New Methods in Applied and Computational Mathematics (NEMACOM'98)
Proceedings of the Centre for Mathematics and its Applications

Simulation of Rigid-Body Dynamics with Impact and Friction

David E. Stewart

*Department of Mathematics, University of Iowa,
Iowa City, IA 52242, USA.*

1. Impact and friction

Rigid-body dynamics is the dynamics of bodies that do not deform. While no body is completely rigid, this is a good model for a wide range of everyday objects which are stiff on the time and length scales of interest to us.

Another aspect that is of great importance in everyday life are Coulomb (or dry) friction, and impact. Whether we are walking, or playing sports (hitting balls, running, etc.), or operating a car (clutching), or picking up objects, impact and friction are commonplace effects. Yet the theoretical understanding of these phenomena is still in its infancy. One of the reasons for this are the discontinuities introduced by standard models of these phenomena. Impacts of rigid bodies clearly leads to discontinuities of the velocity. A lesser, but still very significant, discontinuity is due to Coulomb friction, where the equations of motion involves discontinuous functions.

These discontinuities lead to problems both in terms of theory and in terms of computation. Impacts are perhaps the more severe form of discontinuity: there are impulses in the contact forces which leads to discontinuities in the velocities (which are part of the state vector of the system). Such impulsive forces are best modelled mathematically as measures. While measure differential equations have been around since at least the 1950's (see, for example [12]), in this case the strength of the impulse is determined by the configuration of the mechanical system, and so is not known *a priori.* This leads to measure differential inclusions and complementarity conditions between measures and (Borel measurable) functions.

Coulomb friction forces are bounded relative to the normal contact forces, but the additional discontinuity is that the direction in which they apply is discontinuous in the relative slip velocity when that velocity is zero. Since "sticking" is fairly common with Coulomb friction, it means that this discontinuity has to be dealt with.

1.1. Impact Impacts in mechanical systems are extremely common, yet difficult to realistically model. Since for rigid bodies, impacts are instantaneous, there needs to be some rule which specifies just how the bodies behave in an impact. Consider the difference between two billiard balls colliding and two lumps of play-dough colliding. The former will bounce in a nearly elastic way, almost conserving the apparent kinetic energy, while the latter will undergo plastic deformation in the impact, and have little kinetic energy available for separating afterwards.

This is commonly modelled using Newton's law of impacts which states that the normal component of the relative velocity after collision is $-e$ times the normal component of relative velocity just prior to collision [28]. As it is commonly applied, it is known to sometimes give an increase in the total energy when friction is involved [34]. The quantity e is called the *coefficient of restitution.*

An alternative version of this is Poisson's impact law which states that the impact should be divided into compression and expansion phases: the normal contact impulse for the expansion phase should be e times the impulse for the compression phase. The compression phase should behave like an inelastic impact (in the Newtonian sense). (A fairly thorough derivation of the Poisson impact laws is given in [28].) However Poisson impact laws are known to give loss of energy for some situations with $e = 1$, which should give perfectly elastic impacts which are energy conserving.

Stronge [34] has also given other impact laws which are based on the idea of ensuring that the post-impact kinetic energy associated with the normal velocity is e^2 times the pre-impact kinetic energy associated with the normal velocity.

Recent work [2, 33] suggests that none of the current impact laws is realistic. Further progress in this area will probably require a deeper understanding of the dynamic behaviour of elastic bodies in impact. This area, too, is underdeveloped, and there are not even existence results for simple dynamic impact problems with Signorini (hard contact) conditions, with or without friction. (This is notwithstanding the recent paper of Jarušek and Eck [11].)

1.2. Coulomb friction Coulomb friction was first formulated by C.A. Coulomb in 1781 [5], based on earlier ideas of Amontons in the 1600's. This formulation was essentially the following three rules:

- The friction force between sliding objects is proportional to the normal contact force (the constant of proportionality is the *coefficient of friction* μ).
- The direction of the friction force is opposite to the direction of relative motion between the bodies.
- If there is no slip, then the friction force can take any value with magnitude no greater than μ times the normal contact force.

These three rules cover isotropic friction. More general anisotropic friction laws are needed for dealing with ice-skating, for instance. A more modern re-formulation of Coulomb's laws can be found in Goyal [10]. This formulation is the *maximal dissipation principle*: the friction force is the one that *maximizes dissipation of energy*:

$$\min_{c_f} -c_f^T v_{rel} \qquad \text{subject to} \qquad c_f \in c_n\, FC_0 \tag{1.1}$$

where FC_0 is a closed, balanced, convex set which gives the possible friction forces for unit normal contact force. Note that v_{rel} is the relative velocity vector for the pair of sliding bodies at the point of contact. Also note that c_n is the normal contact force. By making FC_0 highly elliptical, with the major axis perpendicular to the line of the skates, the frictional behaviour of ice-skates on ice can be adequately represented.

Given a representation of FC_0, this optimization problem can be solved using the Karush–Kuhn–Tucker (KKT) conditions *given* c_n. Note however, that the contact forces overall (c_n as well as c_f) at any instant is *not* an optimization problem. Attempting to do so with pseudo-potentials (see, e.g., Moreau [20, 21], Duvaut and Lions [7]) leads to a number of difficulties such as non-convex objective functions, and non-existence of solutions in infinite dimensions. It is rather better to treat it as a complementarity or fixed-point problem.

