Chapter 3
Contact-Impact Force Models for Mechanical Systems

The collision is a prominent phenomenon in many mechanical systems that involve intermittent motion, kinematic discontinuities or clearance joints. As a result of an impact, the values of the system state variables change very fast, eventually looking like discontinuities in the system velocities and accelerations. The impact is characterized by large forces that are applied and removed in a short time period. The knowledge of the peak forces developed in the impact process is very important for the dynamic analysis of multibody mechanical systems having consequences in the design process. The numerical description of the collision phenomenon is strongly dependent on the contact-impact force model used to represent the interaction between the system components. The model for the contact-impact force must consider the material and geometric properties of the colliding surfaces and information on relative positions and velocities, contribute to an efficient integration and account for some level of energy dissipation. These characteristics are ensured with a continuous contact force model, in which the deformation and contact forces are considered as continuous functions during the complete period of contact. This chapter deals with contact-impact force models for both spherical and cylindrical contact surfaces. The incorporation of the friction phenomenon, based on the Coulomb's friction law, is also discussed together with a computational strategy which includes an automatic step size selection procedure based not only on numerical error control but also on the characteristics of the contact.

3.1 Approaches to Contact and Impact of Rigid Bodies

Impact occurs during the collision of two or more bodies, which may be external or belong to a multibody mechanical system. The impact phenomenon is characterized by abrupt changes in the values of system variables, most visible as discontinuities in the system velocities. Other effects directly related to the impact phenomena are the vibration propagation on the system components, local elastic/plastic deformations at the contact zone and some level of energy dissipation. The impact is a very important phenomenon in many mechanical systems such as mechanisms with intermittent motion and with clearance joints (Khulief and Shabana 1986, Ravn 1998).

P. Flores et al., *Kinematics and Dynamics of Multibody Systems with Imperfect Joints.*

The selection of the most adequate contact force model plays a key role in the correct design and analysis of these types of mechanical systems (Flores et al. 2006).

By and large, an impact may be considered to occur in two phases: the compression or loading phase and the restitution or unloading phase. During the compression phase, the two bodies deform in the normal direction to the impact surfaces, and the relative velocity of the contact points/surfaces on the two bodies in that direction is gradually reduced to zero. The end of the compression phase is referred to as the instant of maximum compression or maximum approach. The restitution phase starts at this point and ends when the two bodies separate from each other (Brach 1991). The restitution coefficient reflects the type of collision. For a fully elastic contact the restitution coefficient is equal to the unit, while for a fully plastic contact restitution coefficient is null. The most general and predominant type of collision is the oblique eccentric collision, which involves both relative normal velocity and relative tangential velocity (Maw et al. 1975, Zukas et al. 1982).

In order to evaluate efficiently the contact-impact forces resulting from collisions in multibody systems, such as the contact between the bearing and journal in a revolute joint with clearance, special attention must be given to the numerical description of the contact force model. Information on the impact velocity, material properties of the colliding bodies and geometry characteristics of the contact surfaces must be included in the contact force model. These characteristics are observed with a continuous contact force, in which the deformation and contact forces are considered as continuous functions (Lankarani and Nikravesh 1990). Furthermore it is important that the contact force model can add to the stable integration of the equation of motion of multibody system.

In a broad sense, there are two different methods to solve the impact problem in multibody mechanical systems designated as continuous and discontinuous approaches (Lankarani and Nikravesh 1990). Within the continuous approach the methods commonly used are the continuous force model, which is in fact a penalty method, and the unilateral constraint methodology, based on the linear complementary approach (Pfeiffer and Glocker 1996). The continuous contact force model represents the forces arising from collisions and assumes that the forces and deformations vary in a continuous manner. In this method, when contact between the bodies is detected, a normal force perpendicular to the plane of collision is applied. This force is typically applied as a spring–damper element, which can be linear, such as the Kelvin–Voigt model (Lankarani 1988), or nonlinear, such as the Hunt and Crossley model (Hunt and Crossley 1975). For long impact durations this method is effective and accurate in that the instantaneous contact forces are introduced into the system's equations of motion. The second continuous methodology specifies that when contact is detected a kinematic constraint is introduced in the system's equations. Such a constraint is maintained while the reaction forces are compressive, and removed when the impacting bodies rebound from contact (Ambrósio 2000).

A second approach of a different nature is a discontinuous method that assumes that the impact occurs instantaneously, the integration of the equations of motion being halted at the time of impact. Then a momentum balance is performed to calculate the post-impact velocity, the integration being resumed afterwards with the

updated velocities, until the next impact occurs. In the discontinuous method, the dynamic analysis of the system is divided into two intervals, before and after impact. The restitution coefficient is employed to quantify the dissipation energy during the impact. The restitution coefficient only relates relative velocities after separation to relative velocities before contact and ignores what happens in between. The discontinuous method is relatively efficient but the unknown duration of the impact limits its application, since for large enough contact periods the system configuration changes significantly (Lankarani 1988). Hence the assumption of instantaneity of impact duration is no longer valid and the discontinuous analysis must not be adopted. This method, commonly referred to as piecewise analysis, has been used for solving the intermittent motion problem in mechanical systems (Khulief and Shabana 1986) and it is still the most commonly used approach in vehicle accident reconstruction (PC-Crash 2002).

