A New Algebraic Rigid Body Collision Law Based On Impulse Space Considerations

Anindya Chatterjee[1]
anindya@crash.esm.psu.edu

Andy Ruina[2]
ruina@cornell.edu

Submitted to Journal of Applied Mechanics, July 1997
Accepted, subject to revisions, April 11, 1998
Revised and resubmitted, July 7, 1998
Finally accepted, July 24, 1998
Last edited August 6, 1998

[1]Dept. of Engineering Science and Mechanics, Penn State University
[2]Dept. of Theoretical and Applied Mechanics, Cornell University

Abstract

We present a geometric representation of the set of 3D rigid body collisional impulses that are reasonably permissible by the combination of non-negative post-collision separation rate, non-negative collisional compression impulse, non-negative energy dissipation and the Coulomb friction inequality. The construction is presented for a variety of special collisional situations involving special symmetry or extremes in the mass distribution, the friction coefficient, or the initial conditions. We review a variety of known friction laws and show how they do and do not fit in the permissible region in impulse space as well as comment on other attributes of these laws. We present a few parameterizations of the full permissible region of impulse space. We present a simple generalization to arbitrary 3D point contact collisions of a simple law previously only applicable to objects with contact-inertia eigenvectors aligned with the surface normal and initial relative tangent velocity component (e.g. spheres and disks). This new algebraic collision law has two restitution parameters for general 3D frictional single-point rigid-body collisions. The new law generates a collisional impulse that is a weighted sum of the impulses from a frictionless but non-rebounding collision and from a perfectly sticking, non-rebounding collision. We describe useful properties of our law; show geometrically the set of impulses it can predict for several collisional situations; and compare it with existing laws. For simultaneous collisions we propose that the new algebraic law be used by recursively breaking these collisions into a sequence ordered by the normal approach velocities of potential contact pairs.

1 Introduction

Rigid body simulations involving collisional contact necessarily use some kind of contact constitutive assumptions. Such collisional contact constitutive assumptions fall into three main classes: *full-deformation* laws; *incremental* laws (also called "soft"); and *algebraic* laws (also called "hard" – see Goyal *et al.*, 1994; Walton, 1992). Full-deformation laws are based on continuum mechanics equations for the whole body, most often linear elasto-dynamic equations. Incremental laws involve some kind of micro-mechanical model of the deformation and slip in the contact region, with the assumption that the motion of the bulk of the bodies is described by rigid-body differential equations of motion during the collision. Algebraic laws attempt to describe the *net* collisional interaction by algebraic equations that relate pre-collision and post-collision quantities. Thus, algebraic laws are generally based on less stringent rigidity requirements than incremental laws (see e.g., Chatterjee and Ruina, 1998a). One interpretation of incremental calculations is that they are a means to calculate a reasonable net impulse and not a literal description of the true contact mechanics. In this sense incremental models are equivalent to algebraic laws where function evaluations involve the solution of differential equations.

Unfortunately no known collision law of any kind is both predictive and accurate, in the sense, say, that Newton's laws, or even linear elasticity or the Navier-Stokes equations can be accurate. For a given pair of bodies, the collisional outcome may well depend on not just the initial velocities, masses, and mass moments of inertia, but also on some combination of contact shape, contact-mechanics, surface chemistry, friction, fracture, and vibration phenomena that are not well understood, especially *a priori*. Further, in the case of multiple superficially-simultaneous collisions, the predicton of the collisional outcome often depends so sensitively on initial conditions that sufficiently accurate initial conditions cannot be expected to be known by a simulator. But even if all modeling ingredients are accurately known, which they generally are not, there are technical difficulties that prevent simple calculation of collisional outcomes. Partial differential equations involving bulk and/or contact deformation are time consuming to solve (and to set up to solve). There is no reason to believe that, in general, an accurate continuum model can be well approximated by treating the body as rigid everywhere except in a localized quasistatic region describable by ordinary differential equations (as demanded by incremental laws). Finally, there is no reason to expect that the outcome of detailed modeling or exhaustive experimentation has a tractable summarizing description with standard functions or even lookup tables that apply equally well to a wide variety of bodies and their collisions (as is demanded by algebraic collision laws).

Any generally applicable collision law, whether coming from detailed continuum modeling, approximating ordinary differential equations, or summarizing functions will be highly approximate unless applied to a narrow range of collisional situations.

But simulators will simulate collisions with or without an accurate constitutive law. What outcomes should and should not be reasonably allowed in such simulations, and what basic phenomena may one hope to easily capture in such simulations? How can one compare various available collision laws? Can better collision laws be constructed? To help answer these questions it is perhaps useful to describe algebraically and geometrically the basic restrictions that collisional interaction rules might be reasonably expected to obey, as we do in Section 3 of this paper. In Section 4 we describe how several presently available collision laws fit in to this scheme. In Section 5 we present a new simple algebraic collision law which is a crude description of reality at best, but which overcomes some of the shortcomings in other known laws. And finally in Section 6 we describe a way to use the new law to predict a plausible outcome for multiple simultaneous collisions.

2 Preliminaries

The common basic assumptions of single point contact rigid-body collision theory are that the collisional interaction involves forces that are large compared to body forces and centripetal inertial terms; that the contact forces occur in a region of spatial extent negligibly small compared to body dimensions; and that at times before and soon after the collision velocities of points on the bodies can be accurately calculated using rigid-body relations (using the same positions and orientations before and after the collision). We do not consider here the possibility of nonnegligible contact moments about the contact point[1].

Based on the foregoing assumptions, the net collisional interaction between an arbitrary pair of objects colliding in a known configuration must satisfy impulse-momentum relations of the form

$$\mathbf{P} = \mathbf{M} \cdot \Delta\mathbf{V} = \mathbf{M} \cdot (\mathbf{V}_f - \mathbf{V}_i), \tag{1}$$

where $\mathbf{P}$ is the collisional contact impulse transmitted from one object to the other (we arbitrarily pick one object as the reference object), $\mathbf{V}$ is the velocity of the contact point of the second object relative to the contact point of the reference object, $\Delta\mathbf{V}$ is the change in $\mathbf{V}$ from its pre-collision value of $\mathbf{V}_i$ to its post-collision value of $\mathbf{V}_f$, and $\mathbf{M}$ is a second order rank 3 tensor with units of mass, which depends on the mass distributions of both colliding objects. All symmetric positive definite $\mathbf{M}$ have physical realizations in terms of pairs of unconstrained objects of finite mass (Chatterjee and Ruina, 1998b). $\mathbf{M}$ remains rank 3 for typical single point collisions of linkages consisting of multiple interconnected objects with frictionless constraints/bearings. Given a mechanism's configuration and mass distribution at impending contact $\mathbf{M}$ can be found using standard (non-collisional) rigid body dynamics (see appendix A.1). Derivations of Eq. 1 are outlined in many references (for discussions, see e.g., Smith, 1991; or Chatterjee, 1997). Equation 1 incorporates linear and angular momentum balance; these principles can provide no more information. Either $\mathbf{P}$ or $\Delta\mathbf{V}$ describes the effect of the collision on the rigid body system. For example, if $\mathbf{P}$ is known, then the motion of the system is found by applying $\mathbf{P}$ and $-\mathbf{P}$ to the two contacting bodies[2]. All algebraic rigid body collision laws start from Eq. 1 (or relations equivalent to it; e.g., see Brach, 1991), and then hypothesize three additional conditions (two in 2D) that vary from law to law so that Eq. (1) along with the collision law determine the six unknowns (four in 2D) $\mathbf{P}$ and $\Delta\mathbf{V}$. Incremental laws start with the differential form of Eq. 1 (thus requiring stronger rigidity, see Chatterjee and Ruina, 1998a), and then hypothesize specific types of incremental contact behavior that vary from law to law; integration of the resulting differential equations up to a hypothesized termination criterion then determines the outcome of the collision.

When needed, we assume that at least one of the objects is geometrically smooth at the contact point so a well defined tangent plane exists with normal $\mathbf{n}$ directed out from the reference object in the n or 1 direction. For definiteness, we take the first tangential or 2-direction to oppose the tangential component of pre-collision contact point relative velocity. The other tangential or 3-direction is orthogonal to the 1- and 2-directions. We use this coordinate system for discussion and graphics. Except when noted, the formulas we present have coordinate-free interpretations. In coordinate form Eq. 1 is

$$P = M(V_f - V_i), \tag{2}$$

where the 3×1 column vectors P, V_i and V_f, and the 3×3 local mass matrix M are the component forms of $\mathbf{P}$, $\mathbf{V}_i$, $\mathbf{V}_f$ and $\mathbf{M}$. Further assumptions about the collisional impulse are described below.

[1] See Brach (1991) for an example of a nongeneric collision where the contact moment is qualitatively important. Chatterjee and Ruina (1998a) discuss why the contact moment is likely to be unimportant for most collisions.

[2] Special treatment is sometimes required for systems with special constraints such that an eigenvalue of $\mathbf{M}$ is infinite and $\mathbf{P}$ cannot be determined from $\Delta\mathbf{V}$ alone; or bodies with all mass on a straight line so that $\mathbf{M}$ is singular and $\Delta\mathbf{V}$ cannot be determined from $\mathbf{P}$ alone.

3 The permissible region in impulse space

The following description of permissible collision impulses contains few new ideas, but is absent from much of the published collision literature (including texts and monographs like Goldsmith, 1960; Brach, 1991). Brogliato (1996) discusses the permissible region in the space of collision parameters. The geometric interpretation (in the space of the collision impulse) that we present below motivates the new collision law, and complements Rubin's recent paper (1998) as well as Brogliato's treatment (both of which attempt to describe the constraints on the collisional interaction in terms of mathematical expressions without the equivalent, explicit geometric description in impulse space).

We now describe various points, lines and surfaces in impulse space as discussed in part by many (e.g., Goldsmith (1960), Wang and Mason (1992) and Smith (1991)) but perhaps first brought together in Chatterjee (1997).

An equivalent discussion could be presented in contact-relative-velocity space, but the description of friction is not so convenient there.

3.1 Points, lines, and surfaces in impulse space

See Figs. 1 and 2. Each point in impulse space $P = (P_1, P_2, P_3)$ corresponds to a candidate outcome for a collision with given M and V_i.

The origin $P = 0$ of impulse space corresponds to no collisional interaction and $V_f = V_i$. The normal impulse which brings the normal velocity to zero (perfectly plastic and frictionless) is

$$P_I = -\left(\frac{n^T V_i}{n^T M^{-1} n}\right) n \tag{3}$$

which in our chosen coordinates is $(-V_{i1}/[M^{-1}]_{11}, 0, 0)$. The impulse required to bring the contact points to a stop relative to each other with no bounce and no slip (perfectly plastic and perfectly sticking), is

$$P_{II} = -MV_i. \tag{4}$$

For any impulse P the relative velocity of the contact points satisfies $P - P_{II} = MV_f$ and we can think of P_{II} as a reference point for impulses, corresponding to the reference state of zero relative contact velocity.

A contact kinetic energy E_c can be defined by

$$E_c \equiv \frac{1}{2} V_f^T M V_f = \frac{1}{2}(P - P_{II})^T M^{-1}(P - P_{II}). \tag{5}$$

The contact kinetic energy E_c is not a simple sum of energies of the separate bodies. However, changes in E_c due to contact forces are the same as changes in the system kinetic energy. E_c is a natural quantity to consider in collisional energy calculations since it is frame-invariant, unlike the total kinetic energy which is not.

