PERGAMON

Mechanism and Machine Theory 37 (2002) 1213–1239

MECHANISM AND MACHINE THEORY

www.elsevier.com/locate/mechmt

Literature survey of contact dynamics modelling

G. Gilardi, I. Sharf *

Department of Mechanical Engineering, University of Victoria, P.O. Box 3055, Victoria, BC, Canada V8W 3P6

Received 31 January 2001; accepted 18 March 2002

Abstract

Impact is a complex phenomenon that occurs when two or more bodies undergo a collision. This phenomenon is important in many different areas—machine design, robotics, multi-body analysis are just a few examples. The purpose of this manuscript is to provide an overview of the state of the art on impact and contact modelling methodologies, taking into account their different aspects, specifically, the energy loss, the influence of the friction model, solution approaches, the multi-contact problem and the experimental verification. The paper is intended to provide a review of results presented in literature and some additional insights into existing models, their interrelationship and the use of these models for impact/contact scenarios encountered in space robotic applications.

1. Introduction

The purpose of this paper is to provide an introduction and an overview of the state of the art on the subject of impact and contact dynamics modeling. *Impact* is a complex physical phenomenon, which occurs when two or more bodies collide with each other [1]. Characteristics of impact are very brief duration, high force levels reached, rapid dissipation of energy and large accelerations and decelerations present. These facts must be considered during the design and analysis of any mechanical system [2]. Furthermore, during impact, the system presents discontinuities in geometry and some material properties may be modified by the impact itself.

Contact is a more ambiguous term although it is frequently used interchangeably with impact. In our work, we use this term to describe situations where two or more bodies come in touch with each other at some locations. Inherently, contact implies a continuous process which takes place over a finite time.

* Corresponding author. Fax: +1-514-398-7365.

PII: S0094-114X(02)00045-9

In general, two different approaches can be distinguished for impact and contact analysis. The first approach assumes that the interaction between the objects occurs in a short time and that the configuration of impacting bodies does not change significantly. The dynamic analysis is divided mainly in two intervals, before and after impact, and in secondary phases, such as slipping, sticking and reverse motion. To model the process of energy transfer and dissipation, various coefficients are employed, mainly the coefficient of restitution and the impulse ratio [3,4]. Application of these methods, referred to as *impulse–momentum* or *discrete* methods [5], has been confined primarily to impact between rigid bodies. The extension to flexible systems as well as extension to more general cases involving multiple contacts and intermittent contact is quite complicated. The second approach is based on the fact that the interaction forces act in a continuous manner during the impact. Thus, the analysis may be performed in the usual way, by simply adding the contact forces to the equations of motion during their action period. This allows a better description of the real behavior of the system, in particular, with respect to friction modeling. More importantly, this approach is naturally suitable for *contact* modeling and complex contact scenarios involving multiple contacts and bodies. This approach is referred to as *continuous* analysis or *force based* methods [5].

In the following section, we present basic concepts and definitions used in any impact theory. This is followed by a general historical overview of the research on impact and contact dynamics modeling, starting with the initial models of Newton and Poisson through to the modern formulations capable of dealing with complex contact scenarios. In the body of the paper, starting with Section 2, we allocate one section to each the discrete and continuous models. Given the complexity of impact modeling, it is imperative that results obtained from theoretical analysis are confirmed with experimental measurements. Furthermore, as in any modeling of reality, the goodness of the model depends on the choice and accuracy of model parameters. Therefore, the last section of the paper is devoted to experimental model validation and identification of impact/contact parameters.

It is important to emphasize that the literature on contact/impact analysis is vast and spans many diverse disciplines. To narrow the list of citations, our review focuses on contact (and impact) *dynamics*, rather than contact *mechanics* treatments of the subject. The latter traditionally aims to solve for stress and displacement distributions in the contact patch, as well as the wave propagation problem. Analytical results are often sought for 'simple' geometry and material combinations of the contacting bodies, such as two spheres with identical elastic constants or impact of a mass on an elastic half space. In addition, contact mechanics solutions are obtained for a known loading condition, as in the case of the classical Cattaneo problem [6,7] where the normal loading is held fixed while the tangential load is increased monotonically. Jaeger [8] presents an excellent overview of several contact mechanics analyses. In distinction, contact dynamics models tend to deal with, not surprisingly, dynamic quantities such as forces, impulses and velocities of the contacting bodies. Being motivated by space robotic applications, our review favors the works presenting general contact dynamics formulations for multi-body systems.

1.1. Basic impact theory

Impact of two bodies is characterized by large reaction forces and changes in velocities of the two bodies. As a consequence, the bodies are subject to elastic and/or plastic deformation, with

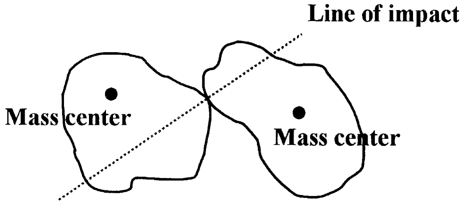

Fig. 1. Impact between two bodies.

dissipation of energy in various forms [9]. A generic impact can be represented as shown in Fig. 1. The line of impact is a straight line normal to the contacting surfaces at the contact point, [1] i.e., normal to the tangential plane at the contact point [1]. Four types of impact can be defined for single-point collision between two bodies: (a) *central* or *collinear*, if the mass centers of the two bodies are on the line of impact; (b) *eccentric*, if the mass centers of one or both bodies are not on the line of impact; (c) *direct*, if the initial velocities of the two bodies are along the line of impact; (d) *oblique*, if the initial velocities of one or both bodies are not along the line of impact. A class of *tangential impacts* was introduced by Wang and Mason [10], which is characterized by zero initial relative velocity along the line of impact.

The dynamics of impact is a very complex event, depending on many properties of contacting bodies such as material, geometry and velocity. In general, two phases can be identified: *compression* and *restitution*, as shown in Fig. 2 [2,3,9,10]. The first phase begins when the two bodies come in contact at the instant t_0 (point O), and finishes when the maximum deformation is reached at the instant t_m (point A), where the relative normal velocity is zero. The second phase begins at the instant t_m and finishes when the two bodies separate, i.e., instant t_f (points B, C or D). For impacts with sufficiently high velocities, not all deformation is recoverable because of the permanent (plastic) deformation and the resulting energy loss. With respect to the latter, impact can also be classified into: (a) perfectly elastic, line O–A–C, where no energy is lost; (b) perfectly plastic, line O–A, where all energy is lost and the deformation is permanent; (c) partially elastic, line O–A–D, with energy loss but no permanent deformation; (d) partially plastic, line O–A–B, with energy loss and permanent deformation.

The objective of impact modeling is to determine the after-impact conditions of the system, given its initial (pre-impact) configuration. Because of the complex dependencies on many parameters, one possible solution is to use experimentally measured coefficients. Coefficient of restitution, defined along normal direction, and coefficients along tangential directions are the most important [3,4].

1.1.1. Coefficient of restitution

The energy loss due to the motion in the normal direction can be expressed in terms of a coefficient, usually denoted by e. To honour the work-energy principle, it should satisfy the condition $0 \leqslant e \leqslant 1$ with the end conditions corresponding to perfectly elastic ($e = 1$) and perfectly

[1] The definition of the line of impact is an approximation since the contacting surfaces can have different normals at the point of contact.

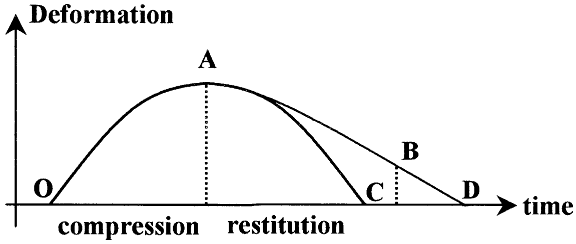

Fig. 2. Deformation during the impact.

plastic impact ($e = 0$). The coefficient of restitution depends on many elements, such as the geometry of the bodies in contact, the approach velocity, the material properties, the duration of contact and, possibly, friction [9]. Related to the energy loss and coefficient of restitution is the phenomenon of wave propagation. Since real bodies are not perfectly rigid, their parts are not instantaneously subjected to the same change of motion following the impact. The disturbance generated at the contact point travels in the body as stress (and deformation) waves with a finite velocity [2,11]. These waves produce oscillations and part of the impact (kinetic) energy is converted to the energy associated with this vibration. A useful parameter to quantify this situation is the ratio between the duration of impact and the period of the fundamental natural mode of vibration of the impacting bodies. The greater is this ratio, the smaller is the energy dissipation associated with the elastic waves since they can travel across the bodies many times before the impact ends. If this ratio is large enough, a quasi-equilibrium state is reached [2]. These observations are complemented by Stronge [11] who states that stress waves contribute to energy dissipation during collision if the relative size of the bodies is different from unity. In this case, the waves in the larger body do not have sufficient time to reflect from any boundary and, in essence, remain 'trapped' in the body. If the two impacting bodies are similar in size and material properties, the vibration energy lost is negligible [11]. The energy flow associated with impact dynamics in the normal direction is illustrated in Fig. 3.