In general, the contact points change during collision. When there is no penetration between the colliding bodies, there is no contact and, consequently, the contact forces are null. The occurrence of penetration is used as the basis to develop the procedure to evaluate the local deformation of the bodies in contact. Although the bodies are assumed to be rigid, the contact forces correspond to those evaluated as if the penetration is due to local elastic deformations. These forces are calculated as being equivalent to those that would appear if the bodies in contact were pressed against each other by an external static force. This means that the contact forces are treated as elastic forces expressed as functions of the coordinates and velocities of the colliding bodies. The methodology used here allows for the accurate calculation of the location of contact points. The direction of the normal contact force is determined from the normal vector to the plane of colliding surfaces at the points of contact.

In short, in dynamic analysis, the deformation is known at every time step from the configuration of the system and the forces are evaluated based on the state variables. With the variation of the contact force during the contact period, the response of the dynamic system is obtained by simply including updated forces into the equations of motion. Since the equations of motion are integrated over the period of contact, this approach results in a rather accurate response. Furthermore this methodology is not limited by the changes in the system configuration during the contact periods.

3.2 Normal Force Models for Spherical Contact Surfaces

The simplest contact force relationship, known as Kelvin–Voigt viscous-elastic model, is modeled by a parallel spring–damper element (Zukas et al. 1982). The spring represents the elasticity of the contacting bodies while the damper describes the loss of kinetic energy during the impact. In most studies, the stiffness and damping coefficients have been assumed to be known parameters, and the analysis has been confined to unconstrained bodies. The spring stiffness in the element can be

calculated using a simple mechanical formula or obtained by means of the finite element method (FEM). Recently Zhu et al. (1999) proposed a theoretical formula for calculating damping in the impact of two bodies in a multibody system. This model assumes that both the spring and the damper are linear. When the contact bodies are separating from each other the energy loss is included in the contact model by multiplying the rebound force with a coefficient of restitution. The restitution coefficient accounts for the energy dissipated during the impact in the form of a hysteresis in the relation between force and deformation.

The normal Kelvin–Voigt contact force, F_N, is calculated for a given penetration depth, δ, as

$$F_N = \begin{cases} K\delta & \text{if } v_N > 0 \quad \text{(loading phase)} \\ K\delta\, c_e & \text{if } v_N < 0 \ \text{(unloading phase)} \end{cases} \tag{3.1}$$

where K is the stiffness, δ is the relative penetration depth, c_e is the restitution coefficient and v_N is the relative normal velocity of the colliding bodies.

The primary drawback associated with this model is the quantification of the spring constant, which depends on the geometry and material properties of the contacting bodies. On the other hand, the assumption of a linear relation between the penetration depth and the contact forces is at best a rough approximation because the contact force depends on the shape, surface conditions and material properties of the contacting surfaces, all of which suggest a more complex relation.

For the linear Kelvin–Voigt model, Fig. 3.1a–c shows the penetration depth δ, the normal contact force F_N and the hysteresis of two internally colliding spheres. The restitution coefficient and the spring stiffness used to build Fig. 3.1 are 0.9 and $1.5 \times 10^8\,\mathrm{N}/m$, respectively.

The best-known contact force law between two spheres of isotropic materials is due to the result of the work by Hertz, which is based on the theory of elasticity (Timoshenko and Goodier 1970). The Hertz (1896) contact theory is restricted to frictionless surfaces and perfectly elastic solids being exemplified by the case shown in Fig. 3.2.

The Hertz law relates the contact force with a nonlinear power function of penetration depth and is written as

$$F_N = K\delta^n \tag{3.2}$$

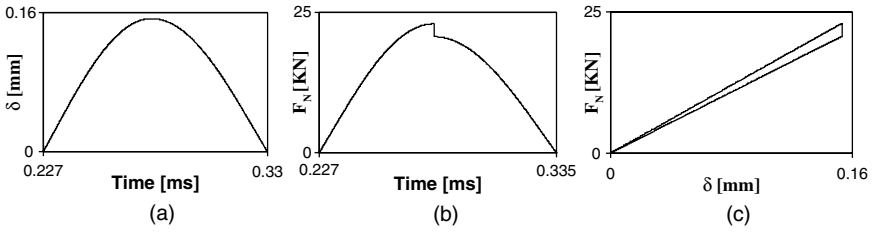

Fig. 3.1 Internally colliding spheres modeled by linear Kelvin–Voigt viscous-elastic contact model: (**a**) penetration depth, d; (**b**) normal contact force, F_N; (**c**) force–penetration relation

