $v^{l;h}$ and $q^{l;h}$ be the computed generalized velocity and position respectively, for step l using step size h. To compare these with the true solution, we use the "interpolants" $v^h(\cdot)$ and $q^h(\cdot)$ defined as follows: Put $t_l = l\,h$. Then set $v^h(t) = v^{l+1;h}$ and $q^h(t) = q^{l;h} + (t-t_l)v^{l+1;h}$ for $t \in (t_l, t_{l+1}]$. This means that $dq^h/dt(t) = v^h(t)$ for almost all t, $q^h(\cdot)$ is locally Lipschitz, and $\mathrm{V}\,v^h = \sum_l \|v^{l+1;h} - v^{l;h}\|$.

The proof of convergence takes a number of steps which are as follows:

1. Show local boundedness for the numerically computed velocities $v^h(\cdot)$ via a discrete Gronwall lemma.
2. Given the (local) boundedness of the computed velocities, show that the variations $\mathrm{V}\,v^h$ are uniformly bounded.
3. Apply the Arzelà–Ascoli theorem to $q^h(\cdot)$ and the Moreau–Valadier theorem to $v^h(\cdot)$ to show convergence of a subsequence, uniformly for $q^h(\cdot)$ and pointwise for $v^h(\cdot)$ (weak* for the measures $dv^h(\cdot)$).
4. Show that limits of converging subsequences are solutions of the measure differential inclusion.
5. Show that the limits also satisfy a maximal dissipation principle.

While the full details are beyond the scope of this paper, some aspects of the convergence proof can be illustrated here.

Some basic assumptions that must be made are: the functions $M(\cdot)$, $n(\cdot)$, $D(\cdot)$ and $k(\cdot,\cdot)$ are all smooth and globally Lipschitz; the mass matrix $M(q)$ is positive definite (uniformly in q); $M(\cdot)$, $n(\cdot)$ and $D(\cdot)$ are all uniformly bounded above; the magnitude of $n(q)$ is bounded away from zero (or for multiple contacts, the smallest singular value of $n(q)$ is bounded away from zero).

4.1. Local boundedness. Local boundedness for the numerical velocities as the step size h goes to zero can be shown using the dissipativeness of the numerical scheme for constant M and k. The result below assumes that the method is implicit in the mass matrix; that is, the value of M used in (3) is $M(q^{l+1})$. The local bound can be shown by determining a local upper bound on the kinetic energy $KE^{l;h} = \frac{1}{2}(v^{l;h})^T M(q^{l;h}) v^{l;h}$ of the numerical solution:

$$\limsup_{h\downarrow 0} KE^{\lfloor t/h \rfloor;h} \le \theta(t)$$

where $\theta(t)$ is the solution of the differential equation $d\theta/dt = a\theta^{3/2} + b\theta^{1/2}$. Although solutions of this differential equations do blow up in finite time, it does give *local* boundedness.

To obtain boundedness for an unbounded time interval, it suffices to obtain convergence of the numerical solution for any sufficiently small time interval, together with boundedness of the exact solutions.

If M is constant (so that $k(q,v) = -\nabla V(q)$, which is assumed bounded), then global boundedness of the numerical solutions can be obtained directly:

$$\limsup_{h\downarrow 0} KE^{\lfloor t/h\rfloor;h} \leq \theta(t)$$

where $d\theta/dt = a + b\sqrt{\theta}$ with solutions $\theta(t) \leq c(1+t^2)$ for some constant c, so that the velocity grows at most linearly in time.

4.2. Bounded variation. The local boundedness of the velocities can then be used to show uniformly bounded variation of the numerical velocity functions $v^h(\cdot)$ (uniform as $h \downarrow 0$), at least on sufficiently small open intervals. The proof relies on the pointedness of the friction cones $FC(q)$ for all q and the assumption that $FC(\cdot)$ has a closed graph. However, the "infinite friction" problem with $\mu = +\infty$ appears to lead to difficulties, although no specific paradoxes or inconsistencies for this case are known to the author. Once the result is known for open intervals, it then follows that $v^h(\cdot)$ has uniformly bounded variation as $h \downarrow 0$ on any compact interval in which the numerical velocities are uniformly bounded as $h \downarrow 0$.