Constant E_c defines a surface which facilitates the geometric description of many collision features (Chatterjee, 1997). Each such E_c surface is ellipsoidal with orientation coinciding with the principal directions of M, and corresponding diameters in inverse proportion to the square roots of the eigenvalues of M. The size and position of the E_c surfaces depends on initial conditions. The center C of each E_c ellipsoid is at P_{II}. An impulse of $P = P_{II}$ leads to zero final E_c and hence the maximum possible collisional dissipation. An impulse of $P = 2P_{II}$ conserves energy and leads to a direct reversal of the contact velocity ($V_f = -V_i$). For a given impulse P the associated contact velocity V is normal to the

E_c surface on which P lies. The E_c surface associated with the original motion (or pre-collision energy) passes through the origin with a normal in the direction of V_i. A collision with no final tangential velocity ($V_f \parallel n$) and a given energy corresponds to one where $P^T n$ is maximum on the corresponding E_c surface. The set of all collisions with $V_f \parallel n$ makes up the *line of sticking* (Goldsmith, 1960; Wang and Mason, 1992) which passes through P_{II} and passes through the E_c surfaces where their normals are parallel to n.

For given E_c the set of impulses with no normal separation velocity ($V_f^T n = 0$) is the ellipse on the ellipsoidal E_c surface where the surface normal is perpendicular to n (places where the surface tangent planes contain n). Over the full range of possible values of E_c this makes up the *plane of maximum compression*[3] (see e.g., Routh, 1897; Goldsmith, 1960; or Wang and Mason, 1992) given by

$$n^T(M^{-1}P + V_i) = 0,$$

which cuts the E_c surfaces in half and includes both P_{II} (at C) and P_I at A. A frictionless collision with kinematic restitution coefficient $e = 1$ has the impulse $2P_I$ and conserves energy. The energy conserving frictionless collision $2P_I$ kills (using P_I) and negates (using P_I again) the normal approach velocity. Thus, A bisects OB.

The dashed line (or plane) in Fig. 1, drawn through point D and parallel to the plane of maximum compression, marks impulses for which the observed kinematic or Newtonian normal restitution e will be some constant. A specific value of e constrains the impulse transmitted to lie on a plane (a line in 2D) parallel to the plane of maximum compression (for which $e = 0$). In Fig. 1 point D is on an $e < 1$ plane but lies outside the initial energy surface and thus corresponds to a net increase in kinetic energy. This construction provides an indication of energy conservation problems with naive algebraic collision laws, such as in Whittaker (1944) or Kane and Levinson (1985).

An identically shaped and oriented family of ellipsoidal surfaces can be drawn in contact force space as level surfaces of $F^T M^{-1} F$. Assuming "force-response rigidity" (Chatterjee and Ruina, 1998a) we can use $F = MA$ (the differential form of Eq. 2). F and A will not be collinear unless they happen to be along an eigenvector of M, which may not be along the normal or tangential direction. The possible lack of alignment of the M eigenvectors with the contact normal and tangent plane is one source of difficulty in collision modeling since it corresponds to inertial coupling between normal and tangential motions and between tangential motions. This misalignment manifests itself by E_c surfaces being crooked with respect to the normal and contact tangent plane.

For every M and pre-collision velocity V_i, there is an E_c ellipsoidal surface in impulse space. Conversely, every ellipsoidal surface transversely intersecting the n or 1-axis at the origin corresponds to the collision of some pair of bodies. Thus, for impulse analysis, consideration of arbitrary rigid body mechanisms is equivalent to considering arbitrary ellipsoidal surfaces as described above.

3.2 Reasonable restrictions on the contact impulse

Here are some natural and common assumptions about collisions :

1. Non-negative dissipation of kinetic energy constrains the impulse to lie on or inside the ellipsoid of the initial E_c surface, so P must satisfy (see, e.g., Smith, 1991)

[3] Routh used the term "greatest compression" to denote the point *during* the collision when, as per his incremental model, the instantaneous normal component of the relative velocity at the contact point becomes zero. That point depends on the path in impulse space followed by the collision. If, say from experiments, only the *net* collision outcome is known, the point of maximum compression cannot be determined. However, the set of all impulses P for which $V_f^T n = 0$ is well-defined, and called the plane of maximum compression.

$$(P - P_{II})^T M^{-1} (P - P_{II}) \leq V_i^T M V_i. \tag{6}$$

2. Non-interpenetration in the final collision velocity ($V^T n \geq 0$) constrains the impulse to lie "above" (in the normal direction) or on the plane of maximum compression (line of maximum compression, in 2D).

3. A non-negative normal impulse (no net tensional impulse) must lie in the upper half space (upper half plane in 2D): $P^T n \geq 0$.

4. If the contact force obeys the Coulomb friction inequality at all instants of time during the collision, then so must the net contact impulse. Thus P must lie in or on the cone given by $\|P - (n^T P)n\| \leq \mu n^T P$. In 2D collisions, this friction cone reduces to a pair of "friction lines".

For positive μ, condition (3) above follows from condition (4).

Conditions (1), (3) and (4) are not fundamental. Mechanical energy could be released by contact chemical reactions or phase transitions in the bodies. There is generally some (usually negligible) contact tension in part of the collision. Coulomb's law is only an approximation. But these restrictions might be reasonable limits for approximate general purpose collision laws. They restrict allowed impulses to a closed, bounded and convex *permissible* region in impulse space. This *permissible region* is the intersection of half an ellipsoid with a solid cone; for 2D collisions, it is a 2D slice of the corresponding 3D region.

In the absence of additional information about the contact mechanism[4], every point inside the permissible region represents a plausible collisional outcome. Although we do not expect every permissible impulse to be witnessed in experiments, we know no basic theoretical reason to further restrict the permissible region, nor do we know of sufficiently general experimental data to restrict it on empirical grounds.

All of the collision laws we discuss in this paper use dimensionless collision parameters such as the coefficient of normal restitution. In calculations with any *fixed* parameter values, the collision impulse is then automatically a homogeneous function of degree one in the pre-collision velocity. P is not simply a linear function of V_i. But, say, doubling V_i while keeping its direction fixed also doubles P while keeping its direction fixed. This velocity scaling is not fundamental, is not obeyed by some real collisions (e.g., spheres; see Goldsmith, 1960), but is apparently obeyed by some other collisions (e.g., vibration dominated interactions; see Stoianovici and Hurmuzlu, 1996).

3.3 2D geometric constructions

We now describe how the regions accessible to the different collision laws are represented graphically for 2D collisions. See Fig. 3a.

Pick a normal or 1-direction; and, perpendicular to it, the tangential or 2-direction.

Draw an ellipse whose shape represents a mass matrix M and size represents an initial contact kinetic energy $V_i^T M V_i/2$ (eigenvectors of M are aligned with the principal axes of the ellipse and the ratio of the eigenvalues of M is the square of the aspect ratio of the ellipse). Ellipses of any shape or orientation are allowed.

[4] An example of such special information would be if the collision is assumed to respect any special symmetry of M and V_i. The collisions of spheres with no spin component about the n axis would then be restricted to the n, V_i plane, a subset of the permissible region. Laws that do not explicitly add a symmetry breaking feature will automatically respect this symmetry, however. Examples include all of the algebraic laws discussed in this paper.

On the ellipse, locate the two points where the tangents are vertical (points A and B in the figure) and join these points by a straight line (the line of maximum compression). Locate the points where the tangents are horizontal (points C and D in the figure) and join them by a straight line (this is the line of sticking). The final velocity has a left or right component depending on whether the collision impulse is to the left or right of the line CD.

Pick the impulse origin point O to be on the arc ADB such that the normal to the ellipse is in the direction of the pre-collision contact velocity. Any point on arc ADB is allowed. If, by convention, the 2-direction is chosen to oppose V_i, then only the points on the portion AD are allowed.

Points A and B correspond to grazing incidence, while point D corresponds to normal incidence. The magnitude of V_i is not uniquely determined by the construction.

The intersection of the normal axis with the ellipse at E (or $2P_I$) represents a perfectly elastic, frictionless collision.

Draw friction lines from point O, at angles $\pm \tan^{-1} \mu$ from the normal direction.

The permissible region is now visible as the region between the friction lines, above line AB, and inside the ellipse.

Through point E, draw a line EF parallel to AB. This line represents kinematic restitution $e = 1$. All laws using kinematic restitution $0 \leq e \leq 1$ must predict an impulse somewhere between the lines AB and EF (possibly extended in both directions).

3.4 Special situations

For the purposes of comparing collision laws we consider various special collision situations. For these collisions, the curves accessible in impulse space to some well-known one-parameter collision laws, as well as the 2D regions accessible to the new collision law, are shown in Figs. 4 through 8. Details of these constructions will be discussed later in the paper. Below, we briefly comment on aspects relevant to the impulse-space description of this section.

A Generic Collision: The fairly generic collision of Fig. 4 approximately duplicates Fig. 1 but without the definitions of lines and points. Special cases of this collision appear below.

A Collision with Diagonal M: In contrast to the generic case of Figs. 1 and 4, consider a collision where the mass matrix is diagonal, as is the case for collisions of spheres and disks (see Fig. 5). The lines of maximum compression and sticking are aligned with the coordinate axes.

A Tangential Collision ($n^T V_i \to 0^-$): For grazing or tangential incidence (as shown in Fig. 6), the normal approach velocity is zero, and the energy ellipse is locally tangent to the 1-axis at the origin. If the line of maximum compression slopes appropriately and the friction coefficient is sufficiently high, a collision can occur. Laws based on kinematic restitution do not predict/capture finite normal rebound for grazing collisions.

A Collision with Infinite Friction ($\mu \to \infty$): For collisions where the coefficient of friction is effectively infinite (see Fig. 7), the friction lines are lined up with the tangential axis. Note that for the case shown in the figure, the center of the ellipse, point C, lies *below* the tangential axis. A perfectly dead, no-slip, no-bounce collision cannot occur in the case shown since it would require a net negative normal impulse component. Thus, $e = 0$, $\mu = \infty$ does not necessarily mean a dead collision, as some authors have assumed (e.g., Goldsmith, 1960).

A Collision with Unbounded M: Some specially constrained problems have mass matrices M which have one or more eigenvalues equal to infinity as for a pendulum striking a floor (see Fig. 8a). This problem has been examined in 3D as a special, "ideally" constrained (effectively two dimensional) problem, using Routh's model (e.g., Stronge, 1994). This problem is treated here as the limit of a sequence of problems where the eigenvalue in the direction along the pendulum rod approaches infinity (say with an arbitrarily large point mass replacing the hinge). As the eigenvalue approaches infinity,

the inital energy ellipse stretches in the corresponding direction. The portion of it contained inside a suitable finite-sized region approaches the parallel lines shown in Fig. 8b (the center of the ellipse is at point C).

3.5 Energy dissipation and $e \leq 1$

For all single-point *frictionless* rigid body collisions, it is well known that for normal restitution $e \leq 1$, nonnegative dissipation of kinetic energy occurs.

However, for frictional collisions, the *kinematic* coefficient of normal restitution can be substantially larger than one even in dissipative collisions. Smith and Liu (1992) have shown this using FEM simulations, and also observed this experimentally, for collisions where M is not diagonal.

M is diagonal for collisions of spheres and disks; such collisions are called "central collisions". Special symmetry-preserving contact deformation behavior can put further limitations on the energy dissipated in such a central collision. Typical experimental data for spheres and disks, too, shows restitution coefficients less than unity. Perhaps due to a combination of these reasons, one occasionally encounters the mistaken opinion that e must be less than unity for central collisions purely from nonnegative energy dissipation considerations. The impulse-space description presented above allows us to briefly address this issue.

Even for diagonal M, energy-dissipating collisions with $e > 1$ can occur, in principle. Figure 9 shows a nearly-grazing central collision (diagonal M), with reasonable friction, where the accessible region (gray) is significantly larger than the region allowed by $e \leq 1$ (black). In the figure, the maximum possible value for e is about 17 (it could be made larger). That such collisions are not often observed in practice is due to details of body-specific deformation mechanisms that are more restricted than the general cases allowed in this paper.

Smith and Liu (1992) discuss three well-known definitions of coefficients of normal restitution. Note that the *kinetic* coefficient of restitution e_k is the same as the kinematic coefficient, for diagonal M (central collisions). Thus, $e_k > 1$ does not preclude positive energy dissipation either. The same applies to the *energetic* restitution coefficient, defined in terms of the work done by the normal component of the contact force.