The dependence of e on the approach velocity can be explained by considering the previously introduced ratio of the impact time to the fundamental period of vibration. If this ratio is large, then the coefficient of restitution is mainly determined by the plastic deformation near the impact point. The velocity of impact determines the extension of plastic deformation and this accounts for the basic velocity sensitivity of the coefficient of restitution of spherical bodies. Usually, bodies with a low ratio of surface area to volume (like spheres) present this kind of behavior [2,9]. For slender bodies like beams or plates, the time of impact tends to be short compared with the period of propagation of stress waves. Thus, a more significant portion of the initial kinetic energy re-

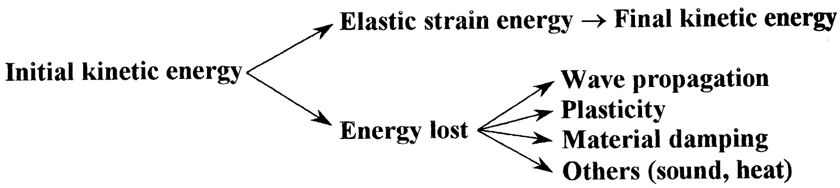

Fig. 3. Energy flow associated with normal direction.

mains in the bodies in the form of vibration. In this case the coefficient of restitution may be significantly influenced by this factor [2]. Later in the paper, we discuss three models proposed for the coefficient of restitution: Newton's [12], Poisson's [13] and Stronge's [14,15].

1.1.2. Tangential coefficients

Friction modelling is another key aspect of impact and contact dynamics, since friction can stop and/or reverse the motion as well as, it contributes to energy dissipation. If friction is taken into account, the relationship traditionally used to determine the force of dry friction is the Coulomb's law. This law states that the magnitude of the frictional force, F_t, can be related to the magnitude of the normal force, F_n, via a coefficient and its direction is always opposite to the relative tangential motion [16]. Two possible situations are distinguished: (a) sliding, with $F_t = \mu_d F_n$; (b) sticking, with $F_t \leqslant \mu_s F_n$, where μ_d is the coefficient of dynamic friction and μ_s is the coefficient of static friction. The two coefficients take into account the nature of the contact surfaces, mainly, the type of material and surface quality. Alternative ways to describe the behavior in the tangential directions include the use of tangential coefficient of restitution, based on Newton's model, and the impulse ratio defined as a ratio of tangential impulse to normal impulse. The latter is a generalization of the coefficient of friction and can take into account other tangential forces [3,4].

1.1.3. Impulse and impact energy

The time integral of the contact force $\mathbf{F}_C$ acting during the impact is called the impulse, $\mathbf{P}$, and is finite [10]:

$$\mathbf{P} = \lim_{\Delta t \to 0} \int \mathbf{F}_C(t)\, \mathrm{d}t \tag{1}$$

The energy loss E_L incurred by impact can be calculated as the negative work done by $\mathbf{F}_C$ during the collision [17], that is:

$$E_L = -\int \mathbf{F}_C \cdot \dot{\boldsymbol{\delta}}\, \mathrm{d}t = -\int \mathbf{F}_C \cdot \mathrm{d}\boldsymbol{\delta} \tag{2}$$

where $\dot{\boldsymbol{\delta}}$ is the velocity of deformation. Alternatively, by making use of the differential relationship between impulse and force, one can write:

$$E_L = -\int \dot{\boldsymbol{\delta}} \cdot \mathrm{d}\mathbf{P} \tag{3}$$

According to the work-energy principle, [2] the work done by $\mathbf{F}_C$ must be equal to the change in the kinetic energy, ΔT [18], so that:

$$\Delta T + E_L = 0 \tag{4}$$

It follows that the change in the kinetic energy depends on the coefficient of restitution and impulse ratios. In [17], Stronge states a theorem which allows to calculate the work done by contact

[2] Usually, in literature on the subject, the term used is the *energy-conservation principle*. This terminology is not completely correct, because of the presence of friction and other energy dissipation.

forces during periods of unidirectional slip as a simple function of the corresponding impulse and the average relative velocity at the contact point.

1.2. Literature overview

1.2.1. Discrete models

The impact between two rigid bodies was analysed initially by Sir Isaac Newton [19], and expanded by Whittaker [12] to account for frictional impulse. In that model the coefficient of restitution is a *kinematic* quantity that defines a relationship between the normal components of the velocities before and after the impact at the contact point (referred to as *Newton's* model). Routh [13] presented a graphic method based on a *kinetic* hypothesis to define the coefficient of restitution (referred to as *Poisson's* model). The coefficient of restitution is defined as a kinetic quantity that relates the normal impulses that occur during the compression and restitution phases. The two approaches are also different in the treatment of the motion in tangential direction during the impact. In Routh's study, the possibility of changes in slip direction during contact is taken into account, while in Whittaker's study it is not. In many simple cases, the two approaches lead to the same result, as shown by Wang and Mason [10], while in other cases, they can produce inconsistent results, as shown by Kane and Levinson [18] and Stronge [14]. This is a consequence of the possible changes in the slip direction. Ignoring these can lead to the overestimation of the final velocity after the impact, as illustrated with the Newtonian approach. Poisson's model, instead, can result in an increase of energy in some configurations for a perfectly elastic impact [14].

Brach [3,4] proposed an algebraic solution scheme, revising Newton's model and introducing *impulse ratios* to describe the behavior in the tangential directions. He defined the tangential impulse as a constant fraction of the normal impulse—the constant ratio of the two being the impulse ratio. It is equivalent to the friction coefficient in many cases. Brach also demonstrated that work-energy and kinematic constraints impose an upper bound on the impulse ratio. He also expanded this approach to include the impulse moments. Alternatively, the motion in the tangential direction was described by using the *tangential coefficient of restitution*. Smith [20] proposed another purely algebraic approach to the problem using the Newtonian definition for the coefficient of restitution. Impulse ratio is determined using as velocity an average value of different slipping velocities. Keller [21] developed an approach which involves the integration of the contact impulse variables. Thus, the system is treated as an evolving process parameterized by cumulative normal impulse. Also, by using a revised Poisson's model, Keller concluded that during impact, no increase in energy is possible. Using Routh's graphical method to analyse the contact models, Wang and Mason [10] identified the impact conditions under which Newton's and Poisson's models give the same solution. Stronge [14] demonstrated the energy inconsistencies in some solutions obtained with Poisson's model when the coefficient of restitution is assumed to be independent of the coefficient of friction. In that case Poisson's model does not lead to vanishing dissipation for a perfectly elastic impact. As a result, Stronge proposed to define the coefficient of restitution as the square root of the ratio of the elastic strain energy released during restitution to the energy absorbed by deformation during compression. With this definition no energetic inconsistencies are present [14].

Multiple frictional contacts in multi-body systems have been studied by several authors. Hurmuzlu and Marghitu [22] examine the problem where a planar rigid-body kinematic chain undergoes an external impact and an arbitrary number of internal impacts. The latter was defined as the situation when two bodies are simply in contact while impact occurs elsewhere in the system [22]. They developed a differential–integral approach, extending Keller's work [21] and using all three models for the coefficient of restitution, and an algebraic approach, based on Newton's model of restitution. Han and Gilmore [23] proposed a similar approach, using an algebraic formulation of motion equations, Poisson's model of restitution and Coulomb's law to define the tangential motion. Different conditions that characterize the motion (slipping, sticking, reverse motion) are detected by analyzing velocities and accelerations at the contact points [24], similarly to Hurmuzlu and Marghitu [22]. Han and Gilmore [23] verified their simulation results with experiments for a two-body and three-body impact.

Haug et al. [25] solve directly the differential equations of motion by using the Lagrange multiplier technique. For impact, Newton's model is used, while Coulomb's law is used for friction. Wang and Kumar [26], Anitescu et al. [27] reduce the problem to a *quadratic programming problem*.

1.2.2. Continuous models

Application of the impulse–momentum methods to model the impact dynamics of rigid bodies leads to several problems. First, in the presence of Coulomb friction, cases arise in which no solution or multiple solutions exist. Examples and analysis of these inconsistencies can be found in Wang and Kumar [26] and Mason and Wang [28]. These ambiguities have been attributed to the approximate nature of Coulomb's model and to the inadequacy of rigid body model, but no clear explanation has been found. The second problem is that energy conservation principles may be violated during frictional impacts, as shown by Stronge [14], as a consequence of the definition of the coefficient of restitution. Finally, the discrete approach is not easily extendible to generic multi-body systems. The use of *compliance* or *continuous* contact models where the impact force is a function of local indentation can overcome the problems encountered in the discrete formulation [26,29].

Different models have been postulated to represent the interaction force at the surfaces of two contacting bodies [2,3]. The first model was developed by Hertz [30], in which an elastostatic theory was used to calculate local indentation without the use of damping. The corresponding relationship between the impact force and the indentation is allowed to be non-linear. In the first and simplest model of damping, referred to as spring-dashpot model [9], the contact force is represented by a linear spring-damper element. Dubowsky and Freudenstein [31] presented an extension of this model called the impact-pair model, where they assumed a linear viscous damping law and a Hertzian spring for modeling the behavior of the impact surfaces. Hunt and Crossley [32] showed that a *linear* damping model does not truthfully represent the physical nature of the energy transfer process. Thus, they proposed a model based on Hertz's theory of contact with a *non-linear* damping force defined in terms of local penetration and the corresponding rate. Lee and Wang [33] proposed a similar model, but with a different function specifying the non-linear damping term. Other damping models have been proposed to describe totally or partially plastic impacts [2,3,15].

Contact stiffness and damping forces are dependent, at the minimum, on two parameters—the coefficient of stiffness and the coefficient of damping. For simple contact between two bodies, the former is determined by the geometry and the material of the contacting objects, while the coefficient of damping can be related to the coefficient of restitution [32,34,35]. An important advantage of continuous contact dynamics analysis is the possibility of using one of many friction models available in literature. Different models have been developed to permit a smooth transition from sticking to sliding friction [29,36–38]. Non-linear models, as well as non-local models have been used to represent the real behavior of the surface irregularities that cause the friction. The use of continuous models for contact forces allows to generalize the contact dynamics methodology to multi-body/multi-contact scenarios, as well as contact involving flexible bodies [5,38].