4.3. Convergence. To complete the convergence theorem, not that uniform boundedness of $v^h(\cdot)$ and the variation of $v^h(\cdot)$ imply that there is a converging subsequence of $(v^h(\cdot), q^h(\cdot))$ with a limit $(v(\cdot), q(\cdot))$ with $v^h(\cdot) \to v(\cdot)$ weakly in BV. The limiting velocity v has bounded variation, and $dq/dt = v$ is satisfied. Further, the measure differential inclusion

$$M(q)\frac{dv}{dt} - k(q,v) \in \widehat{FC}(q)$$

is satisfied. (This step requires an appropriate weak closure property for solutions of measure differential inclusions.)

Also, the maximal dissipativity property can be shown to hold for the limiting solution. From the numerical dissipativity results for constant M and k, it can be shown that the *limiting* solution is dissipative for smooth $M(q)$, $k(q,v)$ and feasible region. Exact dissipativity can improve the situation greatly, as it implies that the limit has *globally* bounded velocities, and convergence is not restricted to a fixed finite interval.

5. Numerical results. The time-stepping method has been implemented and numerical results obtained. Two main test problems have been used: a falling, spinning rod; and a set of four balls.

To demonstrate convergence of the algorithm, graphs of the numerical results for different values of step size h are shown in Figures 5 and 6. Note the absence of numerical chattering in the solutions.

For the balls, Figures 7 and 8 shows the elevation and plan respectively of their trajectories for a step size of 0.01. Also, Table 1 shows how the errors change with the step size; the velocity errors estimated by $\|v^h - v\|_1$, and the position errors estimated by $\|q^h - q\|_\infty$ (the ∞-norm was used on $\mathbf{R}^{24}$). As is evident from Table 1, the errors show a rough $O(h)$ behavior.