Finally it is also clear that even for collisions with diagonal M, for sufficiently high friction coefficient, it is possible to pick an impulse that corresponds to $e < 1$, lies inside the friction cone, and yet outside the maximum E_c ellipsoid. Typical collision laws do not usually pick such points for central collisions, but some naive laws (e.g., Kane and Levison's law, or Whittaker's law) can make such energy "creating" predictions for non-diagonal M, $e < 1$ and $\mu > 0$, as is pointed out in Kane and Levinson's text as well as in appropriate places in this paper.

Thus, for frictional collisions, there is no simple relation between nonnegative energy dissipation and the inequality $e \leq 1$.

4 Some collision laws

We now place the impulses predicted by several collision laws in (and out) of the permissible region and also discuss other strengths and weaknesses of these laws.

4.1 Collision laws that reach the entire permissible region

One candidate class of collision laws are laws that parameterize the entire permissible region defined by M, V_i, and μ, and no points outside. Here are 3 ways that this can be done, for example:

1. Arbitrary parameterization of a convex region. Pick a point P_0 inside the permissible region, say its centroid. Then, since the region is convex, all points can be defined by their spherical angles from P_0, say $0 \leq \psi_1 \leq \pi$, $0 \leq \psi_2 \leq 2\pi$, and the fraction of the distance from P_0 to the boundary, say $0 \leq \rho \leq 1$. Thus the outcome of all accessible collisions can be described with the triple (ψ_1, ψ_2, ρ), and no choice of these parameters (within their given bounds) picks a non-permissible point.

2. Projected impulses. Parameterize the outcome of the collision by the impulse itself, i.e., let the impulse components be the "collision parameters". If this impulse is outside the permissible region, project it back to the surface along lines leading to an arbitrary point P_0 inside the region (say the centroid). In other words, if P is outside the permissible region, diminish the magnitude of $P - P_0$ until the region boundary is reached (the energy surface, the friction cone, or the plane of maximum compression).

3. Brach's approach. Consider the following three-parameter law proposed by Brach (1991): a Newtonian or kinematic restitution parameter e for the normal direction, and two *impulse ratio* parameters μ_2 and μ_3 (not to be confused with the coefficient of friction) for the two tangential directions. If all values of $\mu_{2,3}$ were allowed while e was restricted to lie between 0 and 1, then this law would access an infinite slab of finite thickness in impulse space, parallel to the plane of maximum compression. The parameters $\mu_{2,3}$ are understood to be limited within bounds that respect physical considerations. The bounds on these collision parameters are expressed in terms of (configuration-dependent, implicit, nonlinear) inequality constraints expressing non-negative energy dissipation, in addition to friction limits. Brach's laws can capture all permissible impulses (if we allow $e > 1$).

The laws above that access the full permissible region have the merit of being capable of describing all permissible collisions. But we do not know any reason to expect that any of the parameters in these descriptions might be constant over a useful range of collisions (although Brach's proposed impulse-ratio parameter often stays roughly constant for some range of collisions of spheres and disks, if not too wide a range of initial approach angles is considered). The parameters in the above descriptions also do not have any special intuitive or theoretical interpretation.

The utility of a collision law comes not just from its ability to describe *a posteriori* any given collision once the outcome is known from other sources, but also its ability to predict the outcome of collisions for a range of initial velocities and orientations based either on much collisional data or on separately measurable parameters. The above laws seem lacking in this regard.

Even if not separately measurable and constant, a collision parameter such as a normal restitution coefficient retains some physical interpretation for arbitrary collisions: higher e roughly means that the objects will bounce more. The parameters in the first two laws given above, as well as Brach's impulse-ratio parameters, seem to lack this kind of simple interpretability.

As yet, no collision law we have seen accesses a large 3D subset of the permissible region in impulse space; has collision parameters that have simple physical interpretations; and has collision parameters that are at least roughly constant over some range of some collisions of some bodies for which some data is available.

4.2 Some algebraic collision laws

The permissible region is three dimensional for 3D collisions and two dimensional for 2D collisions. Any collision law that has only *one* free parameter (in addition to a friction inequality using μ) can access a 1D subset (curve) of the 3D permissible region. Similarly, any law with two independent free parameters (in addition to a friction inequality using μ) can access a 2D subset of the permissible region. Currently available algebraic laws, except for the full parameterizations of the permissible region as described above, access only a 1 or 2 dimensional region.

We now discuss some popular algebraic collision laws. For illustration we assume that initial motion

is to the left so the origin O is on arc AD. The impulses accessible to the laws below are shown in Figs. 4 through 8 for a variety of collisional situations.

4.2.1 Routh's law

Routh's (1897) law was re-popularized recently for 2D collisions by Wang and Mason (1992) and examined in some detail in 3D by Bhatt and Koechling (1995a,1995b). It has also been studied by many other authors. Routh's law assumes infinite tangential contact stiffness, any combination of possibly nonlinear springs and dashpots in the normal direction, and standard friction relations using μ. The neglect of tangential compliance allows expression of the incremental relations in terms of normal impulse P_n instead of time, totally independent of details of the normal compliance mechanism. Denoting tangential and normal components with t- and n-subscripts respectively, the tangential impulse P_t (a two-component vector in 3D) is incrementally calculated by

$$\begin{array}{rrcll} \textit{Either} & dP & = & M \cdot \{dV_n, 0, 0\}^T & \text{with} \quad V_t = 0 \quad \text{and} \quad \|dP_t\| \le \mu\, dP_n \quad \text{(no slip)} \\ \textit{or} & dP_t & = & -\mu(V_t/\|V_t\|)\, dP_n & \text{(slip or impending slip)} \end{array} \tag{7}$$

where Eqs. 1 and 2 apply to impulse increments throughout (and in cases of impending slip, the unit vector $V_t/\|V_t\|$ is understood to be the direction of impending slip). Routh's law is perhaps *the* classical incremental law. It may be regarded as algebraic in the sense that ODE solving (for given initial conditions) is function evaluation. Also, in 2D (where V_t has only one component) the governing differential equations can be solved in closed form. Routh's *kinetic* termination condition was

$$P_n \text{ at end of collision } = (1 + e_k) \times P_n \text{ at point of maximum compression}$$

after Poisson (1833)[5]. We loosely refer to all laws that use Routh's incremental contact relations Eq. 7, with any termination criterion, as "Routh's law".

Routh's law, in the absence of jamb-associated difficulties (Batlle and Cardona, 1998) can access a 1D curve that always lies in the permissible region for any $0 \le e_k \le 1$ and $\mu > 0$ (see Ivanov, 1992 for a proof of non-negative energy dissipation in 3D). It has the ability to predict finite *normal* rebound for tangential or grazing collisions (see e.g., Wang and Mason, 1992) even for $e_k \le 1$, unlike algebraic laws that use Newtonian or kinematic restitution, which predict zero normal rebound for finite restitution e. Routh's restrictive assumption of no tangential compliance naturally leads to inaccurate predictions as shown in Maw *et al.* (1981) and Smith and Liu (1992). In particular, Routh's law cannot predict nonzero *tangential* rebound for tangential or grazing collisions nor the slip reversal of a sphere seen in essentially all experiments for at least a small range of nearly normal collisions. The latter effect is easily visible with rubber "super" balls dropped while spinning about a horizontal axis (see e.g., Johnson, 1983; or Garwin, 1969).

See Fig. 3b. To graph Routh's accessible region for 2D collisions (see Wang and Mason, 1992) fix μ and track P as the normal impulse component P_N increases from 0 (at the start, P is at O, outside the parameter range and outside the permissible region). As P_N increases P moves up along the right friction line until it reaches the line of sticking. At the line of sticking the path has a kink. If the line of sticking makes a smaller angle with the normal than the friction lines do, then sticking is possible and the representative point climbs up along the line of sticking. Otherwise, it climbs up along a line parallel to the left friction line (as shown in the figure).

[5]Some ideas of energetic "consistency", and some ambiguities caused by 3D collisions with "jamb", have led to the definition of newer, work-based definitions of restitution (see Batlle and Cardona (1998) for discussion and references).

Depending on the aspect ratio and orientation of the ellipse, the angle of incidence (i.e., location of point O) and the friction coefficient, the intersection with the line of sticking might occur above the line of maximum compression, or even not at all.

The termination criterion, if we use Poisson or kinetic restitution, is that the normal impulse should be $(1+e_k)$ times the value it had when the representative point reached the line of maximum compression. That point is taken to be the collision impulse.

For grazing collisions Routh's law is the only one discussed in this paper that can predict a collision with non-zero post-collision separation. That is, only Routh's law predicts impulses above the plane of maximum compression for grazing collisions (see Fig. 6).

For infinite friction collisions (see Fig. 7) if the center of the initial energy ellipse happens to lie exactly on the tangential axis, then the region accessible to Routh's law shrinks to a point.

For a collision with unbounded M aligned as in Fig. 8, with large enough friction, Routh's law predicts only a dead collision (no-slip, no-bounce, $P = P_{II}$).

4.2.2 Whittaker's law

This classical law (see Whittaker, 1944) is stated more clearly in a 3D context in Kane and Levinson (1985), to whom it is often attributed. It uses a coefficient of normal restitution e (Newtonian or *kinematic* restitution) in addition to μ, and therefore accesses a 1D region in impulse space. By this law, either the collision terminates at a condition of sticking, or the impulse is on the friction cone and opposes the post-collision tangential velocity V_{ft}. That is,

$$V_f^T n = -(1+e)V_i^T n \qquad \text{and}$$

$$\text{Either } V_{ft} = 0 \text{ with } \|P_t\| \le \mu P^T n \quad \text{ or } P_t = -\mu \frac{V_{ft}}{\|V_{ft}\|} P^T n \,.$$

This law requires the solution of nonlinear algebraic equations in 3D, cannot predict spin reversal of spheres, cannot predict finite rebound for grazing collisions, and can predict arbitrarily large increases in kinetic energy in a collision for $e \le 1$ and $\mu > 0$ (see Fig. 8).

Whittaker's law is simplest to view geometrically by starting from the coefficient of kinematic restitution e, which specifies which line the collision terminates on (shown dashed in Fig. 3d). The intersection of the dashed line with the line of sticking is first located. This point is marked with a black circle in the figure. If the point is between the friction lines, then it is the collision impulse. If it is to the right of the right friction line, then the collision impulse is taken to be the intersection of the right friction line with the dashed line. Conversely, if the point is to the left of the left friction line, then the collision impulse is taken to be the intersection of the left friction line with the dashed line.

As can be seen in Figs. 3 or 4, the intersection of the line of sticking and the $e = 1$ line can easily be outside the initial energy ellipse. This is a direct demonstration of Kane's well known collision energy "paradox". This possibility of generating an increase in system kinetic energy persists for $e < 1$ and reasonable μ so long as the initial E_c surface is skewed as in the figures. The situation is extreme in the context of unbounded M as in Fig. 8 where this law fails completely in that it predicts an unbounded increase of kinetic energy and does not respect the hinge constraint in the limit as an increasing point mass at the bar's end behaves like a hinge.

For diagonal M the Whittaker-Kane-Levinson law is the same as Routh's law (see Fig. 5).

For grazing collisions this law can only predict an impulse of $P = P_{II}$ unless P_{II} is outside the friction cone in which case it predicts zero impulse.

4.2.3 Smith's law

This law (Smith, 1991) also has one parameter (kinematic restitution e) in addition to μ. Smith's law expresses the tangential impulse direction using a weighted average of pre- and post- collision tangential relative velocities. This average is an acknowledgement that the tangential impulse comes from frictional sliding in a variety of directions. Smith's law accesses a 1D curve in the permissible region, cannot predict increases in kinetic energy, can capture spin reversal of spheres, but cannot capture non-reversal of spin for positive μ. It cannot predict finite normal rebound for grazing or tangential collisions, and predicts a fixed amount of tangential rebound for such collisions (the amount of rebound is dependent on μ). Nonlinear algebraic equations need to be solved (solutions exist (Chatterjee, 1997), but uniqueness is unproven at this time).