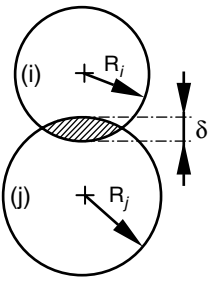

Fig. 3.2 Relative penetration depth during the impact between two spheres

where K is the generalized stiffness constant and δ is the relative normal indentation between the spheres. The exponent n is set to 1.5 for the cases where there is a parabolic distribution of contact stresses, as in the original work by Hertz. Although for metallic materials $n = 1.5$ for other materials such as glass or polymer it can be either higher or lower, leading to a convenient contact force expression that is based on experimental work but that should not be confused with Hertz theory. The generalized parameter K is dependent on the material properties and the shape of the contact surfaces. For two spheres in contact the generalized stiffness coefficient is a function of the radii of the spheres i and j and the material properties as (Goldsmith 1960)

$$K = \frac{4}{3(\sigma_i + \sigma_j)} \left[\frac{R_i R_j}{R_i + R_j} \right]^{\frac{1}{2}} \tag{3.3}$$

where the material parameters σ_i and σ_j are given by

$$\sigma_k = \frac{1 - \nu_k^2}{E_k}, \quad (k = i, j) \tag{3.4}$$

and the quantities ν_k and E_k are the Poisson's ratio and the Young's modulus associated with each sphere, respectively. For contact between a sphere body i and a plane surface body j the generalized stiffness coefficient depends on the radius of the sphere and the material properties of the contacting surfaces, being expressed by (Goldsmith 1960)

$$K = \frac{4}{3(\sigma_i + \sigma_j)} \sqrt{R_i}. \tag{3.5}$$

For two internally colliding spheres modeled by Hertz contact law, Fig. 3.3a–c shows the penetration depth, δ, the normal contact force, F_N, and the relation force–penetration. The generalized stiffness is equal to $6.6 \times 10^{10}\,\mathrm{N}/m^{1.5}$ for the calculations used to generate the graphs.

It is apparent that the Hertz contact law given by (3.2) cannot be used during both phases of contact (loading and unloading phases), since this model does not take into account the energy dissipation during the process of impact. This is a pure elastic contact model, that is, the contact energy stored during the loading phase

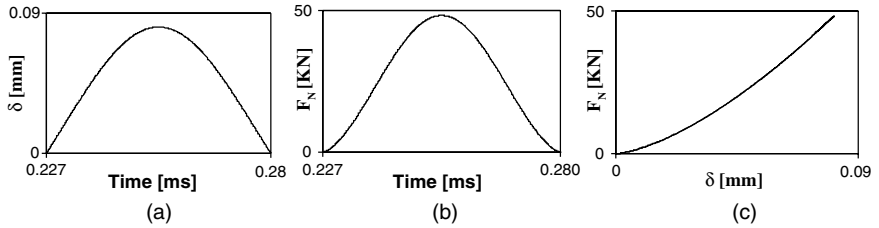

Fig. 3.3 Internally colliding spheres modeled by Hertz contact law: (**a**) penetration depth, δ; (**b**) normal contact force, F_N; (**c**) force–penetration ratio

is exactly the same that is restored during the unloading phase. The advantages of the Hertz law relative to Kelvin–Voigt contact model reside on its physical meaning represented by both its nonlinearity and by the relation between the generalized stiffness and geometry and material of the contacting surfaces. Although the Hertz law is based on the elasticity theory, some studies have been performed to extend its application to include energy dissipation. In fact, the process of energy transfer is a complicated part of modeling impacts. If an elastic body is subjected to a cyclic load, the energy loss due to internal damping causes a hysteresis loop in the force–displacement diagram, which corresponds to energy dissipation.

Hunt and Crossley (1975) showed that the linear spring–damper model does not represent the physical nature of energy transferred during the impact process. Instead they represent the contact force by the Hertz force law with a nonlinear viscous-elastic element. This approach is valid for direct central and frictionless impacts. Based on Hunt and Crossley's work, Lankarani and Nikravesh (1990) developed a contact force model with hysteresis damping for impact in multibody systems. The model uses the general trend of the Hertz law, the hysteresis damping function being incorporated to represent the energy dissipated during the impact. Lankarani and Nikravesh (1990) suggested separating the normal contact force into elastic and dissipative components as

$$F_N = K\delta^n + D\dot{\delta} \tag{3.6}$$

where the first term of the right-hand side is referred to as the elastic force and the second term accounts for the energy dissipated during the impact. In (3.6), the quantity D is a hysteresis coefficient and $\dot{\delta}$ is the relative normal impact velocity.

The hysteresis coefficient is written as a function of penetration as

$$D = \chi\ \delta^n \tag{3.7}$$

in which the hysteresis factor χ is given by

$$\chi = \frac{3K(1-c_e^2)}{4\dot{\delta}^{(-)}} \tag{3.8}$$

$\dot{\delta}^{(-)}$ being the initial impact velocity. By substituting (3.8) into (3.7) and the results into (3.6), the normal contact force is finally expressed as

$$F_N = K\delta^n \left[1 + \frac{3(1-c_e^2)}{4} \frac{\dot{\delta}}{\dot{\delta}^{(-)}} \right] \tag{3.9}$$

where the generalized parameter K is evaluated by (3.3) and (3.4) for sphere-to-sphere contact or by similar expressions for the contact of other types of geometry; c_e is the restitution coefficient, $\dot{\delta}$ is the relative normal penetration velocity and $\dot{\delta}^{(-)}$ is the initial normal impact velocity where contact is detected. The use of the damping scheme included in this model implies the outcome illustrated in Fig. 3.4a–c in which the penetration depth, δ, normal contact force, F_N, and hysteresis of an impact between two internally colliding spheres are presented. The generalized stiffness used to evaluate the relations in Fig. 3.4 is $6.6 \times 10^{10}\,\mathrm{N}/m^{1.5}$.