Two solution approaches can be distinguished in the context of continuous impact models. In the first, the contact model is expressed as an *explicit* functional relationship between the contact force and the generalized coordinates and their rates, with dependencies on certain geometric and material parameters. Application of this approach has been studied by several authors, including Ma [38], Kraus and Kumar [39], Deguet et al. [40], Vukobratovic and Potkonjak [41]. The contact condition is a geometric state, and involves determination of the minimum distance or interference between surfaces [5,38]. Several friction models have been used with the explicit contact model, for example, Coulomb's model and its variations [39,41–43], or the bristle model [38].

The second approach for solving impact/contact dynamics within the continuous framework takes into account the deformation due to contact directly via the flexibility of the contacting bodies. No explicit relationship is employed between the normal contact forces and the indentation, however, the condition of impenetrability at the contact point must be enforced. This approach has been used by Kim [5], Bathe and Bouzinov [44], Farahani et al. [45], Heinstein et al. [46]. Impact can still be detected by checking the minimum distance between the bodies, similarly to the explicit solution. By imposing the geometric condition of impenetrability, it is possible to calculate the contact force, using the Lagrange multipliers method [5,44,46] or with other mathematical techniques [45]. This method is typically used in conjunction with the finite element discretization of the contacting bodies (or contacting regions). It is the closest to reality and makes no assumptions nor approximations on the fundamental nature of contact dynamics.

2. Discrete contact dynamics models

The discrete formulation is based on the assumptions that [10]: the impact process is instantaneous and impact forces are impulsive; kinetic variables have discontinuous changes while no displacements occur during the impact, and that other finite forces are negligible. This model is used mainly if the impact involves *rigid* or very hard and compact bodies, while the effects of deformation at the contact point are taken into account through coefficients. The impact problem is solved by using the linear impulse–momentum principle, the angular impulse–momentum principle, and the relations between the variables before and after impact [3,4]. If m is the mass, $\mathbf{v}$ the center of mass velocity, $\mathbf{P}$ the linear impulse due to impact, $\mathbf{h}$ the angular momentum, $\mathbf{d}$ the distance from the center of mass to the point of impact and $\mathbf{M}$ the angular impulse due to impact, the impact dynamics equations are:

$$\begin{aligned} m_1(\mathbf{v}_1 - \mathbf{v}_{10}) &= \mathbf{P} \\ m_2(\mathbf{v}_2 - \mathbf{v}_{20}) &= -\mathbf{P} \\ \mathbf{h}_1 - \mathbf{h}_{10} &= \mathbf{d}_1 \times \mathbf{P} + \mathbf{M} \\ \mathbf{h}_2 - \mathbf{h}_{20} &= \mathbf{d}_2 \times \mathbf{P} - \mathbf{M} \end{aligned} \tag{5}$$

In the above, indexes 1 and 2 specify the body, while 0 denotes the initial conditions. The unknowns are the linear and angular velocities of the two bodies and the impulses $\mathbf{P}$ and $\mathbf{M}$. The angular impulse $\mathbf{M}$ is neglected in the majority of formulations since, consequently to the basic assumptions, the contact region must be small.

Additional relations are required to solve for the unknown impact variables. For the normal direction, one relation is provided by the coefficient of restitution. In the tangential direction, the relational laws may have to be replaced with kinematic constraints (for instance, during sticking, zero tangential velocity is imposed).

2.1. Coefficient of restitution models

Let the vector triad $(\mathbf{n}, \mathbf{t}, \mathbf{b})$ define a coordinate system with origin at the contact point, where $\mathbf{n}$ is the normal to the two bodies at that point, and vectors $\mathbf{t}$ and $\mathbf{b}$ define the tangent plane [3,4]. Then the linear impulse can be written as

$$\mathbf{P} = P_n\mathbf{n} + P_t\mathbf{t} + P_b\mathbf{b} \tag{6}$$

The relative linear velocity at the contact point, denoted by $\mathbf{C}$, has a component along the normal direction, called the *compression* velocity, and a component along the bi-tangential direction, called the *sliding* velocity [10]. The principle models of restitution are introduced below.

2.1.1. Poisson's model

In Poisson's model [13], the total normal impulse, P_f, is divided in two parts, P_c and P_r, corresponding to compression and restitution phases, respectively. The coefficient of restitution is defined as [10]

$$e = \frac{P_r}{P_c}, \quad P_f = P_c + P_r \tag{7}$$

The condition for the end of compression phase is zero relative velocity along the normal direction, that is $\mathbf{C} \cdot \mathbf{n} = 0$. In the $P_n - P_t$ space, this represents the *line of maximum compression* [10]. Using this definition and Eqs. (5)–(7), it is possible to define the *line of termination* as:

$$C_0 + \frac{C_1}{1+e}P_n + C_2P_t = 0, \quad t = t_f \tag{8}$$

where C_0 is the approach velocity, C_1 and C_2 are parameters depending on initial conditions, geometry and inertia [10].

2.1.2. Newton's model

In Newton's model the coefficient of restitution is defined as [12]:

$$e = -\frac{\mathbf{C}(t_f)\cdot\mathbf{n}}{\mathbf{C}(t_0)\cdot\mathbf{n}} = -\frac{C_f}{C_0} \tag{9}$$

This model is based on a kinematic point of view and only the initial and final values for the relative normal velocity are taken into account. The line of termination is given by [10]

$$(1+e)C_0 + C_1P_n + C_2P_t = 0, \quad t = t_f \tag{10}$$

2.1.3. Stronge's model

This model is based on the *internal energy dissipation hypothesis* [14]. The coefficient of restitution is defined as the square root of the ratio of energy released during restitution to the energy absorbed during compression. In terms of the work done by the normal force during the two phases, the coefficient of restitution can be calculated from:

$$e^2 = \frac{W_r}{-W_c} \tag{11}$$

It can be shown that the energy hypothesis leads to the only model which ensures that the energy loss from sources other than friction is non-negative, and is zero when $e = 1$. In [47], Stronge applies the above definition to derive a theoretical expression for e in terms of W_c and the work required to initiate yield.

In a recent work [15], Stronge considers the problem of oblique impact of a rigid cylinder on a deformable half-space. It is noted that in this and similar cases of collisions between objects of very different sizes, the energy loss to stress waves, W_w, is substantial and can be accounted for with the following definition of the coefficient of restitution:

$$e^2 = \frac{W_r - W_w}{-W_c}. \tag{12}$$

2.2. Additional relations

To obtain additional equations, one can pursue either of two possibilities. The first is to define coefficients of restitution for each of the other directions. For instance during slipping, one can obtain the following relationships by applying Newton's model in the tangential plane [3]:

$$e_t = -\frac{\mathbf{C}(t_f)\cdot\mathbf{t}}{\mathbf{C}(t_0)\cdot\mathbf{t}}, \quad e_b = -\frac{\mathbf{C}(t_f)\cdot\mathbf{b}}{\mathbf{C}(t_0)\cdot\mathbf{b}} \tag{13}$$

In the same manner, it is possible to define three coefficients to model the rotational effects of the impact [3,4]. The second possibility is to define a relationship between quantities in the normal and tangential directions. Brach [3,4] proposed to relate the tangential and normal impulses with equations analogous to Coulomb's law:

$$P_t = \mu_t P_n, \quad P_b = \mu_b P_n \tag{14}$$

where coefficients μ are called impulse ratios. Smith [20] extended the use of Coulomb's law to allow for a change in the direction of tangential velocity. If sticking occurs before the end of the impact, relationships (13) and (14) must be replaced with a condition of zero slip, i.e., $\mathbf{C}\cdot\mathbf{t} = 0$ and $\mathbf{C}\cdot\mathbf{b} = 0$, referred to as *lines of sticking* [23].

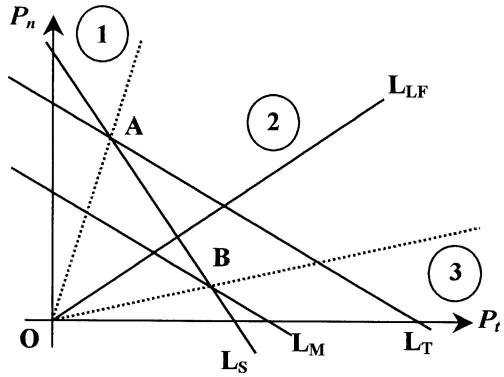

Fig. 4. Graphical impact analysis.

2.3. Solution of impact problem

2.3.1. Planar impact of two bodies

Planar impact analysis is conveniently carried out with a *graphical* approach [10,23]. The impact is represented in the P_n-P_t plane with the following lines as illustrated in Fig. 4:

- Line of limiting friction L_{LF}, defined by Coulomb's friction model.
- Line of sticking L_S, defined by zero velocity along the tangential direction.
- Line of maximum compression L_M, defined by zero relative velocity in the normal direction.
- Line of termination L_T, defined according to the restitution model used.

The graphical construction as in Fig. 4 is completely specified by the following parameters: the geometric conditions of the impact, the inertia properties of the colliding bodies, the velocities before impact, the coefficients of friction and restitution. It is also noted that the use of graphical approach described here assumes Poisson's or Newton's models of restitution.

In Fig. 4, Point A is the intersection between lines L_S and L_T, and point B is the intersection between lines L_S and L_M. The resulting lines OA and OB define three regions that allow to distinguish the different contact modes. In particular, if L_{LF} lies in region #1, slipping without sticking is present; if it lies in region #2, sticking is reached after the maximum compression; if it lies in region #3, sticking is reached before the maximum compression and it may be followed by reverse motion [23]. A detailed analysis of contact processes and their dependence on parameters defining the planar rigid-body collisions is presented in [47]. The main disadvantage of the graphical (and analytical) approaches is that they are not easily extendible to three-dimensional impact, primarily because of the difficulty in describing the limiting friction [10,23].