In Smith's law, the equations describing the 3D collisional interaction are

$$P_n \begin{Bmatrix} 1 \\ -\mu \frac{V_{it}\,\|V_{it}\| + V_{ft}\,\|V_{ft}\|}{\|V_{it}\|^2 + \|V_{ft}\|^2} \end{Bmatrix} = M \begin{Bmatrix} -(1+e)V_{in} \\ V_{ft} - V_{it} \end{Bmatrix}, \tag{8}$$

where n and t subscripts denote normal and tangential components. This nonlinear system of equations needs to be solved for P_n and V_{ft}. One or more solutions with $P_n < 0$ frequently exist for 3D collisions with large enough values of μ, and these non-physical solutions must be rejected.

See Fig. 3c. We can graph the accessible region for Smith's law by parameterizing it with e starting at $e = -1$ (outside the accessible region and parameter range). For e close to -1, the impulse is small, $V_{ft}(e) < 0$, and the impulse ratio is μ. Thus, we draw a curve that climbs from the origin, up the right friction line.

As e increases, P eventually reaches the line of sticking. Now $V_{ft} = 0$. As e increases further, V_{ft} becomes positive, and so the impulse ratio decreases. However, from Eq. 8 it is easy to see that for small V_{ft} the change in the impulse ratio is *quadratic* in V_{ft}. Thus, there is no sudden change in direction (i.e., no kink) and the path of the representative point curves away from the right friction line.

How far and how fast it curves away can be seen from the following. In Fig. 3c, the heavy dashed line is drawn through point F (see Fig. 3a) and parallel to CD. Just as line EF represents impulses where the normal component of relative velocity is exactly reversed, this dashed line represents impulses such that the *tangential* velocity is exactly reversed. From Eq. 8, the impulse ratio is exactly zero when $V_{fT} = -V_{iT}$, i.e., the tangential relative velocity is exactly reversed. Therefore the curving path of the representative point reaches the normal axis at the point where it intersects the dashed line. Since, in the figure, that intersection occurs beyond point E, it corresponds to $e > 1$ and is not shown.

The collision terminates when the appropriate value of e is reached, which corresponds to reaching a line parallel to AB, somewhere between AB ($e = 0$) and EF ($e = 1$).

For grazing collisions Smith's law can predict only one impulse independent of e (dependent on μ) unless P_{II} is outside the friction cone, in which case Smith's law predicts a zero impulse.

4.2.4 Pfeiffer and Glocker's law, and extensions

This three parameter law (Pfeiffer and Glocker, 1996) is an algebraic law somewhat motivated by Routh's law, in that it assumes that the collisional interaction may be divided into well-defined and distinct compression and expansion phases (see, however, footnote 3). It uses kinetic restitution e_k for the normal direction, and models tangential restitution effects using two other parameters ν and e_t. For single point contact the law accesses a 2D subset of the permissible region. Unfortunately the law is restricted to 2D collisions (in its 1996 form) and does not seem to generalize to a simple 3D form with tangential restitution. If tangential restitution is not incorporated, this law may be generalized to a single-parameter 3D version (Anitescu and Potra, 1997).

The single-point versions with no tangential restitution, of these laws (Pfeiffer and Glocker, or Anitescu and Potra), are in some ways a combination of Routh's law and the Whittaker-Kane-Levinson law, avoiding the principal difficulties of both. By doing two separate algebraic calculations for compression and expansion, Routh's integration of nonlinear ODE's is avoided. Meanwhile, at the end of each stage, the friction condition assumed is that either there is no slip or the tangential impulse opposes final slip; this extension of the Whittaker-Kane-Levinson law to two stages ensures nonnegative energy dissipation. Without tangential restitution, these laws just have one collision parameter (normal restitution) and access a 1D subset of the permissible region.

The primary motivation behind the design of these laws is to pose simultaneous multiple collision problems as tractable linear complementarity problems whose solutions provide plausible collision impulses. However, as mentioned in Subsection 6.2 and discussed in Chatterjee (1998), the linear complementarity conditions are not fundamental properties of rigid bodies during collisions, and so there is no reason to prefer the predictions of these laws over those of other, possibly simpler laws.

4.2.5 A bilinear law for spheres and disks

Based on experiments, as well as an approximation of results obtained using the theoretical analysis of Mindlin and Deresiewicz (1951), this algebraic law is only applicable to the collisions of objects like spheres and disks where M has eigenvectors aligned with the coordinate axis in our preferred coordinate system. In other words, the law applies only to collisions where M is diagonal in our coordinate system.

The law is discussed in Brach (1991) and has been used by various researchers (see e.g., Walton, 1992; or Jenkins, 1992). First a normal impulse is calculated using a kinematic restitution coefficient (the same as kinetic restitution, for diagonal M). Then one first assumes that "sticking" occurs, and a candidate tangential impulse is calculated based on the tangential restitution parameter (in this description, the case of "sticking" includes tangential velocity reversal). If the candidate impulse violates the friction inequality, then "slipping" is assumed to occur and the tangential impulse is set equal to μ times the normal impulse. One expression of this law in our chosen coordinates and using kinematic restitution is:

$$\begin{aligned} P_n &= -(1+e)M_{11}V_{in} \qquad (\text{recall, } V_{in} < 0) \\ P_t &= -\frac{V_{it}}{\|V_{it}\|} \cdot \min\{\mu P_n, (1+e_t)M_{22}\|V_{it}\|\} \end{aligned} \tag{9}$$

This law is consistent with experiments in the following ways. For nearly normal collisions and positive e_t the tangential velocity is reversed (at least in part). Tangential velocity is not reversed for collisions closer to grazing incidence.

As emphasized above, this bilinear law is meant for diagonal M, the case shown in Fig. 5. For this case, this bilinear law is identical to the new law described below, and covers the shaded region shown in the figure.

Note that all the laws described above become identical for frictionless ($\mu = 0$) collisions, irrespective of whether M is diagonal or not.

5 A new algebraic collision law

All of the algebraic laws of which we are aware, such as the laws described above, lack one or more of the following features: the law is applicable to 3D, non-reversal of spin of bouncing balls can be predicted, reversal of spin of bouncing balls can be predicted, predictions are in the permissible region, at least a

2D subset of the permissible region is covered, calculation is simple and has unique solutions, and the parameters are easy to interpret.

Towards this end three new collision laws were constructed in Chatterjee (1997) and briefly described in Chatterjee and Ruina (1997). These new laws have all the features described in the paragraph above. The new laws are generalizations to non-diagonal-M of the bilinear law of Eq. 9. We present the simplest of these laws here.

The difficulties in generalizing Eq. 9 to arbitrary 3D collisions are that (1) the direction of the frictional impulse needs to be chosen (since there are two independent tangential directions); and (2) for fixed normal restitution e the magnitude of the normal impulse generally depends on the magnitude and direction of the tangential impulse. That the choice of impulse direction cannot be made in a naive manner is suggested by the possible energy-increases predicted by Whittaker's (Kane and Levinson's) law. Finally, a form should preferably be picked which leads to easy calculation.

The new law has two restitution parameters e and e_t. The *normal* restitution coefficient $0 \leq e \leq 1$ (kinematic or Newtonian) describes the degree to which the normal approach velocity is reversed[6]. The *tangential* restitution coefficient $-1 \leq e_t \leq 1$ is a measure of the extent to which the initial tangential velocity is reversed. Finally, the standard friction coefficient $0 \leq \mu \leq \infty$ is used to bound the impulse within the friction cone.

The new collision law expresses the collisional impulse as a weighted sum of a perfectly plastic, frictionless collision (P_I), and a perfectly plastic, sticking (i.e., no-slip, no-bounce) collision (P_{II}). Thus the 2D region accessed by the new law is in a plane that is a somewhat natural choice for general collisions (the same plane was independently proposed by Rubin, 1998).

This simple choice of impulse directions may be looked upon as a gross approximation that avoids the difficulties encountered with the 3D implementation of incremental laws (see e.g., the discussion of Routh's law in Bhatt and Koechling (1995a,1995b)), taking the view, as explained in the introduction, that even complicated incremental laws are themselves gross approximations with limited applicability. At the same time, this simple choice provides the twin benefits of a simple, extra collision parameter as well as explicit, linear equations that are easy to solve and have unique solutions (compare with Smith's Eq. 8, which is nonlinear and can also have some physically unrealistic (spurious) solutions which iterative numerical procedures have to recognize and avoid).

The calculation proceeds in two steps. First a normal impulse is calculated which enforces the kinematic restitution condition. Then an additional impulse parallel to the plane of maximum compression is computed which does not affect the normal restitution. The second impulse is computed to achieve the tangential restitution but can be clipped by the friction inequality.

5.1 "Sticking" Calculation of Candidate Impulse: No Friction Limitations

The candidate impulse, $\hat{P}$, is taken to be

$$\hat{P} := (1+e)P_I + (1+e_t)(P_{II} - P_I), \text{ where } 0 \leq e \leq 1, \ \ -1 \leq e_t \leq 1. \tag{10}$$

We prove below that the impulse $\hat{P}$ is guaranteed to satisfy the energy and non-interpenetration criteria (conditions 1 and 2 of section 3.1). Our proof consists of the following observations.

The set of possible impulses $\hat{P}$, as given by Eq. 10, is a parallelogram with 4 vertices within or on the boundaries of the the convex half-ellipsoidal region permitted by energy and interpenetration conditions.

[6]We have chosen to keep e between 0 and 1 for simplicity. However, this is not a fundamental restriction, and there can be kinematic restitution greater than 1 in some frictional collisions (see subsection 3.5). If a specific application requires it, e can be allowed to be greater than 1 but assurance of staying in the permissible region is lost and needs to be checked on a case by case basis.

P_I (for $e = 0$, $e_t = -1$) is a perfectly plastic frictionless collision. $2P_I$ (for $e = 1$, $e_t = -1$) is an energy conserving frictionless collision. $2P_{II}$ (for $e = 1$, $e_t = 1$) corresponds to a collision with perfect rebound, $V_f = -V_i$, hence, no interpenetration occurs and energy is conserved. For the fourth vertex, consider applying the corresponding impulse $2P_{II} - P_I$ (for $e = 0$, $e_t = 1$) as a sequence of two impulses, $2P_{II}$ and $-P_I$. After $2P_{II}$ has acted, the energy is the same as before while the velocity is reversed, becoming $-V_i$. The impulse $-P_I$ corresponds to a perfectly plastic, frictionless collision for that (new) initial condition. Thus the fourth vertex corresponds to a collision terminating on the plane of maximum compression and inside the initial energy surface.

Thus the interior of the parallelogram described by Eq. 10 lies entirely in the region permitted by energy and interpenetration conditions (end of proof).

5.2 "Slipping" Check of Candidate Impulse: Consideration of Friction

The candidate impulse, $\hat{P}$, is of the form $(1 + e)P_I + c(P_{II} - P_I)$, for a positive number $c = 1 + e_t$. If $\hat{P}$ lies inside the friction cone, then all criteria of subsection 3.2 are satisfied, and we take $\hat{P}$ to be the impulse transmitted in the collision.