Equation (3.9) is valid only for impact velocities lower than the propagation velocity of elastic waves across the bodies, that is, $\dot{\delta}^{(-)} \leq 10^{-5}\sqrt{E/\rho}$, where E is the Young's modulus and ρ is the material mass density (Love 1944). The quantity $\sqrt{E/\rho}$, velocity of wave propagation, is the larger of two propagation velocities of the elastic deformation waves in the colliding bodies.

Shivaswamy (1997) studied theoretically and experimentally the impact between bodies and demonstrated that at low impact velocities, the hysteresis damping is the prime factor for energy dissipation. Impact at higher velocities, exceeding the propagation velocity of the elastic deformation waves, is likely to dissipate energy in a form not predicted by the current model. In a later work, Lankarani and Nikravesh (1994) proposed a new approach for contact force analysis, in which the permanent indentation is also included. At fairly moderate or high velocities of collision, especially in the case of metallic solids, permanent indentations are left behind on the colliding surfaces. Hence local plasticity of the surfaces in contact becomes the dominant source of energy dissipation during impact. Permanent or plastic deformations are beyond the scope of the present work.

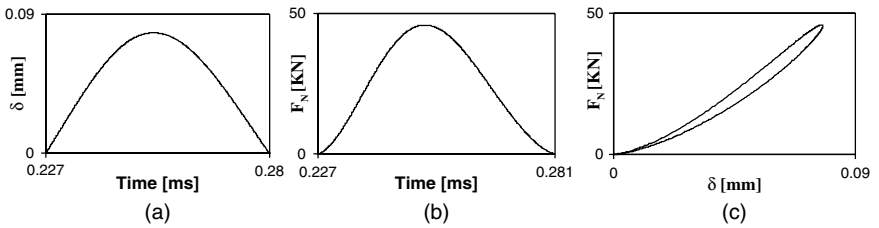

Fig. 3.4 Internally colliding spheres modeled by the Lankarani and Nikravesh contact force model: (**a**) penetration depth, δ; (**b**) normal contact force, F_N; (**c**) force–penetration relation

3.3 Normal Contact Force Models for Cylindrical Surfaces

The contact models given by (3.2) and (3.9) are valid only for colliding bodies that exhibit in the contacting surface a parabolic contact stress distribution, such as in the case of ellipsoidal contact areas. For a cylindrical contact area such as the one pictured in Fig. 3.5 between two parallel cylinders, a literature search reveals few approximate force–displacement relationships.

It is worth noting that line contact assumes a precise parallel alignment of the colliding cylinders. Furthermore a uniform force distribution over the length of the cylinders is also assumed and boundary effects are neglected. For the case of cylindrical contact forces, some authors suggest the use of the more general and straightforward force–displacement relation given by (3.9) but with an exponent, n, in the range of 1–1.5 (Hunt and Crossley 1975, Ravn 1998). Dietl (1997) used the classical solution of contact, presented by Hertz, but with the exponent n equal to 1.08 to model the contact between the journal and the bearing elements.

Based on Hertz theory, Dubowsky and Freudenstein (1971) proposed an expression for the indentation, as function of the contact force, of an internal pin inside a cylinder as

$$\delta = F_N \left(\frac{\sigma_i + \sigma_j}{L} \right) \left[\ln \left(\frac{L^b (R_i - R_j)}{F_N R_i R_j (\sigma_i + \sigma_j)} \right) + 1 \right] \tag{3.10}$$

where $R_{i,j}$ and $\sigma_{i,j}$ are the parameters shown in (3.4), L is the length of the cylinder and the exponent b has a value 3. Since (3.10) is a nonlinear implicit function for F_N, with a known penetration depth, F_N can be evaluated. This is a nonlinear problem and requires an iterative solution scheme, such as the Newton–Raphson method, to solve for the normal contact force, F_N.

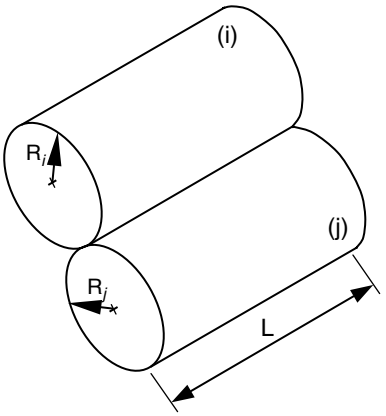

Fig. 3.5 Contact between two external cylinders

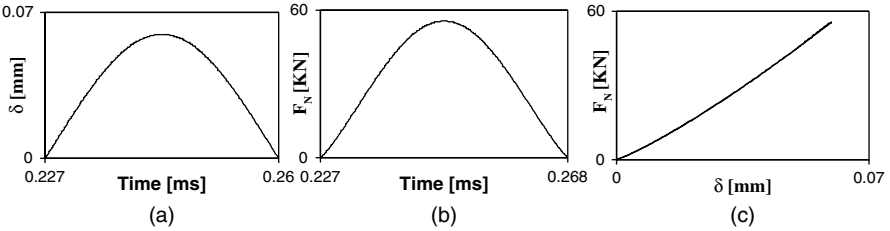

Fig. 3.6 Internally colliding cylinders using Dubowsky and Freudenstein contact force model: (**a**) penetration depth, δ; (**b**) normal contact force, F_N; (**c**) force–penetration ratio

Based on the Dubowsky and Freudenstein contact force model the solutions corresponding to the time variation of the indentation δ, the normal contact force F_N and the force–penetration depth ratio are shown in Fig. 3.6a–c. In the plots presented for the case of a pin contact inside a cylinder, the pin and cylinder radii are 9.5 and 10 mm, respectively. The length of the cylinder is equal to 15 mm, and both the pin and the cylinder are made of steel.