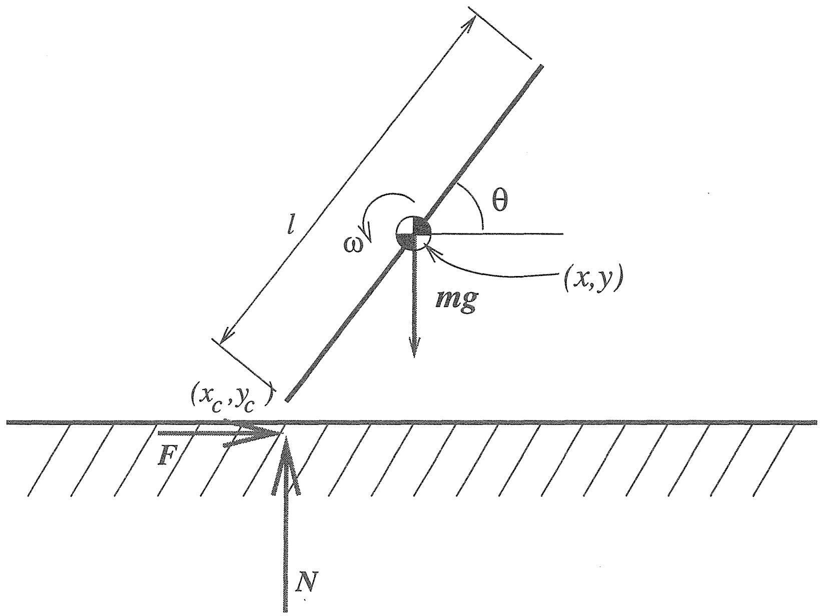


FIGURE 1. Painlevé's problem

1.3. Painlevé's problem In 1895 P. Painlevé published a short paper [22] which seemed to show that rigid-body dynamics with friction was inconsistent. Until very recently, this problem was seen by many people in this area as showing that rigid-body dynamics with Coulomb friction as being inconsistent. More recent work of Moreau, Monteiro–Marques, and Stewart and others have changed this perception, although aspects of the solution to this problem were apparent to some by the 1920's such as Delassus [6].

Painlevé's problem is a simple problem of a rod touching a table while in motion (see Figure 1).

Let N be the normal contact force at the contact point, and F the corresponding horizontal friction force. Let m be the mass of the rod, and J its moment of inertia. Also, let x and y respectively be the horizontal and vertical coordinates of the center of mass of the rod; let θ denote the angle of the rod with respect to horizontal. See the diagram in Figure 1. The contact point is given by $(x_c, y_c) = (x - (L/2)\cos\theta, y - (L/2)\sin\theta)$. Suppose that $\dot{x}_c < 0$ and $\dot{y}_c = 0$ at some instant in time. The equations of motion are given by

$$\begin{aligned} m\,\ddot{x} &= F = \mu N \\ m\,\ddot{y} &= N - mg \\ J\,\ddot{\theta} &= (L/2)[+F\sin\theta - N\cos\theta] \\ &= (L/2)[\mu\sin\theta - \cos\theta]\,N \end{aligned} \tag{1.2}$$

Then

$$\begin{aligned} \ddot{y}_c &= \ddot{y} - (L/2)\cos\theta\,\ddot{\theta} + (L/2)\sin\theta\,\dot{\theta}^2 \\ &= N/m - g - (L^2/4J)\cos\theta[\mu\sin\theta - \cos\theta]\,N + (L/2)\sin\theta\,\dot{\theta}^2. \end{aligned}$$

The contact conditions are that $\ddot{y}_c \geq 0$ (no interpenetration), $N \geq 0$ (no adhesion), and $N\,\ddot{y}_c = 0$ (contact breaks implies no contact force). If $1/m-(L^2/4J)\cos\theta[\mu\sin\theta-\cos\theta] < 0$ and $-g+(L/2)\sin\theta\,\dot{\theta}^2 < 0$ (which is true for certain values of mL^2/J, μ, θ, and $\dot{\theta}$) then $\ddot{y}_c < 0$ for *every* $N \geq 0$. In particular, choose $\dot{\theta} = 0$, make mL^2/J and μ large, and choose a suitable value for θ (say, $\pi/3$ radians).

The problem with this analysis is that we are assuming that $F = +\mu N$ because $\dot{x} < 0$. If impulses are allowed, then the horizontal velocity can be brought to zero instantaneously, allowing any F where $|F| \leq \mu N$. In this case, the problem is solvable. This is an example of an *impulsive force without impact.* This problem causes difficulties with

2. Complementarity problems

A complementarity problem is the task, given a function $f : \mathbf{R}^n \to \mathbf{R}^n$, of finding x where

$$0 \leq f(x) \quad \perp \quad x \geq 0 \tag{2.1}$$

where the inequalities are understood componentwise and "$a \perp b$" means that $a^Tb = 0$. These problems arise in a wide range of different contexts (economics, contact mechanics, game theory, electrical circuits). See [8] for a survey of applications. Since for non-negative vectors a and b, $a^Tb \geq 0$, the complementarity condition $a \perp b$ means that for each i, either $a_i = 0$ or $b_i = 0$. There are many variants on this including mixed complementarity problems ($0 \leq f(x,u) \perp x \geq 0$ in $\mathbf{R}^n$ and $0 = g(x,u)$ in $\mathbf{R}^m$ with $u \in \mathbf{R}^m$) and generalized complementarity problems (where "$x \geq 0$" means that $x \in K$ for a closed convex cone K, and "$0 \leq f(x)$" means that $f(x) \in K^*$, the dual cone $K^* = \{\, w \mid w^Tx \geq 0\,\forall x \in K\,\}$.

If $f(x) = Mx+q$ then the complementarity problem is called a *Linear Complementarity Problem* (LCP) which is the subject of an entire encyclopedic monograph [4]. Not even LCP's have solutions in general, and there is no compact non-trivial characterization of which LCP's have solutions. Even the topology of the solutions to LCP's is far from trivial. Nevertheless, there are a number of well-known classes of M for which there is a well-known and very useful (constructive) existence theory. Foremost amongst these are the *copositive matrices*: a matrix M is copositive if

$$z \geq 0 \quad \Rightarrow \quad z^TMz \geq 0. \tag{2.2}$$

If M is copositive and $s^Tq < 0$ whenever $s \geq 0$ satisfies $s^TMs = 0$, then there is a solution of the complementarity problem $0 \leq z \perp Mz+q \geq 0$. Furthermore, a solution can be found by Lemke's algorithm, which is a variant of the simplex method for solving complementarity problems. See [4] for more details.