Note that $(1 + e)P_I$ corresponds to a frictionless collision and so always lies inside the friction cone. If $\hat{P}$ lies *outside* the friction cone, then there is some constant $\kappa \geq 0$ such that the impulse

$$P^* := (1 + e)P_I + \kappa(P_{II} - P_I) \tag{11}$$

lies exactly *on* the surface of the friction cone. In such cases, we take P^* to be the impulse transmitted in the collision. The calculation is straightforward.

$$\text{If} \quad \|\hat{P} - (n^T\hat{P})n\| > \mu n^T\hat{P} \tag{12}$$

$$\text{then} \quad \kappa = \frac{\mu(1 + e)n^T P_I}{\|P_{II} - (n^T P_{II})n\| - \mu n^T(P_{II} - P_I)} \tag{13}$$

$$\text{else} \quad \kappa = 1 + e_t \tag{14}$$

(note that the check is sufficient even for the problematic cases: (a) when $n^T\hat{P} < 0$, i.e., when $\hat{P}$ has a negative normal component, and (b) when P_{II} and hence $\hat{P}$ is along the normal direction). Finally, obtain P^* from Eq. 11.

Geometrically, the preceding calculation is equivalent to projecting the candidate impulse $\hat{P}$ on to the friction cone along a line in impulse space in the $P_I - P_{II}$ direction. This shows that for frictionless collisions the present law reduces to the other laws reviewed here, with Newtonian restitution (as may also be seen by putting $\mu = 0$ into Eq. 13). Since the impulse $(1 + e)P_I$ is known to be inside the friction cone, and since the calculation of Eq. 13 is only attempted when $\hat{P}$ lies *outside* the friction cone, existence and uniqueness of solutions are assured from geometrical considerations.

One might also construct collision laws that predict approximately the same impulse as suggested above, but with a smoothened transition between "sticking" (no projection required) and "slipping" (projection required). We skip such considerations for simplicity.

To summarize, the collision calculation consists of Eqs. 3, 4 and 10, followed by Eqs. 12 through 14, and finally Eq. 11. For clarity, the calculations described above are given in MATLAB code in appendix A.2. The 2D region in 3D impulse space that is accessible to the new collision law is shown schematically for a generic collision in Fig. 2. See also Fig. 4.

Since the corresponding calculations for the case of spheres and disks are exactly equivalent to those in the bilinear law referred to earlier, this law *also* reduces to that bilinear law whenever the mass matrix M is diagonal (as for collisions of disks and spheres; see Fig. 5).

For grazing collisions (Fig. 6) the region accessible to the new law shrinks to a line that includes the coverage of the other 3D algebraic laws except for Routh's law. Rightly or wrongly, the new law does not predict the normal rebound of Routh's law for the frictional locking type of situation shown. On the other hand, the new law can predict tangential rebound, which Routh's law cannot.

For collisions with infinite friction (see Fig. 7) the new law covers the other laws considered, perhaps reflecting an ability to deal with a broader range of tangential compliances.

Finally, the new law makes reasonable predictions for a constrained problem with sufficiently high friction, with the constraint modeled as the limiting case of M having a large eigenvalue (see Fig. 8).

5.3 Composite collision laws

A simple way to generalize the new law, to (say) allow normal rebound in grazing collisions, is to use it in a weighted sum or interpolation with one or more other laws each of which generates impulses in the permissible region. Since the permissible region in impulse space is convex the interpolated impulses would also be in the permissible region.

For example, we might introduce an interpolation parameter σ, and define a "newer" collision law as follows:

> Given M, V_i, μ and e, calculate the collision impulse P_R as predicted by Routh's law, and the impulse P_{new} as predicted by the new law. Using the parameter $0 \leq \sigma \leq 1$, now compute the collision impulse to be $P := \sigma P_{new} + (1 - \sigma)P_R$.

The extreme of $\sigma = 0$ returns Routh's law, while $\sigma = 1$ makes our Routhless prediction.

At present there does not seem to be enough data to suggest that a good fit with such a combined law would have a usefully greater accuracy than either law alone. But such an interpolation will yield in effect a three-parameter collision law that will be able to capture phenomena such as simultaneous normal *and* tangential rebound in tangential collisions. For general 3D collisions, such a composite law will access a three dimensional subset of the accessible region.

6 Simultaneous multiple collisions

Although simultaneous collisions of independently moving objects seem non-generic, the collision of one object with another that already is in contact with still other objects is common in practice. And although single-point collisions of seemingly ideal mechanisms fit the formalism described in the paper thus far, they may also be treated as simultaneous collisions. Hence a recent surge of interest in predicting the outcome of simultaneous collisions.

The first step in generalizing the impulse space approach is straightforward. Recall that the impulse-momentum relations for a single-point collision are of the form $P = M(V_f - V_i)$, equivalent to $WP = V_f - V_i$ (where $W = M^{-1}$). It is well known (see e.g., Pfeiffer and Glocker, 1996) that a similar equation holds for multiple collisions, which we write as

$$WG = U_f - U_i \,, \tag{15}$$

where the column matrix G contains the components of the impulses P at all k locations, and the column matrices U_f and U_i contain the components of the relative velocities V_f and V_i at all k locations. In what follows, we will loosely call G an impulse, and the U's relative velocities. In general W will be symmetric and positive semidefinite. Often, W will be singular, corresponding to the possibility of self-equilibrating impulse sets (say, equal and opposite impulses along the line joining the two contact points on a body that simultaneously touches two immovable objects). Thus the permissible region will often be infinite in extent unless new restrictive assumptions are added.

6.1 The ill-posed nature of simultaneous multiple collisions

In some sense, a rigid body collision law is a distinguished limit of the properties of a deformable body. The limit is distinguished in that the limiting behavior depends not just on some single measure of deformation going to zero, but also on ratios of deformation moduli (say tangential to normal compliance). For single-point collisions the relevant ratios may not vary too wildly from one collision to another similar collision, hence the hope of deterministic collision laws with some predictive utility. For multiple impacts, however, the key ratios in the distinguished limit include, say, the ratios of (very small) separation distances of various pairs of contacting points. These ratios can differ radically for macroscopically similar-looking situations.

Most notably, all collision laws give vastly different results if the order of near-simultaneous collisions is changed by minute changes in initial conditions, for typical configurations. But since "simultaneous" is only a coarse-grained description of a complex event that does have structure in time, variously ordered near-simultaneous collisions are all different possibilities for the same "simultaneous" collision. Similarly, the outcome of approximately simultaneous collisions also depends sensitively on the relative compliances of the various contacts and the relative degree of rigidity of the various bodies. These issues have been discussed by Ivanov (1995).

The ill-posed nature of "simultaneous" collisions is not a mathematical fine point that can be removed by an appropriate (possibly very impressive looking) formalism. The ill-posedness of some multiple-point collisions is a fundamental physical property of the collisions, not a consequence of an unsuitable modeling approach. The outcome of a "pool break" will vary from game to game, heedless of the unique outcome predicted by any simultaneous-collision law. The classic example of such a situation is the longitudinal impact of three spheres discussed in detail by Ivanov (1995). Another example is a horizontal rod that falls onto a horizontal, frictionless table, making contact at its two endpoints, studied in detail by Goyal *et al.* (1998a, 1998b).

The authors of papers on multiple impacts are aware of these difficulties. The primary motivation behind such studies is often to develop a mathematically consistent framework for simulations. For example, Moreau (1983) says,

> ... physical situations undoubtedly exist, where this concept gives a reasonably accurate description of reality. But, to the author's opinion, the main interest of the concept lies in the internal mathematical consistency of the resulting evolution problem.

He then adds that an existing, simple and consistent theory may more easily allow more realistic empirical corrections.

We summarize the simultaneous collision problem as follows: Inherently inescapable ill-posedness notwithstanding, simulators will simulate simultaneous collisions and there is a need for simultaneous-collision laws that are as reasonable as possible.

6.2 The LCP formulation and multiple collisions

Although we do not adopt such an approach here, we briefly mention the *Linear Complementarity Problem* or LCP (see Cottle, *et al.*, 1992) approach to multiple collisions.

For simplicity, consider what a frictionless, plastic collision might mean in a multiple collision context. We cannot require that the normal components of relative velocities at all contact locations be exactly zero at the end of such a collision (else, some of the normal impulses required may be negative).

Several authors adopt the *complementarity conditions* that at each location, either the normal velocity comes to zero and there is a non-negative impulse, or there is zero impulse and the post-collision

normal velocity is non-negative (see e.g., Löstedt, 1982; Baraff, 1989; Pfeiffer and Glocker, 1996; Anitescu and Potra, 1997). For a plastic frictionless multiple-collision this assumption guarantees a possibly non-unique but always physically permissible G_I.

LCP based methods have been extended to deal with frictional multiple collisions with nonzero normal and tangential restitution in 2D (Pfeiffer and Glocker, 1996), and without tangential restitution in 3D (Anitescu and Potra, 1997).

Note that the realism of such complementarity conditions is not a fundamental property of rigid body collision dynamics, although similar complementarity conditions are fundamental properties of smooth motions (Chatterjee, 1998). We know of no reason to prefer the *predictions* of an LCP-based law to those of any other law.

6.3 Extension of the new law to multiple collisions

One approach to multiple collisions might be to repeat the approach we have taken to single point collisions. One might construct a permissible region in the higher dimensional impulse space of G and then seek a parameterization of a reasonable lower-dimensional subset. Using a linear combination of a perfectly plastic frictionless collision with a perfectly plastic sticking collision (one from the solution set of each would need be selected) might still be a reasonable choice. The details of the projection on to the frictional cone(s) would be more involved. This approach might result in a versatile collision law, depending on the modeling choices made. We have not pursued this, or other similar higher dimensional approaches, thus far.

6.4 Using the new law in a sequence of simple collisions

In the first three sections of this paper we have described a basis for defining reasonable single-point collision laws and have presented one such law. The classical approach to multiple collisions, as discussed among others by Ivanov (1995) and Baraff (1989), is to treat them as a sequence of collisions.

Well-known drawbacks to this approach are that (a) it involves indeterminacy in the sequence to be used (Ivanov, 1995), and (b) it can lead to an infinite set of collision calculations to complete one macroscopic simultaneous collision (Baraff, 1989). Both of these drawbacks are discussed below.

Let us consider the use of our new law in multiple collisions, broken up into a sequence of single collisions. For k "simultaneous" contacts, we use $2k$ collision parameters and k friction parameters. If frictionless hinges are treated as places of simultaneous collision they need no parameters (they have plastic, sticking collisions). The collection of bodies involved in the simultaneous collision are considered as separate bodies. Whether mechanisms are broken into separate bodies or not is a modeling choice (discussed separately below).

We propose that "simultaneous" collisions should be ordered by normal approach velocity. After r single-point collisions (where $r \geq 0$ starts at 0), we pick the contact location where the *normal* component of relative velocity is the lowest (i.e., we find the contact location with the greatest normal approach velocity). If this turns out to be a non-negative number (i.e., separating or sliding but not interpenetrating), then the collision is over. If it is negative, then we carry out a single-point collision calculation for that location *ignoring* the other locations (i.e., assuming that no impulses occur at the other locations). Once the collision impulse is determined, its effect on all the relative velocities at all the contact locations is found, using Eq. 15. This gives the initial velocities for the next stage of the calculation, and the process is repeated.

Having two potential contact points with the same approach velocity *is non-generic* and can only occur if there is special symmetry. Such situations include the horizontally falling rod (Goyal *et al.*, 1998a), two identical spheres hitting a third sphere from two sides, with identical velocities (Ivanov,

1995) and a "pool break" with an exactly centered initial shot. In this non-generic case three options available to the simulator are: 1) Change the initial conditions in the simulation (within a reasonable space of initial uncertainty) to break the symmetry, 2) order the "equal-approach-velocity" collisions arbitrarily (say, by taking the dot product of the position with a predefined non-special direction), or 3) following Ivanov's recommendation for such problems randomize the collision order so that the genuine indeterminacy in the experiments would then be represented by indeterminacy in the collision law.

If the process described above terminates in a finite number of single-point calculations, then we will have a prediction that is reasonable, obeys basic restrictions, and has all the merits of the single-point law used.

However, the prescription may *not* terminate in a finite number of collisions.