Goldsmith (1960) proposed an expression similar to (3.10) but with the value of exponent b equal to 1. However, this value for b leads to a problem of consistency of the units in the expression. Figure 3.7a–c shows the penetration depth δ, the normal contact force F_N and the force–penetration ratio of two internally colliding cylinders modeled with the Goldsmith contact force. This model shows that the force–penetration ratio is almost linear.

The ESDU 78035 Tribology Series (1978) also proposes some expressions for contact mechanics analysis suitable for engineering applications. For a circular contact area the ESDU 78035 model is the same as the pure Hertz law given by (3.2). For rectangular contact, e.g., a pin inside a cylinder, the expression is given by

$$\delta = F_N \left(\frac{\sigma_i + \sigma_j}{L} \right) \left[\ln \left(\frac{4L(R_i - R_j)}{F_N(\sigma_i + \sigma_j)} \right) + 1 \right] \tag{3.11}$$

where all quantities are the same as used for the calculations in Fig. 3.6.

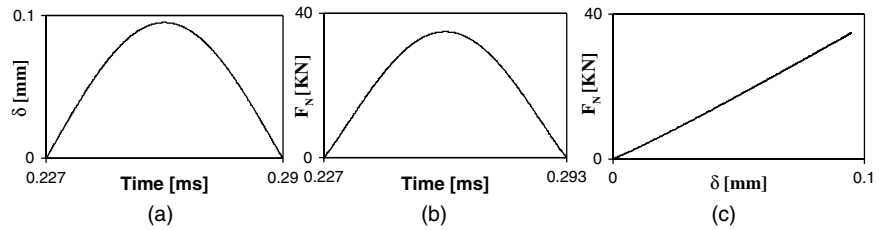

Fig. 3.7 Internally colliding cylinders using Goldsmith contact force model: (**a**) penetration depth, δ; (**b**) normal contact force, F_N; (**c**) force–penetration ratio

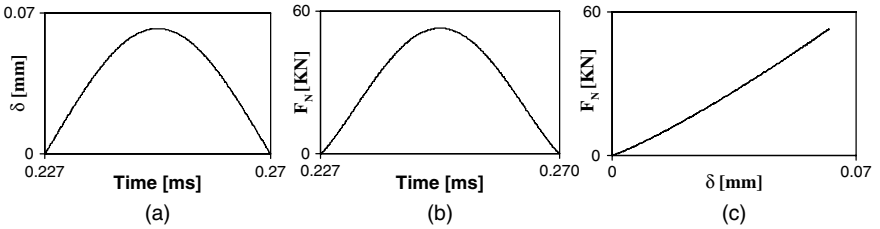

Fig. 3.8 Internally colliding cylinders using the ESDU 78035 contact force model: (**a**) penetration depth, δ; (**b**) normal contact force, F_N; (**c**) force–penetration ratio

Figure 3.8a–c shows the penetration depth δ, the normal contact force F_N and the force–penetration ratio of two internally colliding cylinders modeled by ESDU 78035, given by (3.11).

The contact force due to the spherical and cylindrical contact force models is displayed in Fig. 3.9, where it can be observed that the spherical and cylindrical force–displacement relations are reasonably close. Thus the straightforward force–penetration relation proposed by Lankarani and Nikravesh given in (3.9) is largely used for mechanical contacts not only because of its simplicity and easiness in implementation in a computational program, but also because this is the only model that accounts for the energy dissipation during the impact process (Ryan 1990, Smith and Haug 1990, Bottasso et al. 1999, Pedersen 2001, Pedersen et al. 2002, Silva and Ambrósio 2004, Flores et al. 2006).