2.3.2. General formulation and solution

Several formulations of the motion equations have been applied to solve the impact problem between two (or more) rigid bodies with a discrete approach. The more general form of these equations, previously stated in Eq. (5), is:

$$
\begin{aligned}
m_1(\mathbf{v}_1 - \mathbf{v}_{10}) &= \mathbf{P} + \mathbf{P}_1^{\Phi} + \mathbf{P}_1^{\mathrm{F}} \\
m_2(\mathbf{v}_2 - \mathbf{v}_{20}) &= -\mathbf{P} + \mathbf{P}_2^{\Phi} + \mathbf{P}_2^{\mathrm{F}} \\
\mathbf{h}_1 - \mathbf{h}_{10} &= \mathbf{d}_1 \times \mathbf{P} + \mathbf{M} + \mathbf{d}_1^{\Phi} \times \mathbf{P}_1^{\Phi} + \mathbf{d}_1^{\mathrm{F}} \times \mathbf{P}_1^{\mathrm{F}} \\
\mathbf{h}_2 - \mathbf{h}_{20} &= \mathbf{d}_2 \times \mathbf{P} - \mathbf{M} + \mathbf{d}_2^{\Phi} \times \mathbf{P}_2^{\Phi} + \mathbf{d}_2^{\mathrm{F}} \times \mathbf{P}_2^{\mathrm{F}}
\end{aligned} \tag{15}
$$

where $\mathbf{P}^{\Phi}$ is the vector of unknown impulses associated with geometric constraints on the bodies, $\mathbf{P}^{\mathrm{F}}$ is the vector of external known impulses, $\mathbf{d}^{\Phi}$ and $\mathbf{d}^{\mathrm{F}}$ are the position vectors from the center of mass to the point of application of the respective impulses [3]. To solve Eq. (15), we need to resort to additional relations for the impact parameters. Different models for these parameters lead to different forms of the final equations and possibly different results, depending on the particulars of the impact.

2.3.2.1. Algebraic equations. Using Newton's or Poisson's models to define the coefficients of restitution in any direction (or about any axes), purely algebraic equations are obtained. Together with the impulse ratios, these equations can be written in the form:

$$
\mathbf{e} = \mathbf{e}(\mathbf{P}, \mathbf{M}, \mathbf{v}), \quad \boldsymbol{\mu} = \boldsymbol{\mu}(\mathbf{P}, \mathbf{M}) \tag{16}
$$

Examples of analytical solutions of Eqs. (15) and (16) can be found in Brach [3,4], Smith [20] and Mac Sithigh [48]. Lagrange's equations describing impact between two rigid bodies are presented in [16,25]. These formulations solve for unknown generalized coordinates, the Lagrange multipliers associated with impact forces (or normal impulses) and the friction forces due to stiction. This approach has also been applied by some authors to flexible-body systems (see, for example, Kulief and Shabana [49], Yigit et al. [50]). In this case, the coefficient of restitution value for relatively compact bodies must be used with care as it may be affected by the flexibility. Ref. [50] demonstrates that the "rigid body" concept of the coefficient of restitution can be used for the flexible beam considered by these authors.

2.3.2.2. Integral–differential equations. Another approach to solving the impact problem is to think of the impact as an evolving process parameterized by cumulative normal impulse [21,48]. An application of this approach is reported by Keller [21], where Poisson's model of restitution is used. Stronge [51] employes a similar analytical method to investigate changes in relative velocity, but with the use of the energetic coefficient of restitution. The linear impulse $\mathbf{P}$ is divided in the usual two components, normal and tangential, given by:

$$
P_{\mathrm{n}}(t) = \int_0^t \mathbf{F}_{\mathrm{C}}(s) \cdot \mathbf{n}\, \mathrm{d}s = \int_0^t F_{\mathrm{n}}\, \mathrm{d}s, \quad P_{\mathrm{t}} = -\int_0^{t_{\mathrm{f}}} \mu F_{\mathrm{n}} \mathbf{t}\, \mathrm{d}t = -\int_0^{P_{\mathrm{f}}} \mu \mathbf{t}\, \mathrm{d}P_{\mathrm{n}} = -\int_0^{(1+e)P_{\mathrm{m}}} \mu \mathbf{t}\, \mathrm{d}P_{\mathrm{n}} \tag{17}
$$

The normal component P_{n} is used as an independent variable. The solution of the impact problem is reduced to determining P_{m}, as well as the variation in the slip direction specified by the tangential unit vector $\mathbf{t}$. To this end, differentiating the relative linear velocity with respect to the normal impulse P_{n} gives

$$\frac{d\mathbf{C}}{dP_n} = \frac{d\mathbf{C}}{dt}\frac{dt}{dP_n} = \frac{d\mathbf{C}}{dt}\frac{1}{F_n} \tag{18}$$

Eliminating the velocity time derivative by using the impulse–momentum relations, Eq. (15), allows to express the derivative of **C** as a function of the normal and tangential unit vectors, i.e.,

$$\frac{d\mathbf{C}}{dP_n}\cdot\mathbf{n} = \frac{dC_n}{dP_n}(\mathbf{n}, \mu\mathbf{t}), \quad \frac{d\mathbf{C}}{dP_n}\cdot\mathbf{t} = \frac{dC_t}{dP_n}(\mathbf{n}, \mu\mathbf{t}) \tag{19}$$

where we indicated the corresponding functional dependence. Integrating the equation for the normal component C_n and setting the result to zero yields an integral equation for the impulse P_m. The differential equation for the tangential component can be integrated to solve for the tangential impulse defined in Eq. (17).

2.4. Multi-body/multi-contact impact

One disadvantage of the discrete contact modeling is that it is not easily extendible to impacts involving multiple bodies and multiple contact points. In the application of interest to us—constrained robotics operations—simultaneous collision of several bodies is not likely to occur. It is the second situation, where contact occurs between two bodies with complex geometries at multiple points, that poses difficulties for the discrete approach.

Hurmuzlu and Marghitu [22] present a generalization of the integral–differential method described in Section 2.3.2 to a multi-body system. The equations of motion are written as

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{H}(\mathbf{q}, \dot{\mathbf{q}}, t) = \mathbf{J}_C^T\mathbf{F}_C \tag{20}$$

where **M** is the inertia matrix of the system, **H** contains all non-impact forces (gravitational, external, control), **q** is the vector of generalized coordinates, and $\mathbf{F}_C$ denotes impact forces (normal and tangential) at all contact points. Using kinematics, the acceleration at a contact point can be expressed as

$$\mathbf{a} = \frac{d\mathbf{v}}{dt} = \mathbf{H}_1(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{H}_2(\mathbf{q}, \dot{\mathbf{q}}) \tag{21}$$

By substituting for the generalized accelerations from (20), neglecting the non-impulsive terms and assuming that the generalized coordinates do not change during impact, the above is simplified to

$$\mathbf{a} \approx \mathbf{H}_1\mathbf{M}^{-1}\mathbf{J}_C^T\mathbf{F}_C \tag{22}$$

To proceed with the solution, it is now necessary to distinguish between different situations at the contact points: (a) interaction is present along tangential and normal directions; (b) interaction is present only along the normal direction; (c) there is no interaction. Each of these conditions can be represented with additional equations for each contact force. Conditions of sticking, attachment and detachment can be detected by checking velocity and acceleration at the contact point. The end of the compression phase will be characterized by zero normal velocity while the end of restitution phase will be defined according to the restitution model used (any one of the three models can be used). Hurmuzlu and Marghitu [22] also proposed an algebraic method based on

Eq. (15), with additional equations provided by Newton's model of restitution and the definition of impulse ratio. Note that the algebraic approach provides only a post-impact solution defined by the coefficient of restitution. As a result, the possibility of detachment during the impact cannot be detected at the internal contact points.

2.5. Summary of discrete models

As implied by the discussion above, the definition of restitution is a key aspect of the discrete formulation of impact dynamics. Three theories of restitution have been proposed to date and it is appropriate to comment on how they compare against each other. In particular, energetic analysis of a planar impact of two bodies [10,14] leads to the following conclusions:

- The three restitution models are equivalent if there is no friction, or the impact is central, or there is friction but the motion along the tangential direction does not stop (i.e., there is no slip reversal).
- If friction is present and the impact is eccentric, the normal velocity during and at the end of impact depends on the direction of slip. Since Newton's model does not differentiate between the possible contact modes, it neglects the change in the slip direction.
- Poisson's and Stronge's models dissipate more energy than Newton's model, and this energy is always positive, but different for the two when sticking or reverse motion are present.
- Poisson's and Newton's models are inconsistent when $e = 1$, since for this case, they can produce non-zero energy dissipation in the normal direction.

It has been suggested that a way to resolve some of the problems with the existing restitution models is by allowing an interdependency between the coefficients of restitution and friction [3]. Nevertheless, at this time it appears that Stronge's hypothesis of restitution is the better of the three theories.

Another important aspect of discrete models is the unequivocal use of Coulomb's law to model friction during impact. Several authors have noted the inconsistencies that arise when rigid body models are used with Coulomb's empirical law of friction. Examples are described by Wang and Kumar [26] where the aforementioned inconsistencies are demonstrated by either no feasible solution or by multiple solutions for particular initial conditions. This has been attributed to the approximate nature of Coulomb's model and to the inadequacy of rigid body model, but no clear explanation has been found.