Nonlinear Complementarity Problems (NCP's) can be guaranteed to be solvable under some conditions, although this is often done on a case–by–case basis. For example, if f is a strongly monotone function ($(x-y)^T(f(x)-f(y)) \geq c\|x-y\|^2 > 0$) then solutions exist for the NCP $0 \leq z \perp f(z) \geq 0$. For friction problems (which are not monotone in general), the nonlinear complementarity formulations of [23] are guaranteed to have solutions, which is shown using a homotopy argument. Other formulations, such as the early formulations of [13, 24], however, do not have solutions in general.

3. Formulating rigid-body dynamics

Rigid-body dynamics needs to be formulated as a continuous-time problem, and then for the purpose of computation and simulation, as a discrete-time or numerical problem. Of course, the limit(s) of the numerical solutions as the computational parameters (such as step-size) approach the appropriate limit(s), should be solutions of the continuous-time problem. However, in the case of rigid-body dynamics with impact and friction, it is a non-trivial matter just to formulate what the continuous-time problem should be. Normally, we would think of writing down a set of differential equations which would have solutions by standard results (such as Picard's contraction mapping result, or Carathéodory's existence theorem). Then we would look for a numerical scheme to accurately solve the differential equations.

For rigid-body dynamics with impact and friction, it is not possible to apply standard existence theories. The only existence results obtained so far, have been obtained by developing a numerical model, and showing that the numerical trajectories converge to continuous trajectories that solve the continuous-time problem. See, for example, the work of Monteiro-Marques [17] which gives the first rigorous proof of the existence solution of a rigid-body dynamics problem with impact and friction. Later work also uses this approach [26, 30].

3.1. Continuous time Below we see the continuous-time formulation of rigid-body dynamics with one contact.

The trajectories are given by functions $q(\cdot)$ whose values are generalized coordinates, and $v(\cdot)$ which are generalized velocities ($dq/dt = v$). In these generalized coordinates we have a Lagrangian $L(q,v) = T(q,v) - V(q)$ where $T(q,v)$ is the kinetic energy and $V(q)$ is the potential energy. We assume that $T(q,v) = \frac{1}{2}v^T M(q)v$ where $M(q)$ is a symmetric positive definite matrix (called the mass matrix). The admissible region is specified by a scalar-valued function $f(q)$: $C = \{\, q \mid f(q) \geq 0 \,\}$. The function f is assumed to be smooth and the normal direction vector $n(q) = \nabla f(q) \neq 0$ for any q on the boundary ($f(q) = 0$).

The normal contact force is in the direction $n(q)$. However, since we are using generalized coordinates, the plane of the friction forces is not necessarily perpendicular to $n(q)$. For particles represented in natural Cartesian coordinates, the friction plane is perpendicular to $n(q)$. However, for a body in two-dimensions, the configuration of the system is determined by three coordinates: (x,y), the Cartesian coordinates of the centre of mass, and θ, the angle of the body relative to a reference orientation. For such a body, the friction plane is typically not perpendicular to $n(q)$ due to the angular variable. In fact, this turns out to be crucial in analysing problems such as Painlevé's problem.

To describe the friction cone, we use a homogeneous convex function $\psi(\beta)$: $FC_0(q) = \{\, D(q)\beta \mid \psi(\beta) \leq \mu \,\}$. Homogeneity of ψ means that $\psi(\alpha\beta) = \alpha\,\psi(\beta)$ for any $\alpha \geq 0$. For $FC_0(q)$ to be a balanced convex set, we also need $\psi(\beta) = \psi(-\beta)$. For isotropic friction we can take $\psi(\beta) = \|\beta\|_2$. The complementarity conditions given below are the KKT conditions (as interpreted using generalized gradients *a la* Clarke [3]) for the convex programming problem

$$\min_{\beta} v^T D(q)\beta \quad \text{subject to } \psi(\beta) \leq \mu\, c_n. \tag{3.1}$$

The continuous problem with one contact and inelastic impacts can be formulated as one of finding a function $v(.)$ of bounded variation, and an absolutely continuous $q(.)$ along with the contact forces (measures) c_n the normal contact force, and β describing the frictional forces, and a bounded Borel function λ where

$$\begin{aligned}
M(q)\frac{dv}{dt} &= n(q)\,c_n + D(q)\,\beta - \nabla V(q) + k(q,v) + F_{ext}(t), && (3.2)\\
\frac{dq}{dt} &= v, && (3.3)\\
0 &\le c_n \quad \perp \quad f(q) \ge 0, && (3.4)\\
0 &\in \mu\, D(q)^T v^+ + \lambda\, \partial\psi(\beta), && (3.5)\\
0 &\le \lambda \quad \perp \quad \mu c_n - \psi(\beta) \ge 0, && (3.6)\\
0 &= n(q(t))^T v^+(t) \quad \text{if} \quad f(q(t)) = 0. && (3.7)
\end{aligned}$$

For partly or fully elastic impacts with coefficient of restitution $0 \le \varepsilon \le 1$, we can replace (3.7) with

$$n(q(t))^T(v^+ + \varepsilon v^-) = 0 \quad \text{if} \quad f(q(t)) = 0. \tag{3.8}$$

3.2. Numerically In order to handle problems like those arising with the Painlevé problem, a time-stepping approach which uses the integrals of the force functions (or measures) c_n and β over each time-step interval $[t_l, t_{l+1}]$. Two different numerical formulations are presented here. The first is based on linear complementarity problems and uses a polyhedral approximation $\widehat{FC}(q)$ to the friction cone $FC(q)$. The second is a nonlinear complementarity formulation which uses $\psi(\beta)$ directly.