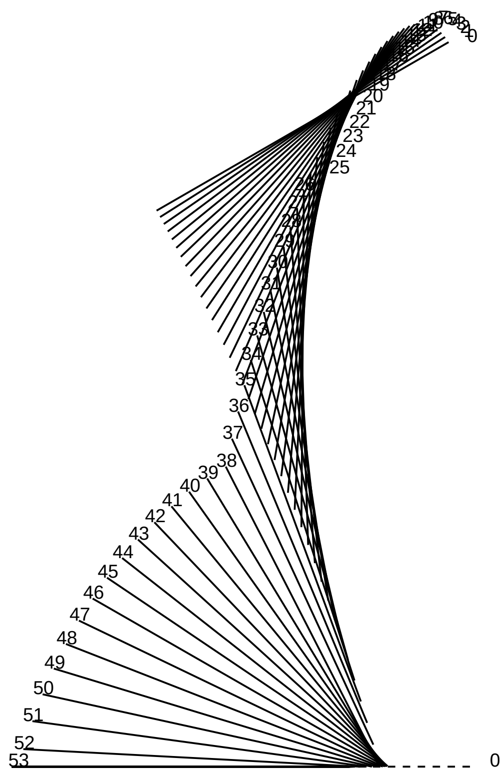

FIG. 4. *Falling and spinning rod, $h = 0.01$*

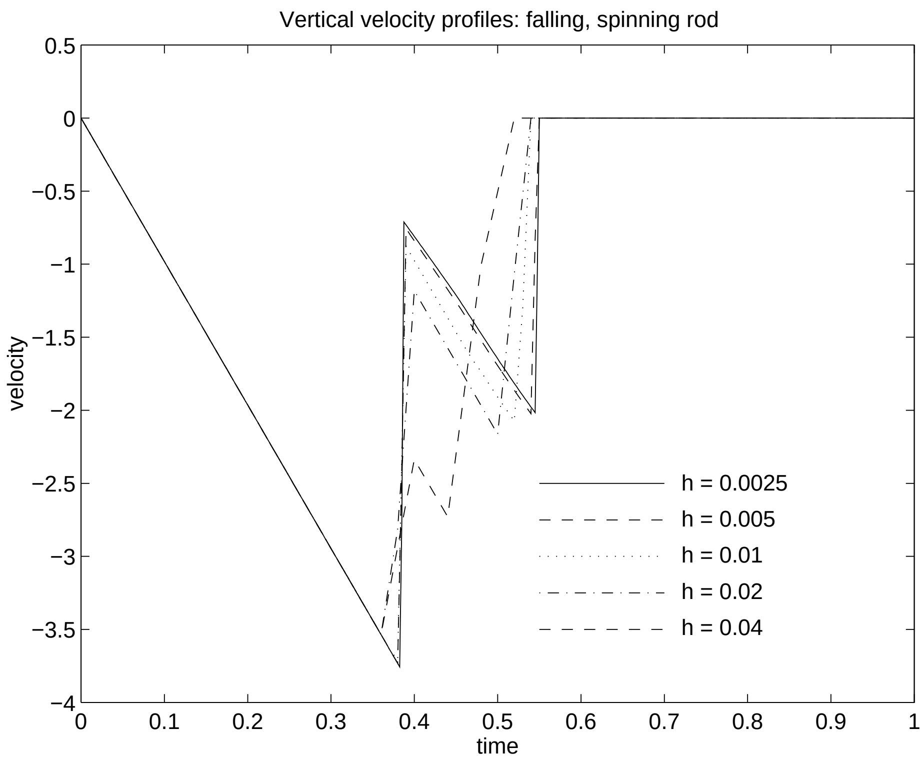

FIG. 5. *Vertical velocity of rod for different values of h*

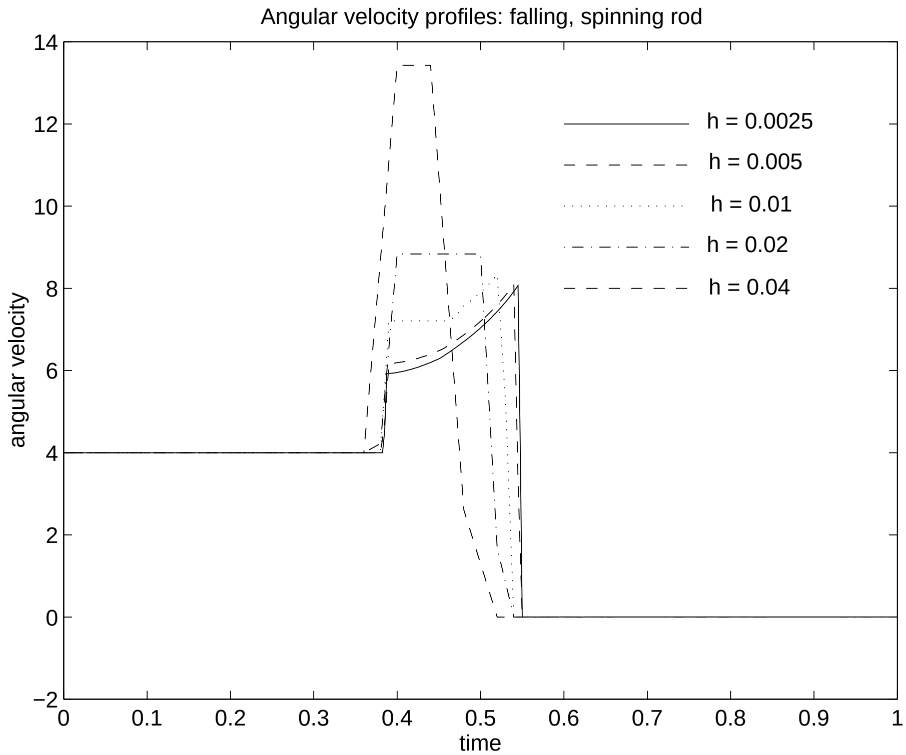

FIG. 6. *Angular velocity of rod for different values of h*

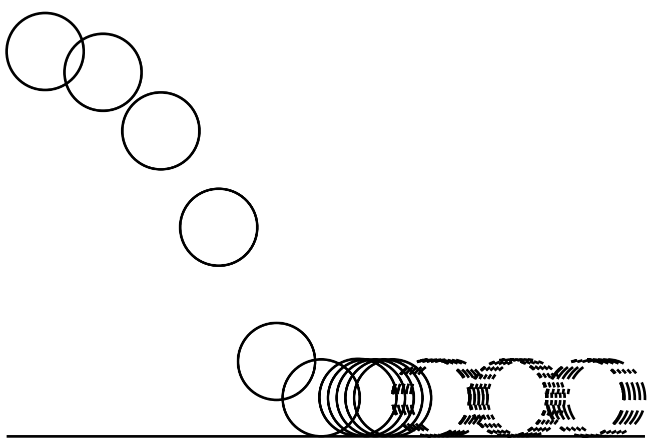

FIG. 7. *Elevation view of balls*

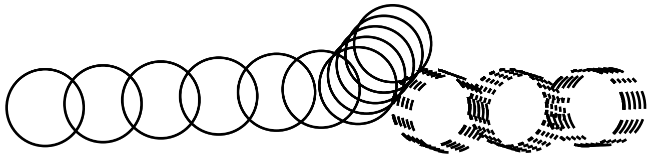

FIG. 8. *Plan view of balls*

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

REFERENCES

[1] E. Allgower and K. Georg, *Numerical Continuation Methods: an introduction*, Springer–Verlag, Berlin, Heidelberg, New York, 1990.

[2] E. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, Tata–McGraw Hill, New Delhi, 1972. Orignally published by McGraw Hill, 1955.

[3] R. Cottle and G. Dantzig, *Complementary pivot theory of mathematical programming*, Lin. Alg. and its Appl., 1 (1969), pp. 103–125.

[4] R. Cottle, J.-S. Pang, and R. Stone, *The Linear Complementarity Problem*, Academic Press, Boston, San Diego, New York, 1992. Series on Computer Science and Scientific Computing.

[5] E. Delassus, *Considérations sur le frottement de glissement*, Nouv. Ann. de Math. (4ème série), 20 (1920), pp. 485–496.