Finally, we observe once again that the discrete models are based on the assumption that impact time is small and the bodies involved in the impact are mainly rigid. The use of these models with flexible bodies is not straightforward because of the "rigid body" concept of the coefficient of restitution [50]. However, the results presented in [52,53] for transverse impact of a rotating flexible beam demonstrate relatively little sensitivity to the coefficient of restitution. The application of discrete modelling to *contact* scenarios such as robotic insertion tasks envisioned for the space station is not straightforward. In these cases, the approach velocities are small and there is time-varying contact between the fixture and the mating object at many points. To deal with the multiplicity of contact points would require ad hoc assumptions regarding the order of impulses [11].

3. Continuous contact dynamics models

The continuous model, also referred to as *compliant contact model*, overcomes the problems associated with the discrete models. The basis of the continuous formulation for contact dynamics is to explicitly account for the deformation of the bodies during impact or contact. In a large class of continuous models, referred to in this paper as *explicit* this is done by defining the normal contact force F_n as an explicit function of local indentation δ and its rate, i.e., [2,3]:

$$F_n \equiv F_n(\dot{\delta}, \delta) = F_{\dot{\delta}}(\dot{\delta}) + F_{\delta}(\delta) \tag{23}$$

In the following, we summarize three existing contact force models, including the initial model of Hertz [30] and the non-linear damping model of Hunt and Crossley [32].

3.1. Contact force models

3.1.1. Spring-dashpot model

The impact is schematically represented with a linear damper (dashpot) for the dissipation of energy in parallel with a linear spring for the elastic behavior [9]. The contact force is defined as [2,3]

$$F_n = b\dot{\delta} + k\delta \tag{24}$$

and is represented schematically in Fig. 5. This model has three weaknesses [35]:

- The contact force at the beginning of impact (point A) is discontinuous, because of the damping term. In a more realistic model, both elastic and damping forces should be initially at zero and increase over time.
- As the objects are separating (point B), i.e., the indentation tends to zero, their relative velocity tends to be negative. As a result, a negative force holding the objects together is present.
- The equivalent coefficient of restitution defined for this model does not depend on impact velocity. As we discuss in Section 4, velocity dependence of e has been demonstrated experimentally [9].

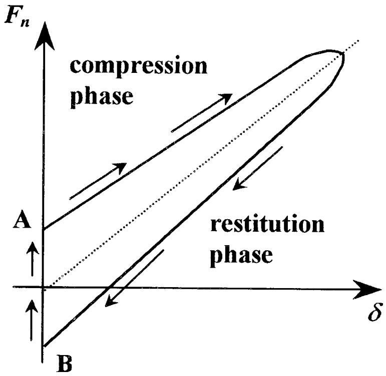

Fig. 5. Contact force history for the spring-dashpot model.

Although the spring-dashpot model is not physically realistic, its simplicity has made it a popular choice [29,39,41,42]. It provides a reasonable method for capturing the energy dissipation associated with the normal forces without explicitly considering plastic deformation issues.

3.1.2. Hertz's model

This is a non-linear model but limited to impacts with elastic deformation and in its original form does not include damping. With this model, the contact process can be pictured as two rigid bodies interacting via a non-linear spring along the line of impact. The hypotheses used states that the deformation is concentrated in the vicinity of the contact area, elastic wave motion is neglected, and the total mass of each body moves with the velocity of its mass centre. The impact force is defined as [2,9,30,54]

$$F_{\mathrm{n}} = k\delta^n \tag{25}$$

where k and n are constants, depending on material and geometric properties and computed by using elastostatic theory. For instance, in the case of two spheres in central impact, $n = 3/2$ and k is defined in terms of Poisson's ratios, Young's moduli and the radii of the two spheres [34]. Since Hertzian model does not account for energy dissipation, its equivalent coefficient of restitution is one. Therefore, as discussed further in Section 4, this model can be used only for low impact speeds and hard materials.

The elastic contact law of Hertz can be augmented to account for plastic deformation by introducing hysteresis in the contact force law. This can be accomplished by using a different force-indentation relationship for the unloading phase of the contact which generally takes the following form [15,55]:

$$F_{\mathrm{n}} = F_{\mathrm{n,max}}\left(\frac{\delta - \delta_{\mathrm{p}}}{\delta_{\mathrm{max}} - \delta_{\mathrm{p}}}\right)^n \tag{26}$$

In the above, $F_{\mathrm{n,max}}$ and δ_{max} are the maximum normal force and indentation reached during the loading phase and δ_{p} is the permanent indentation. Note that in the context of contact dynamics simulation, the maximum quantities in Eq. (26) can be calculated at every instant of the numerical integration. By contrast, the value of δ_{p} must be specified as an additional parameter in this contact force model. The hysteretic force law of Eq. (26) has not been previously applied to solve multi-body contact scenarios, partly because it is somewhat cumbersome to implement and the plastic deformation per se is unimportant in the majority of applications.

3.1.3. Non-linear damping

To overcome the problems of the spring-dashpot model and to retain the advantages of the Hertz's model, an alternative model for energy dissipation was introduced by Hunt and Crossley [32]. It includes a non-linear damping term and hence the impact/contact force is modeled as

$$F_{\mathrm{n}} = b\delta^p\dot{\delta}^q + k\delta^n \tag{27}$$

where it is standard to set $p = n$ and $q = 1$ [32,34,35]. As with the spring-dashpot model, the damping parameter b can be related to the coefficient of restitution, since both are related to the energy dissipated by the impact process. For the central impact of two bodies Hunt and Crossley [32], Lankarani and Nikravesh [34], and Marhefka and Orin [35] established:

$$e = 1 - \alpha\dot{\delta}_0, \quad \alpha = \frac{2b}{3k} \tag{28}$$

An important aspect of this model is that damping depends on the indentation. This is physically sound since contact area increases with deformation and a plastic region is more likely to develop for larger indentations. Another advantage is that the contact force has no discontinuities at initial contact and separation, but it begins and finishes with the correct value of zero. This model has been studied and used by several authors [32–35,38–41,56,57].

3.2. Friction models

It is apparent from our comments in Section 2 that Coulomb's law is frequently used to describe the friction phenomenon for impact problems. The main problem with Coulomb's law is the discontinuity of the friction force due to the difference between static and dynamic behaviors. For this reason and to capture other aspects of frictional interaction, alternative friction laws have been proposed. One improvement to Coulomb's law is obtained by using a *non-local* model of friction [36] where the value of friction at one point (in space) depends on the value of certain quantities in the neighborhood of that point. Another improvement is to use a non-linear model to permit a continuous transition from the phase of sticking to sliding [29,36,37,57,58].

A model worth mentioning is the *bristle* model [37,38], which represents the effect of surface irregularities using bristles. The friction force is defined as

$$\mathbf{F}_t = k_f\mathbf{s}, \quad \mathbf{s}(t) = \begin{cases} \mathbf{s}(t_0) + \int_{t_0}^{t} \mathbf{v}_t \, dt, & \text{if } |\mathbf{s}| < s_{max}, \\ s_{max}\frac{\mathbf{v}_t}{|\mathbf{v}_t|}, & \text{otherwise}, \end{cases} \quad s_{max} = \mu\frac{|F_n|}{k_f} \tag{29}$$

where k_f is the bristle stiffness, $\mathbf{s}$ the vector of bristle displacement, t_0 is the start time of the last sticking at that contact point, $\mathbf{v}_t$ the relative tangential velocity and parameter s_{max} is the maximum allowable deflection of the bristle. Dynamic friction models such as the bristle model above are particularly suited to continuous contact dynamics modeling as they effectively calculate the friction force as a function of time (through dependence on $\mathbf{s}$ or F_n). Most importantly, the friction force is defined *explicitly* and *uniquely* during sticking at the contact point.

Although not presented as a friction model, Stronge [15] uses the concept of tangential compliance to evaluate the tangential force at the contact point during sticking. Using our notation,

$$F_t = k_t\delta_t \tag{30}$$

where δ_t is the tangential component of displacement of a massless particle at the contact point and k_t is the tangential stiffness. Differently from the bristle model, δ_t is evaluated as a solution to simple harmonic motion in the tangential direction.

It is not the scope of the present paper to review in detail the friction models available in literature. The additional complexity of modern friction models comes at a price of a larger number of parameters required to define the model, and a less intuitive connection to the actual physics. Of importance to us is the fact that continuous formulation of impact dynamics allows the user to easily implement any one of the available friction models. By contrast, only Coulomb's law has been used in the discrete formulations of impact dynamics. The reason for this lies in the fact that discrete impact models are based on the impulse–momentum formulation and it is only for the Coulomb's law of friction that one can easily extend the friction force relationship to the impulse domain.

3.3. Formulation and solution

A generic unconstrained system is defined by the vector of generalized coordinates $\mathbf{q}$, including the coordinates used to discretize the flexible bodies [41]. With regard to the latter, we point out that modeling and discretization issues require special consideration when the impacting bodies have significant structural flexibility. We direct our readers to the relevant work by Shabana [52,53], Stronge [11] and Abrate [55,59]. The dynamics equations of a multi-body system in free (unconstrained) motion have the form

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{H}(\mathbf{q}, \dot{\mathbf{q}}, t) = 0 \tag{31}$$

To model the constrained motion of the system, it is convenient to separate the vector of generalized coordinates in two subvectors, $(\mathbf{q}_\mathrm{F}, \mathbf{q}_\mathrm{C})$, where $\mathbf{q}_\mathrm{C}$ contains the coordinates constrained by the contacts, and $\mathbf{q}_\mathrm{F}$ contains the free coordinates.