In such cases, so long as rare energy conserving conditions are not met[7] a single-point collision law dissipates a nonzero fraction of the system's kinetic energy at every collision. Thus, for $e < 1$, if the number of collisions is infinite, the system will *exponentially* (in the number of collisions – macroscopic time is not advanced) approach a state of zero interpenetration velocity at all locations due to the dissipation in each collision. Either by extrapolation or by simple truncation, the end of the collision can still be found in a reasonable time. Since the possibility of infinitely many collisions in finite time is unavoidable even in single-point collisions (as for the classic case of a ball bouncing vertically on the floor, with $e < 1$) series truncation or extrapolation is already a necessary part of general rigid body simulations with collisions.

For the nonphysical cases where one wants to model energy conserving collisions, it is possible to construct examples where the sequence of collisions continues indefinitely with no energy dissipation (we do not mean infinite sequences of separate collisions, but sequences where the system never changes its configuration and the individual impacts are spaced apart by infinitesimal time intervals). Thus for energy conserving collisions our proposed sequencing does not necessarily generate a prediction.

The extension to multiple impacts that we have described above has, by construction, most of the attractive features that originally motivated the single-point version. In particular:

1. It is algebraic, and for each location has two parameters (other than friction) with approximate but clear physical interpretations.

2. It is simple, in that its algorithmic requirements are limited to basic matrix algebra and linear calculations.

3. Existence of solutions is guaranteed whenever $e < 1$, and even for $e = 1$ in all but some non-generic collision configurations. Except for systems with special symmetry, uniqueness of solutions is guaranteed. The non-uniqueness that remains is in keeping with the physics.

4. The law reduces exactly to the single-point version if there is only one contact.

We mention that simulations using the new law in a sequence of collisions with between 5 to 10 bodies and about 30 "simultaneous" contacts have been conducted by B. Mirtich (personal communication), with promising results.

6.5 Mechanism constraints as simultaneous impacts

Although the colliding pendulum of Fig. 8 fits in the formalism of single point collisions with one eigenvalue of M approaching infinity, it may also be viewed as a simultaneous collision at both rod

[7]Namely (i) $e = 1$ and $\mu = 0$, (ii) $e = 1$, $e_t = 1$, $n^T P_{II} > 0$ and μ sufficiently large, or (iii) $e = 1$, and MV_i directed exactly along the normal direction (so that P_{II} is normal.)

ends. Within the sequencing approach we propose for "simultaneous" collisions, this case would be treated as a collision of a free floating rod with the ground followed by a sticking collision of the other end with the hinge. This in turn may or may not lead to an infinite sequence as described.

For collisions of mechanisms there is then the choice of whether to treat the mechanism as a set of bodies or as an intact mechanism. The intact-mechanism approach leads, at least for single point collisions, to a single direct collision calculation but is perhaps not likely to be accurate (since the mechanism constraints cannot be expected to be closely obeyed[8] *during* the collisions). On the other hand, the independent body approach frequently leads to an infinite series of calculations (which needs to be truncated or extrapolated).

7 Concluding comments

The set of reasonably permissible impulses for single point collisions has been described. The properties and shortcomings of some known collision laws have been described, in part using the impulse space description. A new algebraic collision law has been presented for single-point impacts that is automatically in the "permissible" part of the impulse space, is capable of predicting a fairly wide range of phenomena, and has unique solutions that are easy to calculate even in 3D. A generalization to multiple impacts by means of sequenced collisions has been proposed.

We are hopeful that the impulse considerations and new law will facilitate more straightforward rigid body simulations with impacts.

We conclude with a few natural open questions. What part of the permissible region can be accessed by reasonable incremental (micro-contact) laws that involve some tangential compliance? What part of the permissible region is accessed by various experimental situations? Does the proposed sequencing, which renders many simultaneous collision situations determinate, predict outcomes that match those observed in practice?

References

[1] M. Anitescu and F. A. Potra. Formulating dynamic multi-rigid-body contact problems with friction as solvable linear complementarity problems. *Nonlinear Dynamics*, 14:231–247, 1997.

[2] D. Baraff. Analytical methods for dynamic simulation of non-penetrating rigid bodies. *Computer Graphics*, 23(3):223–232, 1989.

[3] J. A. Batlle and S. Cardona. The jamb (self-locking) process in 3D rough collisions. *ASME Journal of Applied Mechanics*, 65:417–423, 1998.

[4] V. Bhatt and J. Koechling. Partitioning the parameter space according to different behaviors during three-dimensional impacts. *ASME Journal of Applied Mechanics*, 62:740–746, 1995a.

[5] V. Bhatt and J. Koechling. Three-dimensional frictional rigid body impact. *ASME Journal of Applied Mechanics*, 62:893–898, 1995b.

[6] R. M. Brach. *Mechanical Impact Dynamics: Rigid Body Collisions.* John Wiley and Sons, New York, 1991.

[8] If the compliance effects in the bearings are comparable to the deformation effects during the collision, the constraint is not ideal.

[7] B. Brogliato. *Nonsmooth Impact Mechanics: Models, Dynamics and Control*. Springer, London, UK, 1996.

[8] A. Chatterjee. *Rigid Body Collisions: Some General Considerations, New Collision Laws, and Some Experimental Data*. PhD thesis, Cornell University, 1997.

[9] A. Chatterjee. Note on the realism of complementarity conditions in rigid body collisions. *In preparation,* 1998.

[10] A. Chatterjee and A. Ruina. Some new algebraic rigid body collision laws. In *Applied Mechanics in the Americas*, volume 4, pages 403–406, 1997. Proceedings of the *Fifth Pan American Congress of Applied Mechanics*, San Juan, Puerto Rico.

[11] A. Chatterjee and A. Ruina. Two interpretations of rigidity in rigid body collisions. Accepted for publication in the *ASME Journal of Applied Mechanics*, 1998a.

[12] A. Chatterjee and A. Ruina. Realizability of Arbitrary Local Mass Matrices in Single-Point Rigid Body Collisions. To appear in the proceedings of the *International Symposium on Impact and Friction of Solids, Structures and Intelligent Machines*, Ottawa, Canada. June, 1998b.

[13] R. W. Cottle, J-. S. Pang and R. E. Stone. *The Linear Complementarity Problem*. Academic Press, San Diego, CA, 1992.

[14] R. L. Garwin. Kinematics of an ultraelastic rough ball. *American Journal of Physics*, 37(1):88–92, 1969.

[15] W. Goldsmith. *Impact: The theory and physical behavior of colliding solids*. Edward Arnold, Ltd., London, 1960.

[16] S. Goyal, J. M. Papadopoulos, and P. A. Sullivan. The dynamics of clattering I: Equations of motion and examples. *ASME Journal of Dynamic Systems, Measurement and Control*, 120(1):83–93, 1998a.

[17] S. Goyal, J. M. Papadopoulos, and P. A. Sullivan. The dynamics of clattering II: Global results and shock protection. *ASME Journal of Dynamic Systems, Measurement and Control*, 120(1):94–102, 1998b.

[18] S. Goyal, E. N. Pinson, and F. W. Sinden. Simulation of dynamics of interacting rigid bodies including friction I: General problem and contact model. *Engineering with Computers*, 10:162–174, 1994.

[19] A. P. Ivanov. Energetics of a collision with friction. *Journal of Applied Mathematics and Mechanics*, 56(4):527–534, 1992.

[20] A. P. Ivanov. On multiple impact. *Journal of Applied Mathematics and Mechanics*, 59(6):887–902, 1995.

[21] J. T. Jenkins. Boundary conditions for rapid granular flow: Flat, frictional walls. *ASME Journal of Applied Mechanics*, 59:120–127, 1992.

[22] K. L. Johnson. The bounce of "superball". *International Journal of Mechanical Engineering Education*, 11(1):57–63, 1983.

[23] T. R. Kane and D. A. Levinson. *Dynamics: Theory and Applications*. McGraw-Hill, New York, 1985.

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

[25] N. Maw, J. R. Barber, and J. N. Fawcett. The role of elastic tangential compliance in oblique impact. *Journal of Lubrication Technology*, 103:74–80, 1981.

[26] J. J. Moreau. Standard inelastic shocks and the dynamics of unilateral constraints. In *Unilateral problems in structural analysis*. International Centre for Mechanical Sciences, Courses and Lectures - No. 288, 173–221, 1983. Edited by G. D. Piero and F. Maceri.

[27] F. Pfeiffer and C. Glocker. *Dynamics of rigid body systems with unilateral constraints*. Wiley series in nonlinear science. John Wiley and Sons, New York, 1996. Series editors: Ali H. Nayfeh and Arun V. Holden.

[28] S. D. Poisson. *Traité de Mécanique*. Bachelier, Paris, second edition, 1833.

[29] E. J. Routh. *Dynamics of a System of Rigid Bodies*. Macmillan and Co., London, sixth edition, 1897.

[30] M. B. Rubin. Physical restrictions on the impulse acting during three- dimensional impact of two "rigid" bodies. *ASME Journal of Applied Mechanics*, 65:464–469, 1998.

[31] C. E. Smith. Predicting rebounds using rigid-body dynamics. *ASME Journal of Applied Mechanics*, 58:754–758, 1991.

[32] C. E. Smith and P.-P. Liu. Coefficients of restitution. *ASME Journal of Applied Mechanics*, 59:963–969, 1992.

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

[34] W. J. Stronge. Swerve during three-dimensional impact of rough bodies. *ASME Journal of Applied Mechanics*, 61:605–611, 1994.

[35] O. Walton. *Particulate Two-Phase Flow*, chapter 25. Butterworth-Heinemann, Boston, 1992. Edited by M. C. Roco.

[36] Y. Wang and M. T. Mason. Two-dimensional rigid-body collisions with friction. *ASME Journal of Applied Mechanics*, 59:635–642, 1992.

[37] E. T. Whittaker. *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*. Dover, New York, fourth edition, 1944.

A Details of Some Calculations for Single-Point Impacts

A.1 Calculating the Mass Matrix M

The mass matrix M is a property of the configuration of the rigid body system at impending contact. How M is most easily calculated in practice depends on what one is doing. In the worst case a rigid body simulator can find the three columns of M^{-1} by doing four non-collisional rigid body calculations.

This method is based on the fact that the relative acceleration between the contact points on the two bodies is the sum of two parts: a part that includes centripetal terms as well as the effects of other applied forces but not the effect of equal and opposite contact forces, and another part that is strictly linear in the equal and opposite contact forces. The calculation described below can be performed in any orthonormal coordinate system.

First, in the reference calculation, the two impending contact points are allowed to move as they would if they were not in contact but all the non-collisional forces in the system were acting. The difference between the acceleration vectors of the two contact points is the reference relative acceleration.

Next, a unit force is applied in the positive x_1 direction, on the contact point of body 2, and in the negative x_1 direction on the contact point of body 1. The two impending contact points are allowed to move as they would if they were not in contact but all the non-collisional forces in the system as well as the new forces in the x_1 direction were acting. The difference between the acceleration vectors of the two contact points is calculated. Subtracting from this the reference relative acceleration, the first column of M^{-1} is obtained.

A similar calculation, but with the equal and opposite forces acting along the positive and negative x_2, and then x_3, directions, gives the second and third columns of M^{-1}. The matrix M^{-1} thus obtained is inverted to give M.