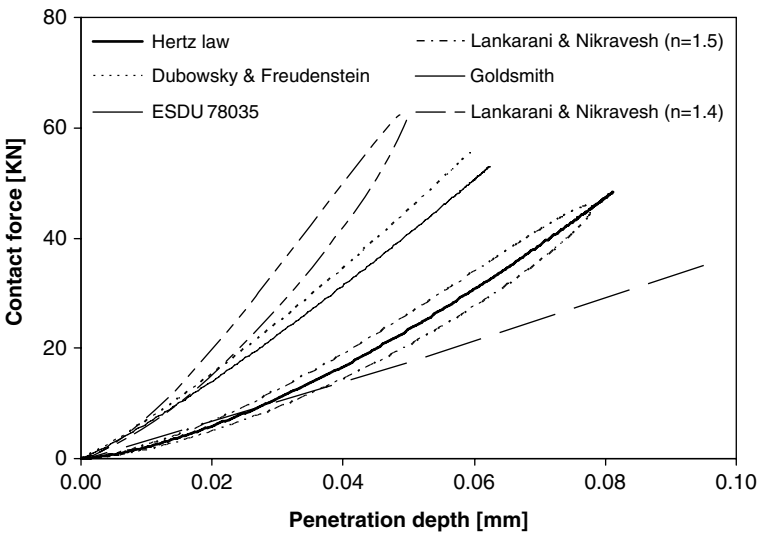

Fig. 3.9 Force–deformation curves for spherical and cylindrical contact surfaces

3.4 Tangential Friction Force Models

When contacting bodies slide or tend to slide relative to each other tangential forces are generated. These forces are usually referred to as friction forces. Three basic principles have been experimentally established, namely: (1) the friction force acts in a direction opposite to that of the relative motion between the two contacting bodies; (2) the friction force is proportional to the normal load on the contact; (3) the friction force is independent of a nominal area of contact. These three statements constitute what is known as the laws of sliding friction under dry conditions (Stolarski 1990). Based on the experimental observations by Angmonds in the field of optics, Coulomb developed what is today the most commonly used friction law, but there is still no simple model which can be universally used by designers to calculate the friction force for a given pair of bodies in contact. In face of the shortcomings of the Coulomb friction law, in recent years there has been much interest on the subject of friction, and many research papers have focused on the subject (Keller 1986, Wang and Manson 1992, Han and Gilmore 1993).

The presence of friction in the contact surfaces makes the contact problem more complicated as the friction may lead to different contact modes, such as sticking or sliding. For instance, when the relative tangential velocity of two impacting bodies approaches zero, stiction occurs. Indeed, as pointed out by Ahmed et al. (1999), the friction model must be capable of detecting sliding, sticking and reverse sliding to avoid energy gains during impact. This work was developed for the treatment of impact problems in jointed open-loop multibody systems. Lankarani (2000) extends Ahmed's formulation to the analysis of impact problems with friction in any general multibody system including both open- and closed-loop systems.

The Coulomb's friction law of sliding friction can represent the most fundamental and simplest model of friction between dry contacting surfaces. When sliding takes place, the Coulomb law states that the tangential friction force F_T is proportional to the magnitude of the normal contact force, F_N, at the contact point by introducing a coefficient of friction c_f (Greenwood 1965). The Coulomb's friction law is independent of relative tangential velocity. In practice, this is not true because friction forces can depend on many parameters such as material properties, temperature, surface cleanliness and velocity of sliding, which cannot all be accumulated for by a constant friction coefficient. Therefore a continuous friction force–velocity relationship is desirable. Furthermore the application of the original Coulomb's friction law in a general-purpose computational program may lead to numerical difficulties because it is a highly nonlinear phenomenon that may involve switching between sliding and stiction conditions. Also from this point of view, more realistic friction force models are required.

In the last decades, a number of papers addressed the issue of the tangential friction forces (Bagci 1975, Threlfall 1978, Rooney and Deravi 1982, Haug et al. 1986, Wu et al. 1986a, b). Most of them use the Coulomb friction model with some modification in order to avoid the discontinuity at zero relative tangential velocity and to obtain a continuous friction force. Dubowsky (1974) assumed the friction force to be equal to a constant value opposing the direction of velocity, given by

$$\mathbf{f}_T = -c_c \frac{\mathbf{v}_T}{v_T} \tag{3.12}$$

where c_c is a coefficient independent of normal contact force and $\mathbf{v}_T$ is the relative tangential velocity. This model does not take the effect of zero velocity into account, that is, it has the disadvantage of an infinite gradient at null relative tangential velocity. This causes computational difficulties in the integration process since the force instantaneously changes from $-\mathbf{f}_T$ to $+\mathbf{f}_T$. This model is qualitatively illustrated in Fig. 3.10a, which shows the Coulomb's friction force versus relative tangential velocity.

A friction model with better numerical features is found in Rooney and Deravi (1982), where the friction force is calculated from two sets of equations. When the relative tangential velocity is not close to zero the Coulomb's friction law is given by

$$\mathbf{f}_T = -c_f \mathbf{f}_N \frac{\mathbf{v}_T}{v_T} \tag{3.13}$$

and when the relative tangential velocity of the contacting bodies is close to zero the friction force is a value within a range given by

$$-c_f F_N < F_T < c_f F_N \tag{3.14}$$

where c_f is the coefficient of friction, $\mathbf{v}_T$ is the relative tangential velocity and F_N is the normal contact force, which is always positive. This model is illustrated in Fig. 3.10b.

Wilson and Fawcett (1974) used the model expressed by (3.13) in the dynamic study of a slider–crank mechanism with a clearance joint between the slider and the guide. More recently, Ravn (1998) also used (3.13) to include the friction effect in revolute joints with clearance. Threlfall (1978) proposed another friction force model, in which the transition between $-\mathbf{f}_T$ and $+\mathbf{f}_T$ is made using a curve as follows:

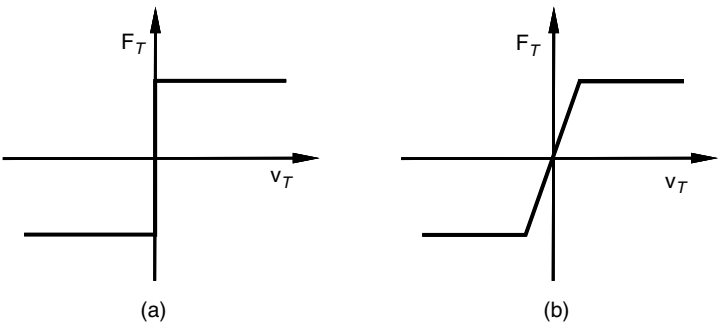

Fig. 3.10 (**a**) Standard Coulomb's friction law, (**b**) Rooney and Deravi friction force

$$\mathbf{f}_T = c_f \mathbf{f}_N \frac{\mathbf{v}_T}{v_T} \left(1 - \mathrm{e}^{-\left(\frac{3v_T}{v_r}\right)}\right), \quad \text{if } |\mathbf{v}_T| < v_r \tag{3.15}$$

where c_f is the friction coefficient, $\mathbf{f}_N$ is the normal contact force, v_r is a small characteristic velocity as compared to the maximum relative tangential velocity encountered during the simulation. The value of v_r is a specified parameter that for small values results in slowing down the integration method as it gets closer to the idealized model of Fig. 3.10a. In practice, the regulation factor $1 - \exp(-3v_T/v_r)$ smoothes out the friction force discontinuity. The shape of this curve is illustrated in Fig. 3.11a.

Ambrósio (2002) presented another modification for Coulomb's friction law, in which the dynamic friction force is expressed as

$$\mathbf{f}_T = -c_f c_d \mathbf{f}_N \frac{\mathbf{v}_T}{v_T} \tag{3.16}$$

where c_f is the friction coefficient, F_N is the normal contact force, $\mathbf{v}_T$ is the relative tangential velocity and c_d is a dynamic correction coefficient, which is expressed as

$$c_d = \begin{cases} 0 & \text{if} \quad v_T \le v_0 \\ \frac{v_T - v_0}{v_1 - v_0} & \text{if} \quad v_0 \le v_T \le v_1 \\ 1 & \text{if} \quad v_T \ge v_1 \end{cases} \tag{3.17}$$

in which v_0 and v_1 are given tolerances for the velocity. This dynamic correction factor prevents the friction force from changing direction for almost null values of the tangential velocity, which is perceived by the integration algorithm as a dynamic response with high-frequency contents, thereby forcing a reduction in the time-step size. The great merit of this modified Coulomb's law is that it allows the numerical stabilization of the integration algorithm. This friction force model, illustrated in Fig. 3.11b, does not account for other phenomena like the adherence between the sliding contact surfaces, which can be added as a complementary model.

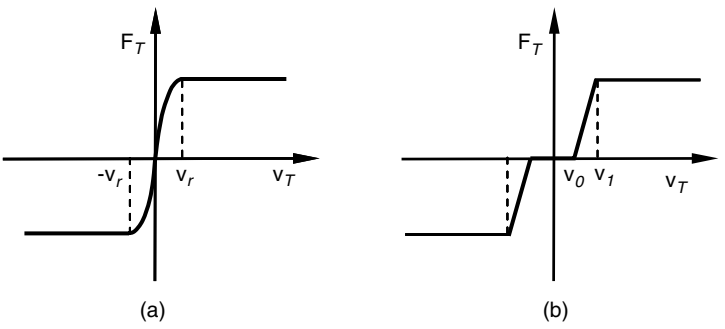

Fig. 3.11 (**a**) Threlfall friction force, (**b**) friction force of (3.16)

3.5 Numerical Aspects in Contact Analysis

In a dynamic simulation, it is very important to find the precise instant of transition between the different states, that is, contact and noncontact situations. This requires close interaction with the numerical procedure to continuously detect and analyze all situations. If not, errors may build up and the final results will become inaccurate.

When a system consists of fast- and slow-moving components, that is, the *eigenvalues* are widely spread, the system is designated as being stiff (Nikravesh 1988). Stiffness in the system equations of motion arises when the gross motion of the overall mechanism is combined with the nonlinear contact forces that lead to rapid changes in velocity and accelerations. In addition, when the equations of motion are described by a coupled set of differential and algebraic equations, the error of the response system is particularly sensitive to constraints violation. Constraints violation inevitably leads to artificial and undesired changes in the energy of the system. Yet, by applying a stabilization technique, the constraint violation can be kept under control (Baumgarte 1972). During the numerical integration procedure, both the order and the step size are adjusted to keep the error tolerance under control. In particular, the variable step size of the integration scheme is a desirable feature when integrating systems that exhibit different time scales, such as in multibody systems with impacting bodies (Shampine and Gordon 1975). Thus large steps are taken when the system's motion does not include contact forces, and when impact occurs the step size is decreased substantially to capture the high-frequency response of the system.

One of the most critical aspects in the dynamic simulation of the multibody systems with collisions is the detection of the precise instant of contact. In addition, the numerical model used to characterize the contact between the bodies requires the knowledge of the pre-impact conditions, that is, the impact velocity and the direction of the plane of collision. The contact duration and the penetration cannot be predicted from the pre-impact conditions due to the influence of the kinematic constraints imposed in the bodies on the overall system motion. Thus, before the first impact, the bodies can freely move relative to each other and, in this phase, the step size of the integration algorithm may become relatively large. The global motion of the system may be characterized by relatively large translational and rotational displacements during a single time step. Therefore, if the numerical integration is not handled properly, the first impact between the colliding bodies is often made with a high penetration depth and, hence, the calculated contact forces become artificially large.