[6] ———, *Sur les lois du frottement de glissement*, Bull. Soc. Math. France, 51 (1923), pp. 22–33.

[7] A. Dontchev and F. Lempio, *Difference methods for differential inclusions: A survey*, SIAM Rev., 34 (1992), pp. 263–294.

[8] R. Dugas, *A History of Mechanics*, Éditions du Griffon, Neuchatel, Switzerland, 1955. Original in French; translation to English by J.R. Maddox.

[9] B. Eaves and C. Lemke, *equivlance of linear complementarity problems and piecewise linear systems*, Math. Oper. Res., 6 (1981), pp. 475–484.

[10] C. Elliott, *On the convergence of a one-step method for the numerical solution of ordinary differential inclusions*, IMA J. Numer. Anal., 5 (1985), pp. 3–21.

[11] A. Filippov, *Differential equations with discontinuous right-hand side*, Amer. Math. Soc. Tranlations, 42 (1964), pp. 199–231. Original in Russian in *Math. Sbornik*, **5**, pp. 99–127, (1960).

[12] ———, *Differential Equations with Discontinuous Right-Hand Side*, Kluver Academic, 1988.

[13] J. Franklin, *Methods of Mathematical Economics*, Springer-Verlag, New York, Heidelberg, Berlin, 1980.

[14] S. Goyal, *Planar sliding of a rigid body with dry friction: limit surfaces and dynamics of motion*, PhD thesis, Mechanical Eng., Cornell University, Jan. 1989.