The polyhedral approximation to the friction cone is the cone generated by $\{\, n(q) + \mu\, d_i(q) \mid i = 1, 2, \ldots, m \,\}$ where $\mu\, d_i(q)$ is a collection of direction vectors in $FC_0(q)$. Write $\widetilde{D}(q) = [d_1(q),\, d_2(q),\, \ldots,\, d_m(q)]$. The friction forces $D(q)\beta$ are approximated by $\widetilde{D}(q)\widetilde{\beta}$ where $\widetilde{\beta}_i \ge 0$ and $\sum_i \widetilde{\beta}_i \le \mu\, c_n$. The relationship between $\widehat{FC}(q)$ and $FC(q)$ is illustrated in Figure 2. It is assumed that for each i there is a j where $d_i(q) = -d_j(q)$. This is related to the assumption that $FC_0(q)$ is a *balanced* set: $FC_0(q) = -FC_0(q)$.

The discretization of is the problem of finding q^{l+1} and v^{l+1} (and the force integrals c_n^{l+1}, $\widetilde{\beta}^{l+1}$, and Lagrange multiplier λ^{l+1}) given q^l and v^l for a time-step of size $h > 0$ that satisfy the following conditions:

$$\begin{aligned}
M(q^{l+1})(v^{l+1} - v^l) &= n(q^l)c_n^{l+1} + \widetilde{D}(q^l)\widetilde{\beta}^{l+1} && (3.9)\\
&\qquad + h[-\nabla V(q^l) + k(q^l, v^l) + F_{ext}(t_l)] &&\\
q^{l+1} - q^l &= h\, v^{l+1}, && (3.10)\\
0 \le c_n^{l+1} &\perp n(q^l)^T(v^{l+1} + \varepsilon v^l) \ge 0, && (3.11)\\
0 \le \widetilde{\beta}^{l+1} &\perp \lambda^{l+1}\, e + \widetilde{D}(q^l)^T v^{l+1} \ge 0, && (3.12)\\
0 \le \lambda^{l+1} &\perp \mu\, c_n^{l+1} - e^T\widetilde{\beta}^{l+1} \ge 0, && (3.13)
\end{aligned}$$

where $f(q^l + h\, v^l) < 0$; if $f(q^l + h\, v^l) \ge 0$, then we set $c_n^{l+1} = 0$ and $\widetilde{\beta}^{l+1} = 0$ and solve the first two equations. Note that e is a vector of 1's of the appropriate size.

Given q^l and v^l the problem to compute q^{l+1} and v^{l+1} can be reduced in size by eliminating v^{l+1} and q^{l+1} in terms of the impulses c_n^{l+1} and β^{l+1} via equations (3.9,3.10).

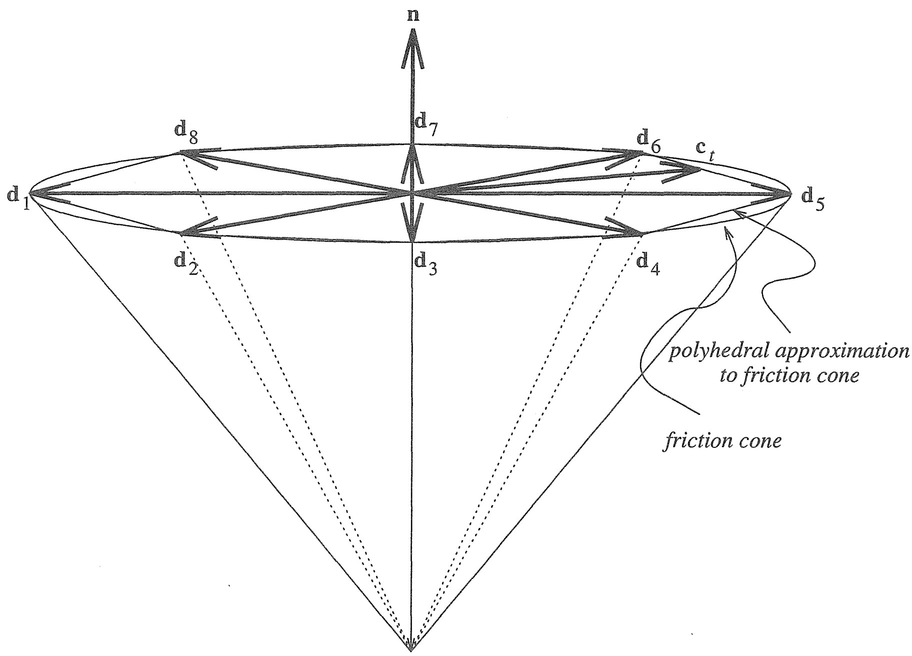


FIGURE 2. Polyhedral approximation to the friction cone

The resulting problem is an LCP of the form:

$$0 \leq \begin{bmatrix} n^T M^{-1} n & n^T M^{-1} D & 0 \\ D^T M^{-1} n & D^T M^{-1} D & e \\ \mu & -e^T & 0 \end{bmatrix} \begin{bmatrix} c_n \\ \beta \\ \lambda \end{bmatrix} + \begin{bmatrix} n^T b_1 \\ D^T b_2 \\ 0 \end{bmatrix} \perp \begin{bmatrix} c_n \\ \beta \\ \lambda \end{bmatrix} \geq 0.$$

The matrix in this LCP can be easily checked to be copositive since M is positive definite and symmetric, $\mu > 0$, and the e and $-e^T$ entries form an antisymmetric sub-matrix. Solutions exist for such LCP's (and can be found using Lemke's algorithm) since the constant term satisfies the conditions of [4, Cor. 4.4.12, p. 277].