3.3.1. Rigid impact formulation

We will start by assuming the impact is rigid, that is similarly to the discrete approach, there is no deformation associated with the impact, and furthermore, there is no account of energy dissipation. Such a model can be valid in some situations, for example when the surfaces in contact are very stiff [41]. Letting $\mathbf{F}_\mathrm{C}$ represent the vector of contact normal forces [3] the motion equations for a multi-body system in contact become

$$\begin{aligned} &\mathbf{M}\ddot{\mathbf{q}} + \mathbf{H} + \left\{ \begin{array}{c} \mathbf{0} \\ \mathbf{J}_\mathrm{C}^\mathrm{T}\mathbf{F}_\mathrm{C} \end{array} \right\} = 0 \\ &\mathbf{\Phi}_\mathrm{C}(\mathbf{q}_\mathrm{C}, t) = 0, \quad \mathbf{J}_\mathrm{C} = \frac{\partial \mathbf{\Phi}_\mathrm{C}}{\partial \mathbf{q}_\mathrm{C}} \end{aligned} \tag{32}$$

where $\mathbf{J}_\mathrm{C}$ is the constraint Jacobian, and $\mathbf{\Phi}_\mathrm{C}$ represents the constraints associated with the contact conditions, in particular that the relative distance at the contact point is zero. It is further appropriate [4] to separate the vector of constrained generalized coordinates as $(\mathbf{q}_\mathrm{C}^\mathrm{sl}, \mathbf{q}_\mathrm{C}^\mathrm{st})$where $\mathbf{q}_\mathrm{C}^\mathrm{sl}$ contains coordinates associated with slipping contact constraint, while $\mathbf{q}_\mathrm{C}^\mathrm{st}$ contains those associated with the sticking constraint. In case of slipping, in the most general case, friction is defined as a function of contact forces, coordinates and their rates:

$$\mathbf{F}_\mathrm{t}^\mathrm{sl} = \mathbf{F}_\mathrm{t}^\mathrm{sl}(\mathbf{F}_\mathrm{C}, \mathbf{q}_\mathrm{C}^\mathrm{sl}, \dot{\mathbf{q}}_\mathrm{C}^\mathrm{sl}) \tag{33}$$

In case of sticking, the corresponding friction force $\mathbf{F}_\mathrm{t}^\mathrm{st}$ is unknown and the sticking constraints must be included in the formulation. Thus, the system equations become

[3] In Section 2, $\mathbf{F}_\mathrm{c}$ was used to denote the total contact force, including the normal and tangential components. The presentation here is made easier by explicitly separating these forces, however, we chose to keep the same symbol for the sake of convenience.

[4] This discussion assumes that the static friction force is not explicitly defined but is unknown.

$$\mathbf{M}\ddot{\mathbf{q}} + \mathbf{H} + \left\{ \begin{array}{c} \mathbf{0} \\ \mathbf{J}_C^T \mathbf{F}_C \end{array} \right\} + \left\{ \begin{array}{c} \mathbf{0} \\ \mathbf{F}_t^{sl} \\ \mathbf{J}_C^{st^T} \mathbf{F}_t^{st} \end{array} \right\} = 0$$
$$\mathbf{\Phi}_C(\mathbf{q}_C, t) = 0, \quad \mathbf{J}_C = \frac{\partial \mathbf{\Phi}_C}{\partial \mathbf{q}_C} \tag{34}$$
$$\mathbf{\Phi}_t^{st}\left(\dot{\mathbf{q}}_C^{st}\right) = 0, \quad \mathbf{J}_C^{st} = \frac{\partial \mathbf{\Phi}_t^{st}}{\partial \mathbf{q}_C^{st}}$$

where $\mathbf{\Phi}_t^{st}$ and $\mathbf{J}_C^{st}$ are associated with sticking constraints [41]. The unknowns are the generalized coordinates $\mathbf{q}$, the contact (normal) forces $\mathbf{F}_C$, and the friction forces $\mathbf{F}_t^{st}$ and are calculated by solving the differential–algebraic system (34).

The assumptions of the rigid impact model—no deformation and no energy dissipation—are not likely to be valid for impacts involving high impact forces, although surface and material properties play a role as well. Furthermore, it may be impractical to apply the rigid impact analysis to scenarios involving complex geometries where multiple contacts need to be monitored during simulation.

3.3.2. Explicit continuous formulation

A generic contact model can be expressed as an *explicit* functional relationship between the contact force and the generalized coordinates and their time derivatives [39,41], i.e.,

$$\mathbf{F}_C = \mathbf{F}_C(\mathbf{q}_C, \dot{\mathbf{q}}_C) \tag{35}$$

with additional dependencies on certain geometric and material parameters. Then, the geometric contact constraints defined by the vector $\mathbf{\Phi}_C$ in the system (34) can be dropped. The motion is now "constrained" by the elastodynamic forces defined by Eq. (35). Note that a check of the contact constraints is still necessary to determine when contact begins or ends. This is an important aspect of the compliant impact dynamics formulation and is discussed in Section 3.4.

3.3.3. Implicit continuous formulation

This solution to contact dynamics problem is distinguished here because it implicitly allows deformation due to contact, however, it does not explicitly evaluate the relationship between contact force and deformation. This method is typically used in the finite element formulations of contact problems, as in [5,44–46]. The system of equations takes the same general form as in Eq. (34) except that the generalized coordinates must include the elastic coordinates associated with the discretization of the contact region. The constraints enforce the condition of impenetrability at the contact nodes or elements.

3.4. Contact detection and interference calculation

A general algorithm to solve the impact/contact problem must address the problem of detecting contact. The contact condition of *real* impacts between bodies is that no material overlap can occur, that is a condition of impenetrability [5]. One way to determine if contact is present is by checking the minimum distance between bodies—the contact is declared if this distance decreases to zero. Many algorithms for minimum distance calculation and contact detection have been proposed in literature [60]. One approach involves solution of a constrained optimization problem of the form:

$$\min\left\{\frac{1}{2}(\mathbf{p}_1-\mathbf{p}_2)^{\mathrm{T}}(\mathbf{p}_1-\mathbf{p}_2)\right\},\quad \begin{array}{l}\mathbf{g}_1(\mathbf{p}_1)\leqslant 0\\ \mathbf{g}_2(\mathbf{p}_2)\leqslant 0\end{array} \tag{36}$$

where $\mathbf{p}_1$ and $\mathbf{p}_2$ are the position vectors of two points on the two bodies and $\mathbf{g}_1$ and $\mathbf{g}_2$ are the bounding surface constraints. The contact point is the point where the minimum distance is zero. Examples of application of this approach can be found in [5,38,46].

The explicit formulation of the continuous contact dynamics requires evaluation of the contact force, which in turn depends on the deformation associated with the contact. Practical implementations of the explicit contact force model have used the concept of *interference* distance or *penetration* in order to define the contact force. The interference distance calculation can also be formulated as an optimization problem [38,60]:

$$\min\{-d\},\quad \begin{array}{l}\mathbf{g}_1(\mathbf{p}_1)\leqslant -\frac{d}{2}\mathbf{e}_1\\ \mathbf{g}_2(\mathbf{p}_2)\leqslant -\frac{d}{2}\mathbf{e}_2\end{array} \tag{37}$$

where $\mathbf{e}_1$ and $\mathbf{e}_2$ are vectors containing 1's and d is the interference distance. The optimization approach provides a general and robust solution to minimum distance and interference problems.

3.5. Solution of contact dynamics equations

The solution procedure must be tailored to the particular formulation of the contact dynamics equations. Three general solution methodologies can be defined. The first involves solving a system of differential–algebraic equations, such as that defined by Eq. (34). Although several algorithms are available for solution of such systems, they are less advanced than methods for solving ordinary differential equations. Alternatively, the explicit continuous model can be solved as a system of ordinary differential equations with a number of algorithms for initial-value problems. Another method is the *penalty method* [16,40] where constraints are included in the motion equations with a penalty parameter. The main weakness of this method is the ill-conditioning which worsens as the penalty value is increased. Other problems are that the solution strongly depends on the particular choice of the penalty parameter and that it violates the work-energy balance of the system since some of the energy is stored in the penalty term.

4. Experimental verification

The complexity of impact dynamics requires verification through experiments. Such a validation can be conducted at two levels, the first being a validation of the *basic* theories of restitution or contact force models, while the second is a validation of the overall contact dynamics simulation. As well, experimental measurements or other means are required to determine parameters characterizing the impact. In the case of discrete models, this involves determination of the coefficient of restitution, while for continuous models, at least contact stiffness and damping are needed. Both modelling approaches also use one or several parameters to define the friction model—it is usually the coefficient of friction.

Much of the experimental work related to impact/contact dynamics has focused on measuring model parameters and verifying the contact force models. The main results of this research are summarized in the following sections. We do not review the literature dealing with the validation of the complete contact dynamics models, but direct the interested readers to Van Vliet, Sharf and Ma [58].

4.1. Coefficient of restitution

The principal and most general study of this subject is due to Goldsmith [9], who measured the displacements of the impacting bodies, the duration of impact, the geometry of the crater and the stress waves generated by the impact. Using the initial impact velocity and the measured displacements, the final (post-impact) velocity was calculated and used to compute the coefficient of restitution. Fig. 6 shows qualitatively the relationship between the coefficient of restitution and impact velocity for a central direct impact of two spheres. The experimental data were used to find the dependencies between different quantities, such as coefficient of restitution and the initial impact velocity, as well as to check the limits of the known impact theories. The results obtained for the coefficient of restitution clearly demonstrated the dependence of this parameter on the geometry and material of the impacting bodies, as well as the initial impact geometry and velocity. More specifically, the coefficient of restitution decreases with the increase of the initial impact velocity, and for most materials, it is significantly smaller than unity, even at very low impact speeds. This implies that Hertz's theory [30] of perfectly elastic impact is not valid in most impact situations and that some plastic deformation always takes place. As a conclusion, Goldsmith [9] states that Hertz's theory provides a good description for impact of two spheres or a sphere and a plate, if the materials are hard and the initial speed is low. Under more general conditions, plastic deformation and/or energy loss associated with wave propagation should be taken into account. As shown in Fig. 6, the coefficient of restitution tends to level out with increase in impact velocity—the fact demonstrated analytically by Brach [3].