[15] O. Hájek, *Discontinuous differential equations I & II*, J. Diff. Equ., 32 (1979), pp. 149–185.

[16] R. Howe and M. Cutkosky, *Practical force-motion models for sliding manipulation*, Internat. J. Robotics Research, (1995). In press.

[17] A. Kastner-Maresch, *Implicit Runge–Kutta methods for differential inclusions*, Numer. Funct. Anal. and Optim., 11 (1991), pp. 937–958.

[18] C. Lemke and J. J.T. Howson, *Equilibrium points of bimatrix games*, J. SIAM, 12 (1964), pp. 413–423.

[19] P. Lötstedt, *Coulomb friction in two-dimensional rigid-body systems*, Z. Angewandte Math. Mech., 61 (1981), pp. 605–615.

[20] M. M. Marques, *Differential Inclusions in Nonsmooth Mechanical Problems: Shocks and Dry Friction*, vol. 9 of Progress in Nonlinear Differential Equations and Their Applications, Birkhäuser Verlag, Basel, Boston, Berlin, 1993.

[21] J. Moreau, *Frottement, adhésion, lubrification*, Comptes Rendus, Série II, 302 (1986), pp. 799–801.

[22] ———, *Unilateral contact and dry friction in finite freedom dynamics*, in Nonsmooth Mechanics and Applications, J. Moreau and P. Panagiotopoulos, eds., Springer–Verlag, Vienna, New York, 1988, pp. 1–82. International Centre for Mechanical Sciences, Courses and Lectures #302.

[23] H.-D. Niepage, *Inverse stability and convergence of difference approximations for boundary value problems for differential inclusions*, Numer. Funct. Anal. and Optim., 9 (1987),

pp. 761–778.

[24] H.-D. Niepage and W. Wendt, *On the discrete convergence of multistep methods for differential inclusions*, Numer. Funct. Anal. and Optim., 9 (1987), pp. 591–617.

[25] P. Painlevé, *Sur le lois du frottement de glissemment*, Comptes Rendus Acad. Sci. Paris, 121 (1895), pp. 112–115. Following articles under the same title appeared in this journal, vol. 141, pp. 401–405 and 546–552 (1905).

[26] J. Sanz-Serna, *Symplectic integrators for Hamiltonian problems: an overview*, in Acta Numerica 1992, A. Iserles, ed., Cambridge Uni. Press, 1992, pp. 273–279.

[27] F. Singer, *Engineering Mechanics: Statics and Dynamics*, Harper & Row, 1975.

[28] D. Stewart, *A high accuracy method for solving ODEs with discontinuous right-hand side*, Numer. Math., 58 (1990), pp. 299–328.

[29] ———, *Asymptotic cones and measure differential inclusions*, tech. report, Dept. Mathematics, Texas A&M University, 1995.

[30] ———, *Numerical methods for friction problems with multiple contacts*, J. Austral. Math. Soc., Ser. B, (1995). Accepted May, 1994.

[31] D. Stewart and J. Trinkle, *An implicit time-stepping scheme for rigid body dynamics with inelastic collisions and coulomb friction.* Submitted to *International J. Numerical Methods in Engineering*, 1995.

[32] W. Stronge, *Rigid body collisions with friction*, Proc. Royal Society of London, A431 (1990), pp. 168–181.

[33] K. Taubert, *Differenz verfahren für gewöhnliche anfangswertaufgaben mit unstetiger rechte seite*, in Numerische Behandlung nichtlinearer Integrodifferential- und Differentialgleichungen, A. Dold and B. Eckmann, eds., 1974, pp. 137–148. Lecture notes ser#395.

[34] ———, *Differenzverfahren für schwingungen mit trockener und zäher reibung und für regelungsysteme*, Numer. Math., 26 (1976), pp. 379–395.

[35] ———, *Converging multistep methods for initial value problems involving multivalued maps*, Computing, 27 (1981), pp. 123–136.