For smooth friction cones, another approach uses the ψ function used in the previous section. Conditions (3.12,3.13) should be replaced by the discrete analogues of the continuous time conditions (3.5,3.6):

$$\begin{aligned} 0 &\in \mu D(q^l)^T v^{l+1} + \lambda^{l+1} \partial\psi(\beta^{l+1}), \\ 0 \leq \lambda^{l+1} &\perp \mu c_n^{l+1} - \psi(\beta^{l+1}) \geq 0. \end{aligned}$$

This leads to highly nonlinear complementarity problems. Nevertheless, such complementarity problems can be solved.

This discretization is a partly implicit Euler method. Therefore it can only give $O(h)$ accuracy at best. However, unlike conventional discretizations, it can handle impulsive forces — in particular, it can handle Painlevé's problem. Note that the complementarity condition $0 \leq f(q) \perp c_n \geq 0$ does not appear explicitly in (3.9–3.13); (3.11) is essentially the differentiated form of this condition. Using the differentiated constraint only can result in the true constraint "drifting" into the inadmissible region, which is an effect that has been noticed in relation to DAE formulations of rigid-body dynamics with bilateral (i.e., equality) constraints [1]. It is tempting to replace (3.11) with $0 \leq f(q^{l+1}) \perp c_n^{l+1} \geq 0$. This does not work: the resulting discretization behaves as if it had a "random" coefficient of restitution when impacts occur. (The effective coefficient of restitution depends on the time within the time-step $[t_l, t_{l+1}]$ that contact occurs.) Since (3.11) uses differentiated constraints, it may be occasionally advisable to project q^{l+1} back to the feasible region. This can be done without disturbing the time-stepping, since the time-stepping method is a one-step method.

The numerical formulation based on polygonal friction cones has been implemented and used to simulate a number of different systems. Typical simulation results are shown in Figure 3. This shows a thrown ball falling and then colliding with the first of three balls on a flat table. The coefficient of restitution used is $e = 0.9$, and the coefficient of friction is $\mu \approx 0.4$.

4. Measure differential inclusions and convergence

4.1. What are measure differential inclusions? To explain measure differential inclusions, we will first have a look at differential inclusions. Differential inclusions were originally introduced by A.F. Filippov in the early 1960's [9] as a means of "regularizing" ODE's with discontinuous right-hand sides. These differential inclusions have the form

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

where $F(x,t)$ is a set-valued function with the following properties:

1. The values $F(x,t)$ are closed, convex sets.
2. $F(\cdot, t)$ is an upper semi-continuous function $\mathbf{R}^n \to \mathcal{P}(\mathbf{R}^n)$; that is V is an open set containing $F(x,t)$, then there is an open set U containing x in $\mathbf{R}^n$ such that $F(y,t) \subset V$ for any $y \in U$.
3. $\|F(x,t)\| = \sup_{z \in F(x,t)} \|z\|$ is an L^1 function of t for any fixed x.

Then for any initial conditions $x(t_0) = x_0$, $dx/dt \in F(x,t)$ at least has local solutions, and provided $F(x,t)$ satisfies a "no blow-up" condition like $x^T F(x,t) \leq C(\|x\|^2 + 1)$, global solutions exist. These solutions are absolutely continuous functions $x(\cdot)$ where for any $t_1 < t_2$ in the solution domain,

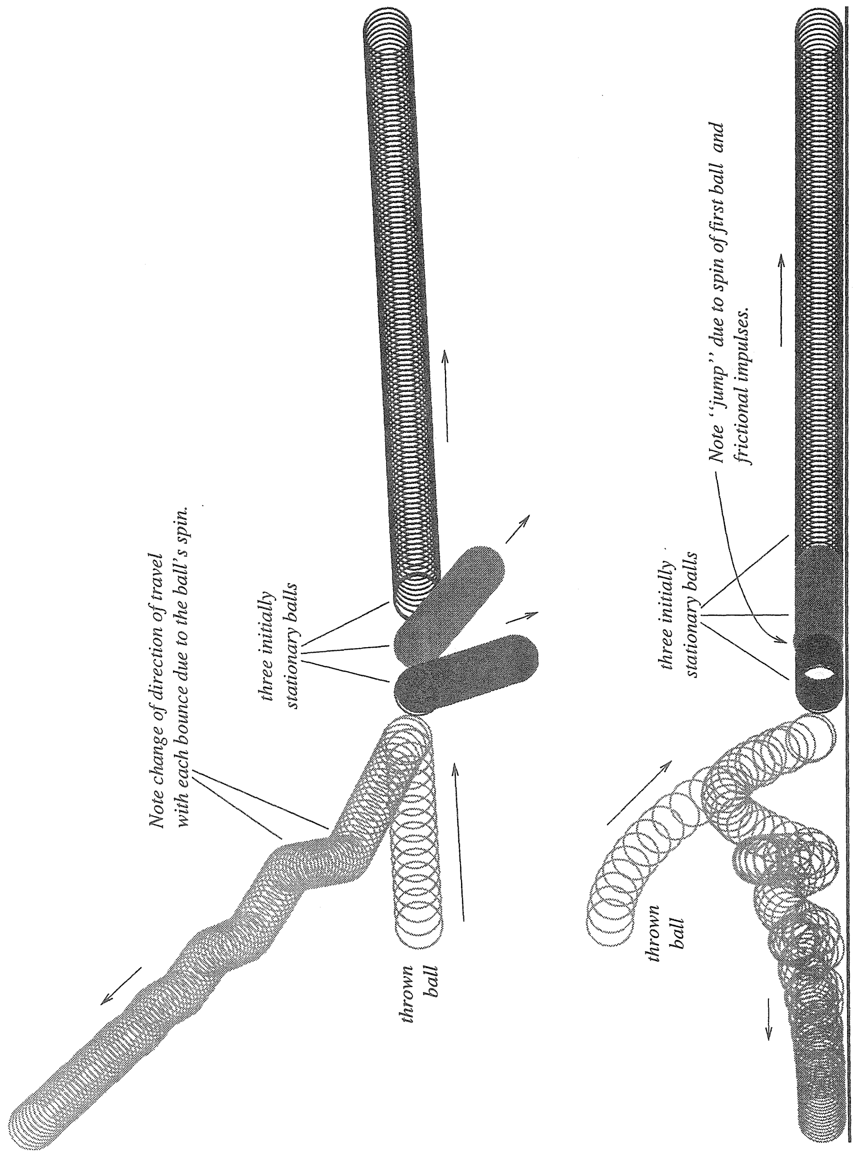
$$x(t_2) - x(t_1) \in \int_{t_1}^{t_2} F(x(\tau), \tau)\, d\tau = \left\{ \int_{t_1}^{t_2} f(\tau)\, d\tau \mid f(\tau) \in F(x(\tau), \tau) \forall \tau;\ f \in L^1 \right\}.$$