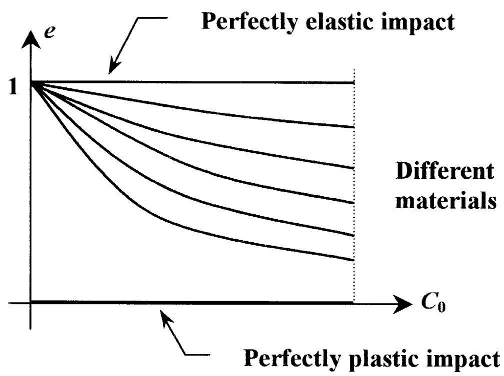

Fig. 6. Coefficient of restitution as a function of approach velocity.

Based on the studies reported by Goldsmith, it is possible to establish the following relationship:

$$e = 1 - \alpha\dot{\delta}_0^m \tag{38}$$

where parameter α and the exponent m are dependent on the material and geometry. Other researchers have measured e for specific contact situations, such as the collision of steel bars with external surfaces [56], impact of an unconstrained slender bar with a massive surface [43], impacts of a steel sphere moving between two plates [61], and the oblique impact between a sphere and a plate [62]. All experiments demonstrate that coefficient of restitution depends on many properties and impact characteristics, in addition to the material properties of impacting bodies. It is therefore difficult to use this parameter to model complex impact scenarios with multiple-point contacts.

4.2. Contact stiffness and damping

Contact stiffness and damping, also known as contact parameters, are used to define the contact force law in the explicit continuous formulation of impact dynamics. Although this model is conceptually simple, the physical meaning of contact parameters is not obvious and thus, it is not straightforward to define their values for complex contact scenarios. This is particularly true for robotic insertions where the contacting bodies—parts to be mated—are subcomponents of a multi-body chain, which itself may comprise many compliant and dissipative elements. In such situations, the characteristics of impact are very much determined by the relative compliance and damping of local (contact) regions, the structural flexibility and damping of the contacting bodies and/or other constituents of the system. Nevertheless, for simple impact geometries, one can use the analytical approach to estimate the contact parameters. For example, one can calculate the contact stiffness by applying Hertz' theory of contact [54]. For contact damping, one can use the energy-balance to find a relationship between contact damping and the coefficient of restitution [32,34,35], although, inevitably, such a result is limited by the accuracy of the coefficient of restitution.

An alternative estimation of contact stiffness was used in Van Vliet et al. [58] for a relatively complex case of a peg contacting a hole wall. There, the contact stiffness was calculated by using the stiffness of the most compliant element of the wall—the load cells between the hole sides and the fixed support. Possibly the most practical approach is to determine the contact parameters by direct experimental measurements. Some research in this area has been carried out in the robotics community where several identification algorithms have been developed to estimate the stiffness of the robot environment [63]. As done by the authors of [56], one can always tune the parameter values used in the numerical simulation to achieve an agreement with the experimental results.

4.3. Tangential coefficients

Extension of experiments to oblique impacts allows a verification of tangential models, in particular, Coulomb's model of friction as well as tangential compliance. The former is effectively done by measuring the coefficient of friction for different impact geometries and demonstrating that it is approximately constant [56,62]. Results, qualitatively displayed in Fig. 7, indicate the

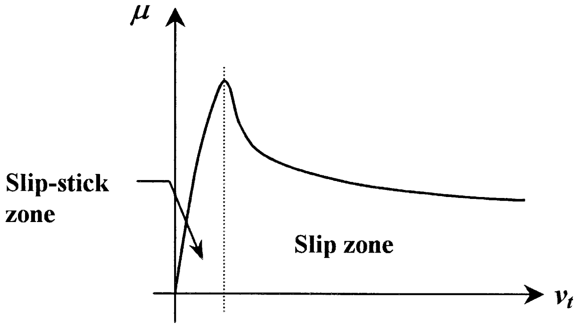

Fig. 7. Impulse ratio as a function of initial tangential velocity.

presence of two main zones. In the first zone, the initial tangential velocity is low (i.e., the angle of incidence is high) and the impulse ratio increases rapidly. In this zone, both slipping and sticking is possible. In the second zone, the tangential velocity is high (i.e., the angle of incidence is low). For any tangential velocity in this zone, only slipping is present until the termination of impact. In this zone, the coefficient of friction decreases slowly with increase in tangential velocity.

4.4. Validation of contact force models

It is difficult to make general statements on the validity of discrete impact dynamics models. It is commonly said that they depend on the applicability of the rigid-body hypothesis, more specifically, whether it is reasonable to neglect deformation at the contact point. Discrete models have been used, however, to predict the impact dynamics of flexible bodies, such as flexible beams [49,50,52,53]. The strong dependency of the energy lost during impact on vibrations of the contacting bodies has been confirmed. Although the discrete formulation is capable of capturing the energy loss—through the coefficient of restitution—this parameter depends on many factors and it is difficult to use for general impact scenarios.

If the rigid body hypothesis is not applicable, a continuous contact model can be used. The non-linear contact force model has been validated experimentally by using flexible beams, as in [43,56,64]. It was shown that the dynamic behavior of the system is not sensitive to the value of the damping coefficient over a wide range. Comparison of simulated and measured velocities of the contact point shows good agreement, especially for low speed impact.

5. Conclusions

In this paper, we reviewed the state of the art on the subject of impact and contact dynamics modeling. All models were classified in two categories: *impulse–momentum* or *discrete* and *continuous*. In the former, the impact analysis is divided into discrete events and energy dissipation is accounted for via coefficient of restitution and the impulse ratio. In the continuous approach, the dynamics analysis is conducted continuously, by admitting (explicitly or implicitly) a relationship between contact force and deformation.

The discrete formulation has been applied mainly to rigid-body collisions since its main premise is the rigid-body hypothesis which imposes a limit on the energy loss. Definition of the coefficient of restitution is a key aspect of the discrete approach and three such definitions exist in literature. Closed-form solutions for simple impact geometries demonstrate that under general impact conditions, the three models of restitution do not produce the same results. Moreover, Newton's and Poisson's models may produce solutions which are energetically inconsistent. Naturally, predictions of the discrete impact analysis depend on the accuracy of the coefficient of restitution. However, experiments have proven that this parameter depends on many impact characteristics, which makes accurate estimation very difficult. The use of the coefficient of restitution for impacts involving flexible bodies is precarious.

Various solution methods to solve the discrete impact dynamics equations have been presented, as well as the generalization of the methodology to multi-body systems subject to multiple contacts. It was noted that the discrete formulation is not easily extendible to handle general impact scenarios, more specifically those where impacts occur at many points in the system. Finally, the necessary use of Coulomb's law with the discrete approach may lead to inconsistencies or multiple solutions.

The continuous approach has several advantages over the discrete formulation. Very importantly, it does not require one to differentiate between impact and contact situations and permits the use of solution methods employed for non-impact dynamics problems. The approach extends quite naturally to contact scenarios with multiple bodies and/or multiple points of contact. The added complexity of the minimum distance/interference determination problem seems minimal when compared to making ad hoc assumptions on the impulse histories at different contact locations. Different models for contact force were presented, and as noted by several researchers and verified experimentally, the model with a non-linear damping term represents quite well the real behavior of the system during impact. However, it may require tuning contact parameter values since for impacts involving multi-body systems, their values are not easily obtained. Unlike the discrete formulation, the continuous approach allows for use of *any* friction model. Two principal solution methods were identified in the context of continuous formulation: explicit and implicit. The latter is typically used in conjunction with a finite-element discretization of the contacting bodies.

In our final overall consideration of the different approaches, we conclude that the continuous formulation combined with an implicit solution for the contact forces appears to be the best for analyzing flexible multi-body systems subject to multiple impacts and contacts. This procedure is most general, does not involve approximations or assumptions regarding the impact process and does not require additional contact parameters.

Acknowledgements

Funding support from the National Science and Engineering Research Council (NSERC) and McDonald Dettwiler Space and Advanced Robotics Limited, through the Collaborative Research grant #222803-98, is acknowledged. Technical support from Dr. Ou Ma of McDonald Dettwiler Space and Advanced Robotics Limited is also acknowledged.