A.2 MATLAB code for collision calculation

```
function P = collrule( M, V_i, n, mu, e, e_t)
% This matlab function calculates the collision impulse P, as
% predicted by the algebraic model proposed by Chatterjee and Ruina.
% Note: single quote (') means transpose
% M is a 3X3 symmetric positive definite matrix, V_i is 3X1,
% n is 3X1 with n'*n = 1,
% mu >= 0, 0 <= e <= 1, -1 <= e_t <= 1
% Assume that n'*V_i <= 0 (else, no collision occurs)

P_I = - n*n'*V_i /( n'*inv(M)*n );       % perfectly plastic, frictionless: Eq. 3
P_II = - M*V_i;                          % perfectly plastic, sticking:     Eq. 4
P_hat = (1+e)*P_I + (1+e_t)*(P_II - P_I);          % candidate impulse:    Eq. 10

if norm(P_hat - n*n'*P_hat) > mu*n'*P_hat          % friction check:       Eq. 12
  kappa = mu*(1+e)*n'*P_I/(norm(P_II - n*n'*P_II) - mu*n'*(P_II - P_I));%   Eq. 13
else
  kappa = 1 + e_t;                                                    %   Eq. 14
end

P = ( 1+e )*P_I  + kappa* (P_II - P_I);  % the collisional impulse:         Eq. 11
```

List of Figures

1 2D schematic of physically permissible region in impulse space for a generic collision. . . 24
2 Schematic diagram showing 2D region in 3D impulse space accessible to the new collision model. 24
3 Sketching regions in 2D impulse space accessible to various collision models. 25
4 A generic 2D collision. Shaded region is accessible to new model. 26
5 A collision with diagonal M (as for spheres and disks). Shaded region is accessible to new model. 26
6 A tangential or grazing collision. Thick line is accessible to new model (as shown). Note, Routh's method predicts finite normal rebound, but no tangential rebound. New model predicts no normal rebound, but permits tangential rebound. 27
7 A collision with infinite friction ($\mu \to \infty$). Shaded region is accessible to new model. Note that if the center of the ellipse, C, were to lie on the t or 2-axis, then Routh's model would predict only a perfectly plastic, sticking collision. 27
8 A 2D pendulum strikes a wall. Modeled by letting one eigenvalue of M approach ∞. Schematic diagram shows a finite portion of energy ellipsoid. Shaded region is accessible to new model. Kane and Levinson's (Whittaker's) model predicts unbounded gain in kinetic energy, does not converge to the kinematically constrained problem, and is not shown. 28
9 For a central collision (i.e., diagonal M) with sufficiently high friction and sufficiently close to grazing incidence, energy-dissipating collisions can occur with large values of e. 29

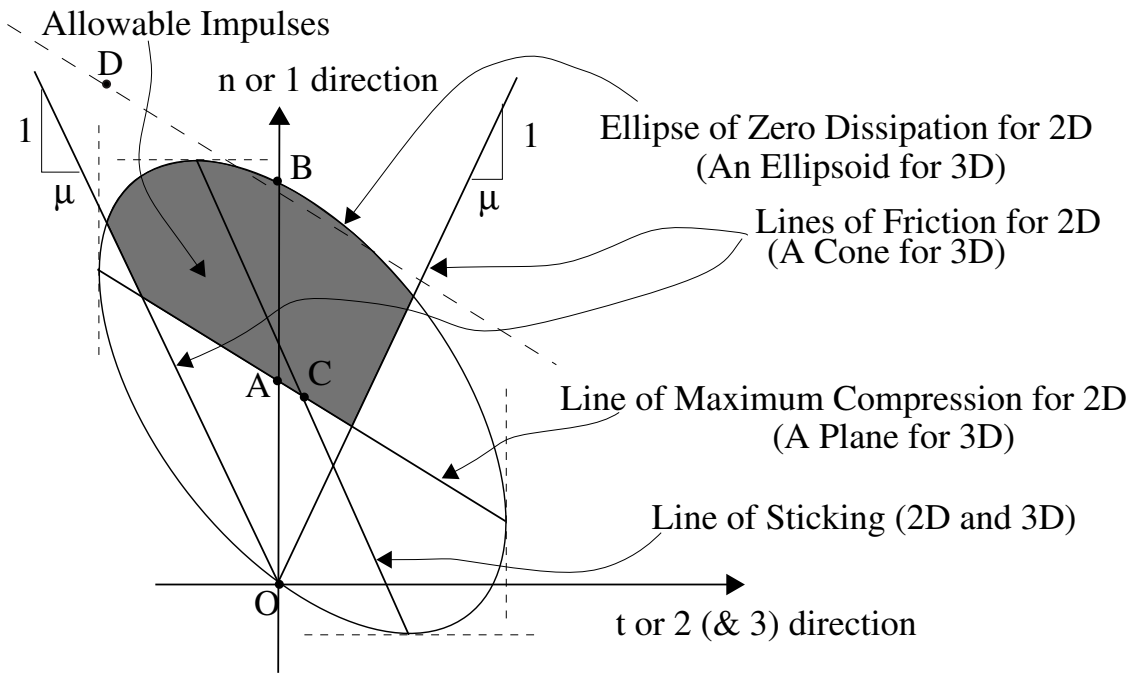

Figure 1: 2D schematic of physically permissible region in impulse space for a generic collision.

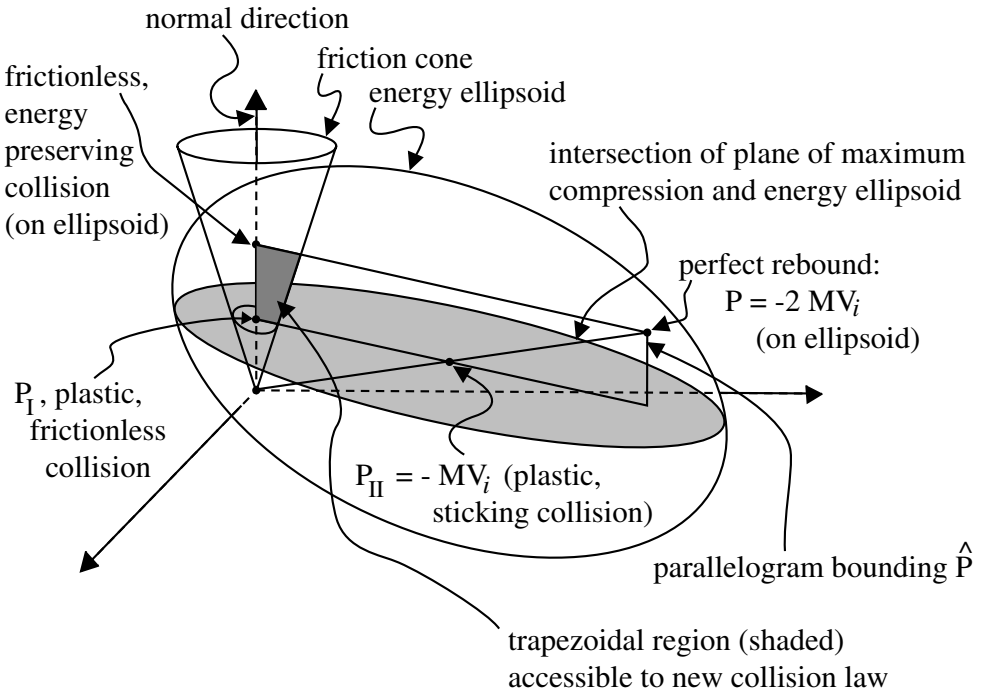

Figure 2: Schematic diagram showing 2D region in 3D impulse space accessible to the new collision model.

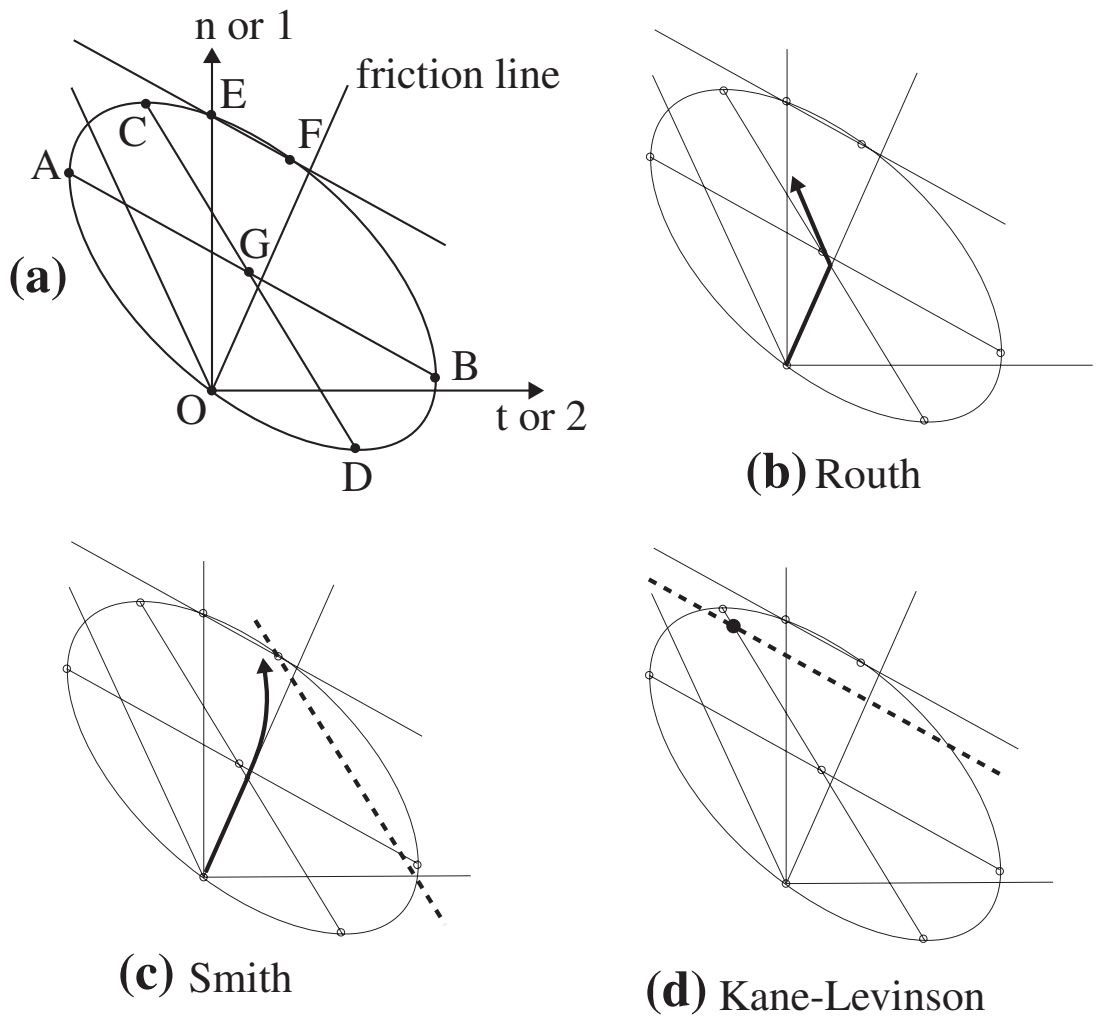

Figure 3: Sketching regions in 2D impulse space accessible to various collision models.

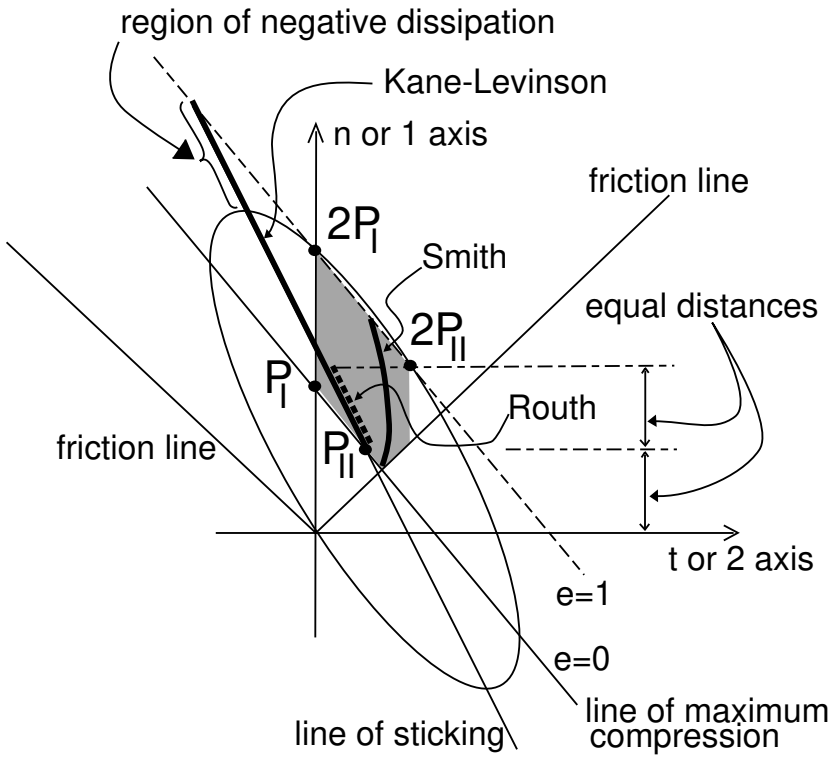

Figure 4: A generic 2D collision. Shaded region is accessible to new model.