This is equivalent to requiring that $x(\cdot)$ is absolutely continuous, and $dx/dt(t) \in F(x(t), t)$ for almost all t.

We consider measure differential equations of the form

$$\frac{dx}{dt} \in F(y(t), t)$$

where $y(t)$ is an auxiliary function, perhaps defined through another differential equation such as $dy/dt = g(x, y, t)$.



(a) Plan view

(b) Elevation view

FIGURE 3. Elevation and plan views of "billiards" problem with partly elastic impacts

Measure differential inclusions allow for unbounded $F(y,t)$; condition (2) should be replaced by the condition that $F(\cdot,t)$ has a closed graph (that is, $z_k \in F(y_k,t)$ and $y_k \to y$, $z_k \to z$ implies that $z \in F(y,t)$); condition (3) can be dropped altogether. However, the solutions are not necessarily absolutely continuous: they can be functions of bounded variation. Then dx/dt is no longer a function, but a measure. Jumps in $x(t)$ give impulses (or Dirac-δ functions) in dx/dt. Since we can no longer properly talk about the values $dx/dt(t)$ even almost everywhere.

Measure differential equations first appear in the literature in the work of Schatzman [29], but were first called "measure differential inclusions" by Moreau in connection with "sweeping processes" [18, 19]. The most general formulations of measure differential inclusions can be found in [31, 32]. In these most general formulations there are strong and weak formulations. The strong formulation is based on the Lebesgue decomposition of measures, and Radon–Nikodym derivatives, and most closely resembles the "$dx/dt(t) \in F(x(t),t)$ for a.a. t" condition for ordinary differential inclusions. The other is based on integrals and resembles the integral formulation of ordinary differential inclusions.

If λ_0 is the Lebesgue measure, then we can decompose the measure $dx/dt = \xi\,\lambda_0 + \zeta$ where ξ is a function in L^1 and ζ is a singular measure. The strong formulation requires that $\xi(t) \in F(y(t),t)$ for Lebesgue almost all t, and $d\zeta/d|\zeta|(t) \in F(y(t),t)_\infty$ where L_∞ is the *asymptotic* or *recession cone* of a convex set L. The simplest definition of the recession cone is

$$L_\infty = \left\{ \lim_{k\to\infty} t_k x_k \mid t_k \downarrow 0,\, x_k \in L \right\}.$$

L_∞ corresponds to the directions in L "at infinity"; the condition $d\zeta/d|\zeta|(t) \in F(y(t),t)_\infty$ says that the unbounded and impulsive forces must be in directions lying in $F(y(t),t)$ "at infinity".

The weak formulation is much better for proving convergence results: for every continuous ϕ, non-negative and not everywhere zero,

$$\frac{\int \phi(t)\,\nu(dt)}{\int \phi(t)\,dt} \in \overline{\mathrm{co}} \bigcup_{\tau:\phi(\tau)\neq 0} K(\tau). \tag{4.1}$$

4.2. What kind of convergence? From functional analysis we learn that it can be quite important what *kind* of convergence we can obtain. Since the numerically computed forces c_n^h and β^h are sums of δ-functions, we need to work in the space of bounded Borel measures. Unfortunately, the usual topology of this space is difficult to work with numerically for these kinds of problems. To illustrate, suppose that we have some numerical procedure for investigating a single bounce of a rigid ball with a rigid table subject to some external forces (like air resistance, for instance). Then for a given step-size $h > 0$ it will give us a certain time t_h of bounce and a certain strength of the impulse s_h. Now it is clearly reasonable to expect that $t_h \to t^*$ and $s_h \to s^*$ as $h \downarrow 0$ for some values t^* and s^*. The impulses computed numerically are then $s_h\delta(t-t_h)$ and the "limiting" impulse is $s^*\delta(t-t^*)$. However, unless t_h is exactly equal to t^*, $\|s_h\delta(\cdot - t_h) - s^*\delta(\cdot - t^*)\| = \int |s_h\delta(t-t_h) - s^*\delta(t-t^*)|\,dt = |s_h| + |s^*|$ which does not (in general) go to zero. So we cannot use the usual strong (normwise) convergence in the space of measures. Furthermore, we do not even get convergence in the weak topology for bounded Borel measures. We finally get convergence in the weak* topology of bounded Borel measures.