References

[1] W.F. Riley, L.D. Sturges, Engineering Mechanics Dynamics, John Wiley and Sons, New York, 1996.
[2] P. Barkan, Impact design, in: Mechanical Design and Systems Handbook, Section 31, McGraw-Hill, New York, 1974.
[3] R.M. Brach, Mechanical Impact Dynamics: Rigid Body Collisions, John Wiley and Sons, New York, 1991.
[4] R.M. Brach, Formulation of rigid body impact problems using generalized coefficients, International Journal of Engineering Science 36 (1) (1998) 61–71.
[5] S.W. Kim, Contact Dynamics and Force Control of Flexible Multi-Body Systems, Ph.D. Thesis, Department of Mechanical Engineering, McGill University, Montreal, 1999.
[6] M. Ciavarella, Generalized cattaneo partial slip plane contact problem I—Theory, International Journal of Solids and Structures 35 (18) (1998) 2349–4362.
[7] M. Ciavarella, Generalized cattaneo partial slip plane contact problem II—Examples, International Journal of Solids and Structures 35 (18) (1998) 2363–2378.
[8] J. Jaeger, Analytical solutions of contact impact problems, Applied Mechanics Reviews 47 (2) (1994) 35–54.
[9] W. Goldsmith, Impact: The Theory and Physical Behavior of Colliding Solids, Edward Arnold Publishers Ltd, London, 1960.
[10] Yu. Wang, M.T. Mason, Two dimensional rigid-body collisions with friction, Journal of Applied Mechanics 59 (1992) 635–642.
[11] W.J. Stronge, Impact Mechanics, Cambridge University Press, Cambridge, 2000.
[12] E.T. Whittaker, A Treatise on the Analytical Dynamics of Particles and Rigid Bodies, Cambridge University Press, Cambridge, 1904.
[13] E.T. Routh, Dynamics of a System of Rigid Bodies, MacMillan & Co, London, 1905.
[14] W.J. Stronge, Unraveling paradoxical theories for rigid body collisions, Journal of Applied Mechanics 58 (1991) 1049–1055.
[15] C.T. Lim, W.J. Stronge, Oblique elastic–plastic impact between rough cylinders in plane strain, International Journal of Engineering Science 37 (1) (1999) 97–122.
[16] J.G. De Jalon, E. Bayo, Kinematic and Dynamic Simulation of Multi-body Systems—The Real Time Challenge, Mechanical Engineering Series, Springer-Verlag, New York, 1993.
[17] W.J. Stronge, Energy dissipated in planar collision, Journal of Applied Mechanics 59 (3) (1992) 681–682.
[18] T.R. Kane, D.A. Levinson, Dynamics: Theory and Applications, Series in Mechanical Engineering, McGraw-Hill, New York, 1985.
[19] I. Newton, Philosophia Naturalis Principia Mathematica, S. Pepsy, Reg. Soc, London, 1686.
[20] C.E. Smith, Predicting rebounds using rigid-body dynamics, Journal of Applied Mechanics 58 (1991) 754–758.
[21] J.B. Keller, Impact with friction, Journal of Applied Mechanics 53 (1986) 1–4.
[22] Y. Hurmuzlu, D.B. Marghitu, Rigid body collision of planar kinematic chains with multiple contact points, The International Journal of Robotics Research 13 (1) (1994) 82–92.
[23] I. Han, B.J. Gilmore, Multi-body impact motion with friction—analysis, simulation and experimental validation, Journal of Mechanical Design 115 (1993) 412–422.
[24] B.J. Gilmore, R.J. Cipra, Simulation of planar dynamic mechanical systems with changing topologies—Part 1: Characterization and prediction of the kinematic constraint changes, Journal of Mechanical Design 113 (1991) 70–76.
[25] E.J. Haug, S.C. Wu, S.M. Yang, Dynamics of mechanical systems with coulomb friction, stiction, impact and constraint addition–deletion—I, Mechanism and Machine Theory 21 (5) (1986) 401–406.
[26] Y.T. Wang, V. Kumar, Simulation of mechanical systems with multiple frictional contacts, Journal of Mechanical Design 116 (1994) 571–580.
[27] M. Anitescu, J.F. Cremer, F.A. Potra, Formulating three dimensional contact dynamics problems, Mechanics of Structures and Machines 24 (4) (1996) 405–437.
[28] M.T. Mason, Y. Wang, On the inconsistency of rigid-body frictional planar mechanics, IEEE International Conference on Robotics and Automation 1 (1988) 524–528.

[29] W.S. Howard, V. Kumar, A minimum principle for the dynamic analysis of systems with frictional contacts, IEEE International Conference on Robotics and Automation 1 (1993) 437–442.
[30] H. Hertz, Miscellaneous papers, in: H. Hertz, Jones and Schott, London, 1896.
[31] S. Dubowsky, F. Freudenstein, Dynamic analysis of mechanical systems with clearances—Part 1: Formation of dynamical model, Journal of Engineering for Industry 93 (1971) 305–309.
[32] K.H. Hunt, F.R.E. Crossley, Coefficient of restitution interpreted as damping in vibroimpact, Journal of Applied Mechanics 42, Series E (1975) 440–445.
[33] T.W. Lee, A.C. Wang, On the dynamics of intermittent-motion mechanisms, Part 1: Dynamic model and response, Journal of Mechanisms, Transmissions, and Automation in Design 105 (1983) 534–540.
[34] H.M. Lankarani, P.E. Nikravesh, A contact force model with hysteresis damping for impact analysis of multi-body systems, Journal of Mechanical Design 112 (1990) 369–376.
[35] D.W. Marhefka, D.E. Orin, A compliant contact model with nonlinear damping for simulation of robotic systems, IEEE Transactions on Systems, man, and Cybernetics—Part A: Systems and Humans 29 (6) (1999) 566–572.
[36] J.T. Oden, E.B. Pires, Non-local and non-linear friction laws and variational principles for contact problems in elasticity, Journal of Applied Mechanics 50 (1983) 67–76.
[37] D.A. Haessig, B. Friedland, On the modeling and simulation of friction, Journal of Dynamic Systems, Measurement, and Control 113 (1991) 354–362.
[38] O. Ma, Contact dynamics modeling for the simulation of the space station manipulators handling payloads, IEEE International Conference on Robotics and Automation 2 (1995) 1252–1258.
[39] P.R. Kraus, V. Kumar, Compliant contact models for rigid body collisions, IEEE International Conference on Robotics and Automation 2 (1987) 1382–1387.
[40] A. Deguet, A. Joukhadar, C. Laugier, A collision model for deformable bodies, IEEE International Conference on Robotics and Automation 1 (1998) 636–641.
[41] M.K. Vukobratovic, V. Potkonjak, Dynamics of contact tasks in robotics. Part I: General model of robot interacting with environment, Mechanism and Machine Theory 34 (1999) 923–942.
[42] K. Mirza, M.D. Hanes, D.E. Orin, Dynamic simulation of enveloping power grasps, IEEE International Conference on Robotics and Automation 1 (1993) 430–435.
[43] D.B. Marghitu, The impact of flexible links with solid lubrification, Journal of Sound and Vibrations 205 (1997) 712–720.
[44] K.J. Bathe, P.A. Bouzinov, On the constraint function method for contact problems, Computers and Structures 64 (5/6) (1997) 1069–1085.
[45] K. Farahani, M. Mofid, A. Vafai, A solution method for general contact-impact problems, Computer Methods in Applied Mechanics and Engineering 187 (2000) 69–77.
[46] M.W. Heinstein, F.J. Mello, S.W. Attaway, T.A. Laursen, Contact-impact modeling in explicit transient dynamics, Computer Methods in Applied Mechanics and Engineering 187 (2000) 621–640.
[47] W.J. Stronge, Theoretical coefficient of restitution for planar impact of rough elasto-plastic bodies, American Society of Mechanical Engineers, Applied Mechanics Division AMD 205 (1995) 351–362.
[48] G.P. Mac Sithigh, Rigid-body impact with friction—various approaches compared, in: Impact Mechanics: Experiment, Theory and Calculation, ASME-AMD, vol. 205, 1995, pp. 307–317.
[49] Y.A. Kulief, A.A Shabana, Impact responses of multi-body systems with consistent and lumped masses, Journal of Sound and Vibration 104 (2) (1986) 187–207.
[50] A.S. Yigit, A.G. Ulsoy, R.A. Scott, Dynamics of a radially rotating beam with impact, Part 1: Theoretical and computational model, Journal of Vibration and Acoustics 112 (1990) 65–70.
[51] W.J. Stronge, Swerve during three-dimensional impact of rough rigid bodies, Journal of Applied Mechanics 61 (3) (1994) 605–611.
[52] H. Palas, W.C. Hsu, A.A. Shabana, On the use of momentum balance and the assumed modes methods in transverse impact problems, Journal of Vibration and Acoustics 114 (3) (1992) 364–373.
[53] K.H. Hwang, A.A. Shabana, Effect of mass capture on the propagation of transverse waves in rotating beams, Journal of Sound and Vibration 186 (3) (1995) 495–525.
[54] K.L. Johnson, Contact Mechanics, Cambridge University Press, Cambridge, 1985.
[55] S. Abrate, Impact on Composite Structures, Cambridge University Press, Cambridge, 1998.

[56] D. Stoianovici, Y. Hurmuzlu, A critical study of the applicability of rigid-body collision theory, Journal of Applied Mechanics 63 (1996) 307–316.
[57] A.S. Yigit, A.G. Ulsoy, R.A. Scott, Spring-dashpot models for the dynamics of a radially rotating beam with impact, Journal of Sound and Vibrations 142 (1990) 515–525.
[58] J. Van Vliet, I. Sharf, O. Ma, Experimental validation of contact dynamics simulation of constrained robotic tasks, The International Journal of Robotics Research 19 (12) (2000) 1203–1217.
[59] S. Abrate, Modeling of impacts on composite structures, Composite Structures 51 (2) (2001) 129–138.
[60] I. Sharf, M. Nahon, Interference distance calculation for two objects bounded by quadratic surfaces, ASME Design Engineering Technical Conference 1 (1995) 634–641.
[61] M.A. Veluswami, F.R.E. Crossley, Multiple impacts of a ball between two plates, Journal of Engineering for Industry 97B (1975) 820–827.
[62] J.P. Nobre, A.M. Dias, R. Gras, A study on elasto-plastic impact friction, Wear 230 (1999) 133–145.
[63] D.A. Erickson, Contact stiffness of damping estimation for constrained robotic system, M.Sc. Thesis, Department of Mechanical Engineering, University of Victoria, Victoria, 2000.
[64] A.S. Yigit, A.G. Ulsoy, R.A. Scott, Dynamics of a radially rotating beam with impact, Part 2: Experimental and simulation results, Journal of Vibrations and Acoustics 112 (1990) 71–77.