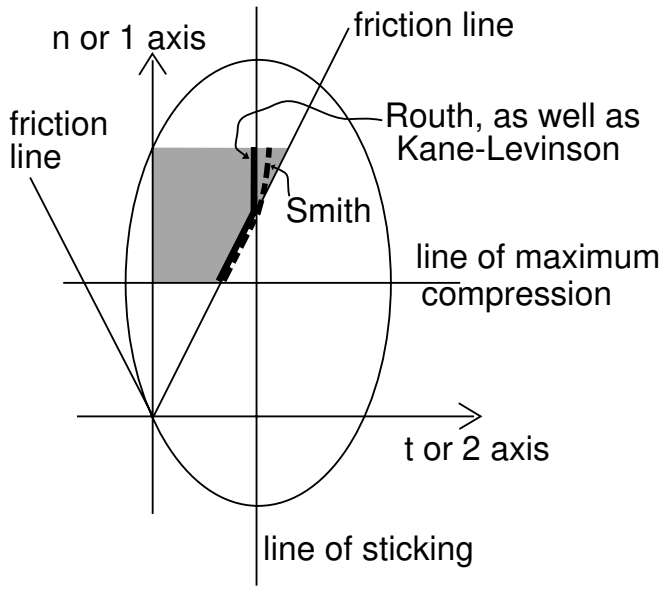

Figure 5: A collision with diagonal M (as for spheres and disks). Shaded region is accessible to new model.

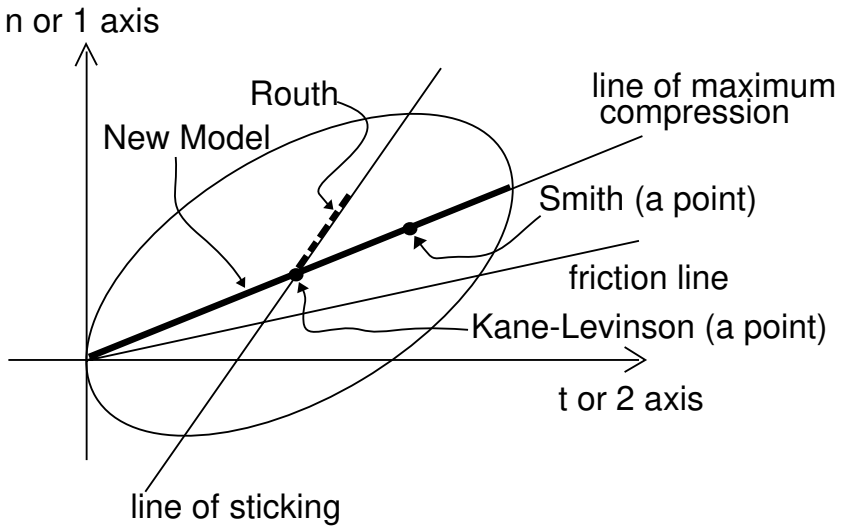

Figure 6: A tangential or grazing collision. Thick line is accessible to new model (as shown). Note, Routh's method predicts finite normal rebound, but no tangential rebound. New model predicts no normal rebound, but permits tangential rebound.

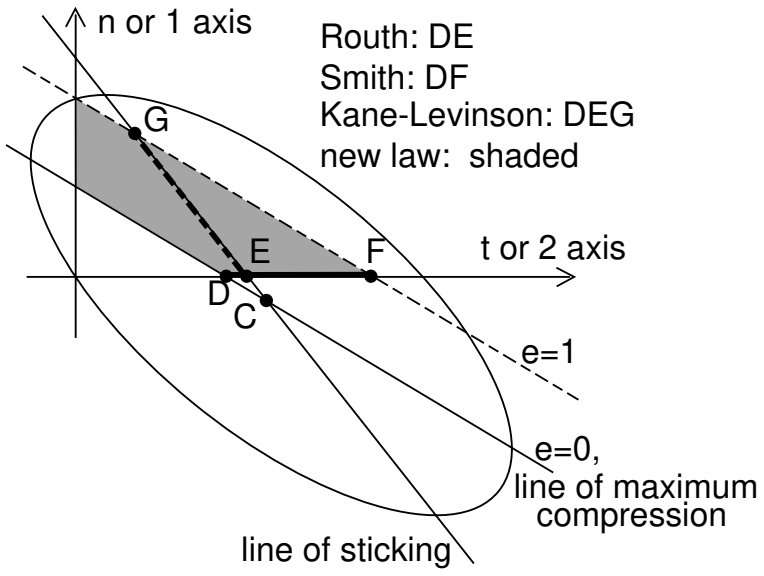

Figure 7: A collision with infinite friction ($\mu \to \infty$). Shaded region is accessible to new model. Note that if the center of the ellipse, C, were to lie on the t or 2-axis, then Routh's model would predict only a perfectly plastic, sticking collision.

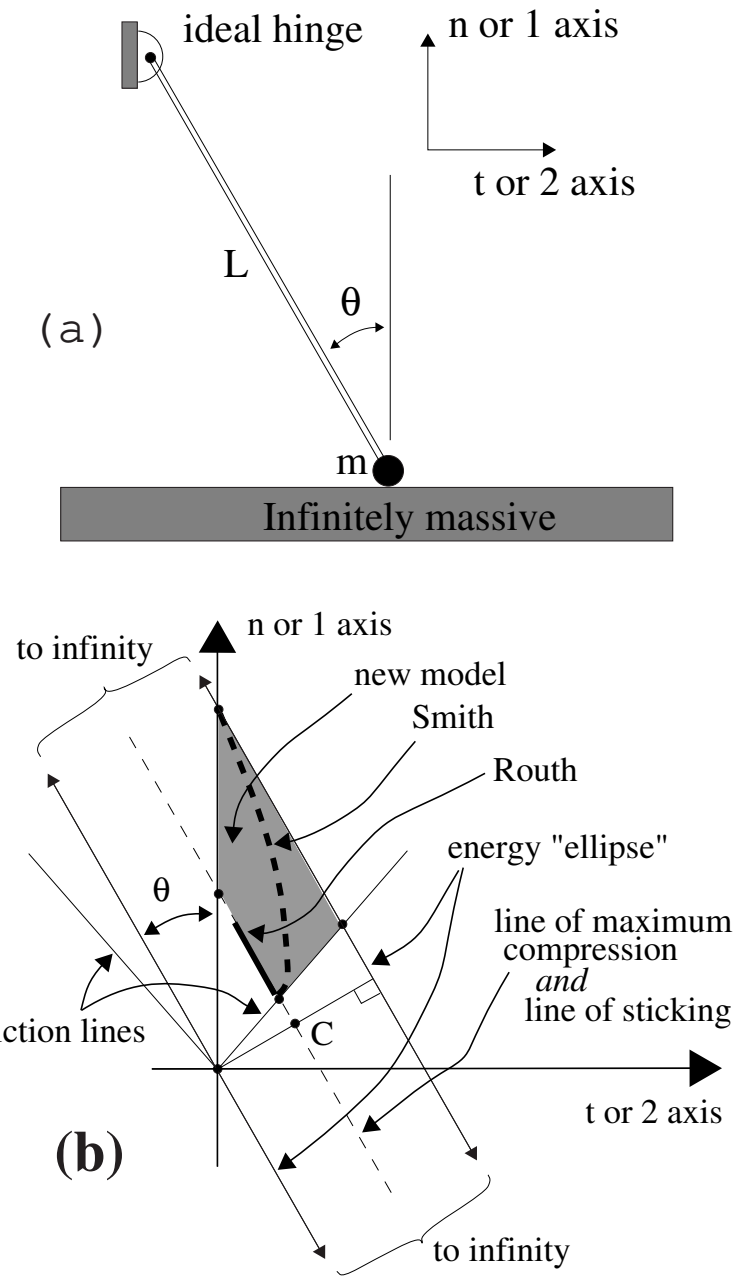

Figure 8: A 2D pendulum strikes a wall. Modeled by letting one eigenvalue of M approach ∞. Schematic diagram shows a finite portion of energy ellipsoid. Shaded region is accessible to new model. Kane and Levinson's (Whittaker's) model predicts unbounded gain in kinetic energy, does not converge to the kinematically constrained problem, and is not shown.

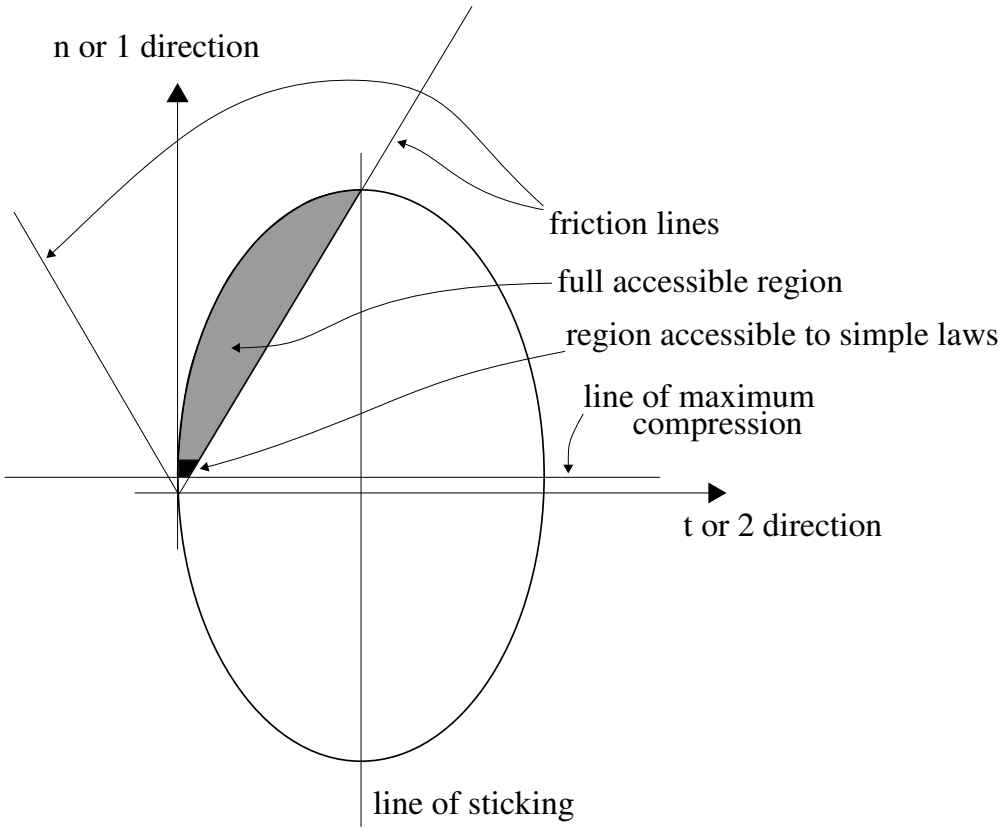

Figure 9: For a central collision (i.e., diagonal M) with sufficiently high friction and sufficiently close to grazing incidence, energy-dissipating collisions can occur with large values of e.