A sequence of measures ν_h converges weak* to ν means that for all continuous functions ϕ, $\int \phi(t)\, d\nu_h(t) \to \int \phi(t)\, d\nu(t)$ as $h \downarrow 0$. There are a number of classical theorems of functional and classical analysis that give us tools to formulate convergence theorems. First we note the Reisz representation theorem for continuous functions: every linear functional on the space of continuous functions on $[0, T]$ is precisely the space of bounded Borel measures on $[0, T]$, which can in tern be represented by Riemann–Stieltjes integrals: $l(\phi) = \int \phi\, d\nu = \int \phi(t)\, dg(t)$ where the final integral is a Riemann–Stieltjes integral with a function g of bounded variation. Then there is Helly's theorem, which is stated in terms of functions of bounded variation: if g_h is a family of functions with a uniform bound on their variation and values: $\bigvee g_h \leq M$, and $|g_h(t)| \leq M$ for all $h > 0$ and t, then there is a pointwise limit $g(t)$ for all but countably many t, and $\bigvee g \leq M$. The most general result that we will find useful here is Alaoglu's theorem, which states that any closed, bounded, convex set in a dual Banach space is (sequentially) compact in the weak* topology. A good reference for functional analysis is [16].

4.3. The main result The strongest result that has been obtained for the above numerical schemes is as follows:

Theorem 1 *We assume the following (H1)–(H6):*

- *(H1) The functions $M(q)$, $n(q)$, $D(q)$ and $V(q)$ are all smooth and globally Lipschitz continuous functions of q, with Lipschitz constants L_M, L_n, L_D and L_V respectively.*
- *(H2) The matrices $M(q)$ is uniformly positive definite; that is, $\lambda_{min}(M(q)) \geq \beta_M > 0$ for some β_M, for all q.*
- *(H3) The functions $M(q)$, $n(q)$, and $D(q)$ are all uniformly bounded in the 2-norm by, B_M, B_n and B_D respectively.*
- *(H4) The singular values of $n(q)$ are all bounded away from zero: $\|n(q)\| \geq \beta_n > 0$ for all q.*
- *(H5) The friction cone $FC(q)$ is pointed for all q.*
- *(H6) For each i there is a k such that $d_i^{(j)}(q) = -d_k^{(j)}(q)$.*

Then there is a subsequence $h_k \downarrow 0$ where

$$\begin{array}{lll} q^{h_k}(\cdot) & \to q(\cdot) & \textit{uniformly}, \\ v^{h_k}(\cdot) & \to v(\cdot) & \textit{pointwise almost everywhere}, \\ dv^{h_k}(\cdot) & \to dv(\cdot) & \textit{weak} * \textit{ as Borel measures}, \end{array} \tag{4.2}$$

on $[0, T]$, and every such subsequence converges to a solution $(q(\cdot), v(\cdot))$ of the measure differential inclusion . Every such limit is exactly dissipative in that the total energy $\frac{1}{2}v^T M(q)v + V(q)$ is non-increasing. In the one-contact case the impacts are inelastic in the sense that $n(q)^T v^+ = 0$ whenever there is contact. Also, in the one-contact case, the approximate Coulomb friction law (for the polygonal approximation to the friction cone)

$$\frac{d\widetilde{\beta}}{dc_n}^T \widetilde{D}^T v^+ = -\mu \|\widetilde{D}^T v^+\|_\infty$$

holds for c_n-almost every point where v is continuous. The approximate Coulomb law also holds everywhere *in the limit for the one-contact case if (1) $n(q)^T M(q)^{-1} z > 0$ for all $0 \neq z \in \widehat{FC}(q)$, or (2) the friction planes are one-dimensional.*

While this result is a significant advance on previous work (such as [17]), there are still a number of areas where improvement is clearly desirable. One of these is the issue of multiple contacts. Even in the frictionless case it is not clear what are suitable laws are for multiple simultaneous impacts. Another is the requirement that the friction planes are one-dimensional. On the other hand, this result not only shows the convergence of a numerical scheme, but also shows the *existence* of a solution to rigid-body dynamics that was not previously shown.

5. Open questions

There are a great many open questions remaining about rigid-body dynamics and how to simulate it. Here we give just a few of the issues that deserve study.

5.1. Velocity-dependent friction coefficients Velocity-dependent friction coefficients ($\mu = \mu(v)$ or $\mu = \mu(\|v\|)$) are commonly observed in practice, but present some significant theoretical challenges. There are two main models of velocity-dependent friction coefficients: the two-coefficient model where $\mu = \mu_{dynamic}$ if $v \neq 0$ and $\mu = \mu_{static}$ if $v = 0$; there is also the continuous model where $\mu = \mu(\|v\|)$ where $\mu(s)$ is a decreasing positive function in s which varies smoothly from $\mu(0)$ to $\mu_\infty = \lim_{s\to\infty} \mu(s) > 0$. The former introduces a (probably spurious) discontinuity into the formulation which makes the theory substantially harder. The latter is much more reasonable. Experimental results could confuse the two since the region where $s > 0$ and $\mu'(s) < 0$ is very often dynamically unstable. Thus the transition from $v = 0$ (no slip) to large $\|v\|$ can happen quickly giving the impression that there are just two coefficients of friction $\mu(0)$ and μ_∞.

5.2. Partly elastic impacts Partly elastic impacts have not been dealt with here. Important work on this in the frictionless case has been done by Paoli and Schatzman [27, 26, 25] and by Mabrouk [15, 14]. Both groups have based their work on the study of particles and have focussed on Newtonian impact laws $n^T v^+ = -e\, n^T v^-$ where e is the coefficient of restitution. However, the a priori assumption that e is a constant is clearly not true, even approximately. See, for example, the experimental and simulation results of Stoianovici and Hurmuzlu [33]. Another critique of conventional impact models is given by Chatterjee [2]. What is needed to resolve this situation is a way of incorporating elastic vibrations into rigid-body dynamics.

5.3. Multiple simultaneous contacts Multiple simultaneous impacts is another area which needs development. Even giving theoretically justifiable impact laws for multiple simultaneous impacts *without friction* is a difficult task. As noted in the previous subsection, resolving this issue will require incorporating elastic vibrations into "rigid" body dynamics. This is an area which will probably see considerable development in the coming years.

References

[1] Campbell, S. L., and Leimkuhler, B. Differentiation of constraints in differential-algebraic equations. *Mech. Structures Mach. 19*, 1, 1991, 19–39.

[2] Chatterjee, A. On the realism of complementarity conditions in rigid-body collisions. *Nonlinear Dynamics 20*, 2, 1999, 159–168.

[3] Clarke, F. H. *Optimization and Nonsmooth Analysis.* Classics in Appl. Math. #5. SIAM, Philadelphia, PA, 1990. 1st edition published J. Wiley and Sons, 1983.

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

[5] de Coulomb, C. A. Théorie des machines simples, en ayant égard au frottement de leurs parties, et la roideur des cordages. Piéce qui a reporté le Prix double de l'Académie des Sciences pour l'année 1781. *Mémoires des Savans Etrangers **X***, 1785, 163–332. Réimprimé 1809 chez Bachelier, Paris.

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

[7] Duvaut, G., and Lions, J.-L. *Inequalities in Mechanics and Physics*, vol. 219 of *Grundlehren der mathematischen Wissenschaften.* Springer Verlag, Berlin, Heidelberg, New York, 1976. Orig. in French: "Les inéquations en mécanique et en physique" (1972).

[8] Ferris, M. C., and Pang, J.-S. Engineering and economic applications of complementarity problems. *SIAM Review* ***39***, 4, 1997, 669–713.

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

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

[11] Jarušek, J., and Eck, C. Dynamic contact problems with small Coulomb friction for viscoelastic bodies. Existence of solutions. *Math. Models Methods Appl. Sci.* ***9***, 1, 1999, 11–34.

[12] Lee, E. B., and Markus, L. *Foundations of Optimal Control.* Appl. Math. Ser. J. Wiley and Sons, 1967.

[13] Lötstedt, P. Time-dependent contact problems in rigid-body mechanics. *Math. Prog. Study* ***17***, 1982, 103–110.

[14] Mabrouk, M. Liaisons unilatérales et chocs élastiques quelconques: un résultat d'existence. *Comptes Rendus Acad. Sci., Sér. I* ***326***, 1998, 1353–1357.

[15] Mabrouk, M. A unified variational model for the dynamics of perfect unilateral constraints. *Eur. J. Mech. A Solids* ***17***, 5, 1998, 819–842.

[16] Megginson, R. E. *An Introduction to Banach Space Theory*, vol. 183 of *Graduate Texts in Mathematics.* Springer, Berlin, New York, 1998.

[17] Monteiro Marques, M. D. P. *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.

[18] Moreau, J.-J. Sur les mesures différentielles de fonctions vectorielles et certain problèmes d'évolution. *Comptes Rendus Acad. Sci. Paris, Sér. A* ***282***, 1976, 837–840.

[19] Moreau, J.-J. Evolution problem associated with a moving convex set in a Hilbert space. *J. Differential Equations* ***26***, 3, 1977, 347–374.

[20] Moreau, J.-J. Application of convex analysis to some problems of dry friction. in *Trends in applications of pure mathematics to mechanics, Vol. II (Second Sympos., Kozubnik, 1977)*, vol. 5 of *Monographs Stud. Math.* Pitman, Boston, Mass., 1979, pp. 263–280.

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

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

[23] Pang, J.-S., and Stewart, D. E. A unified approach to frictional contact problems. *Internat. J. Eng. Science* ***37***, 13, 1999, 1747–1768.

[24] Pang, J.-S., Trinkle, J. C., and Lo, G. A complementarity approach to a quasistatic rigid body motion problem. *J. Computational Optimization and Applications* ***5***, 2, 1996, 139–154.

[25] Paoli, L., and Schatzman, M. Approximation et existence en vibro-impact. *C. R. Acad. Sci. Paris, Sér. I* ***329***, 1999, 1103–1107.

[26] Paoli, L., and Schatzmann, M. Mouvement à un nombre fini de degrés de liberté avec contraintes unilatérales: cas avec perte d'énergie. *Math. Model. Numer. Anal.* ***27***, 6, 1993, 673–717.

[27] Paoli, L., and Schatzmann, M. Vibrations avec contraintes unilatérales et perte d'énergie aux impacts, en dimension finie. *Comptes Rendus Acad. Sci. Paris, Sér. I* ***317***, 1993, 97–101.

[28] Pfeiffer, F., and Glocker, C. *Multibody Dynamics with Unilateral Contacts.* Wiley Series in Nonlinear Science. J. Wiley and Sons, New York, 1996.

[29] Schatzman, M. Sur une classe de problèmes hyperboliques non linéaires. *C. R. Acad. Sci. Paris Sér. A-B* ***277***, 1973, A671–A674.

[30] Stewart, D. E. Convergence of a time-stepping scheme for rigid body dynamics and resolution of Painlevé's problem. *Archive Rational Mech. Anal.* ***145***, 1998, 215–260.

[31] Stewart, D. E. Reformulations of measure differential inclusions and their closed graph property. Tech. Rep. Comp. Math. Tech. Rep. 113, University of Iowa, Department of Mathematics, October 1998. Available via `http://www.math.uiowa.edu/faculty/comp-math-1998.htm`.

[32] Stewart, D. E. Formulations of measure differential inclusions in infinite dimensions. *J. Set-Valued Analysis*, February 2000. Accepted.

[33] Stoianovici, D., and Hurmuzlu, Y. A critical study of the applicability of rigid-body collision theory. *ASME J. Applied Mech.* ***63***, 2, 1996, 307–316.

[34] Stronge, W. J. Rigid Body Collisions with Friction. *Proc. Royal Society of London* ***A431***, 1990, 168–181.