The importance of the initial penetration control, in the framework of the integration of the equations of motion, is better discussed using a simple example. Take the case of the falling ball illustrated in Fig. 3.12, with a mass $m = 1.0\,\mathrm{kg}$, a moment of inertia equal to $0.1\,\mathrm{kg\,m^2}$, a radius $R = 0.1\,\mathrm{m}$, animated by an initial horizontal velocity $v = 1.0\,\mathrm{m}/s$ and acted upon by gravity forces only. The motion of the ball is such that during its falling trajectory it strikes the ground. The penetration of the ball in the ground, in the integration time step, for which contact is first detected, is

$$\delta^{(-)} = (h - R) - y_b \tag{3.18}$$

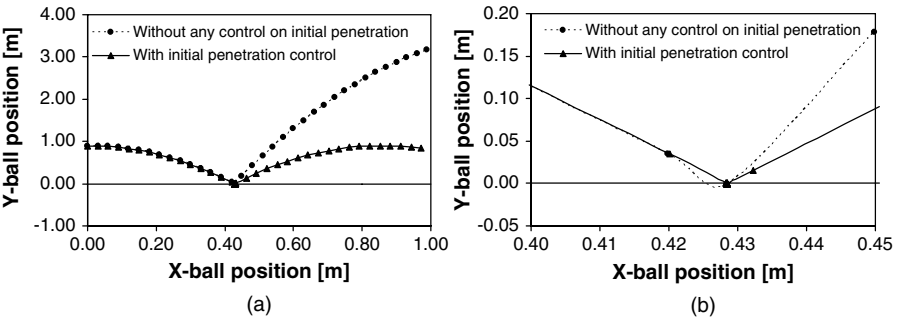

Fig. 3.12 (**a**) Trajectory of a falling ball obtained with integration algorithms with and without initial penetration control; (**b**) detailed view in the vicinity of contact

where y_b is the y coordinate of the ball's center of mass. The superscript $(-)$ on δ means that it is the penetration when contact is first detected. Note that $\delta^{(-)}$ must be a positive value for contact. Therefore, by monitoring the sign of the penetration at every time step $t + \Delta t$ the start can be identified from

$$\delta^{(-)}(\mathbf{q}, t)\ \delta^{(-)}(\mathbf{q}, t + \Delta t) \leq 0 \tag{3.19}$$

When (3.19) is verified the start of contact is defined as occurring at $t + \Delta t$. The integration of the equations of motion of the system could be proceeded with no numerical problem if the penetration first detected is close to zero, or at least below a pre-defined threshold, i.e., if $\delta^{(-)}(\mathbf{q},\ t + \Delta t) \leq \delta_{max}$. Because this is not always the case, strategies to limit the time step in the vicinity of contact must be implemented when solving contact problems.

Define as δ^- the distance between the two surfaces in the time step t^- that precedes the time step t^+, at which penetration δ^+ is first detected. In between these time steps, say at t^c, the penetration $\delta^c = 0$ exists. Assuming constant velocity for the multibody system in the vicinity of contact, the time at which contact starts can be calculated by

$$t^c = t^- + \frac{\delta^-}{\delta^+ - \delta^-}\Delta t \tag{3.20}$$

Consequently the ideal situation, during the integration of the multibody equations of motion, would be a time step in the vicinity of contact of

$$\Delta t^{ideal} = t^c - t^- + \varepsilon \tag{3.21}$$

where ε is a very small number to effectively ensure that $\delta_{max} > \delta^c > 0$. Several procedures are suggested to ensure that $\delta^+ < \delta_{max}$, which can be implemented in any code, depending on the access that exists to the numerical integrator.

Procedure 1: Assume that in the vicinity of contact the motion of the multibody system is such that each body moves approximately with constant velocity. Note that the assumption only needs to be valid within a simple time step. Then the time for contact is calculated using (3.20) and the ideal time step using (3.21). Now the positions and velocities of the multibody system, at the time of contact t^c, are estimated as

$$\mathbf{q}^c = \mathbf{q}^- + (\mathbf{q}^+ - \mathbf{q})\frac{\Delta t^{ideal}}{\Delta t} \tag{3.22}$$

$$\dot{\mathbf{q}}^c = \dot{\mathbf{q}}^- + (\dot{\mathbf{q}}^+ - \dot{\mathbf{q}})\frac{\Delta t^{ideal}}{\Delta t} \tag{3.23}$$

where the superscripts $-$, $+$ and c mean that the quantity in which they are applied is evaluated at the instant before contact, after contact and at the time of contact, respectively. The integration algorithm is now restarted at time t^c with the initial positions and velocities given by (3.22) and (3.23).

The procedure proposed, being approximate, does present slight violations of the position and velocity constraint equations. Because a constraint stabilization method or a constraint elimination method is being used, according to the discussion in Chap. 2, it is expected that such violations remain under control. Notice also that when a variable time-step integrator is used during its start the time steps are small. Therefore in the vicinity of contact, small time steps are used by the integrator and even if the conditions calculated by (3.22) and (3.23) are just before contact the integration process continues with the guarantee that the initial penetration never exceeds the prescribed threshold.

Procedure 2: The numerical algorithms used for integration of first-order differential equations with variable time steps, such as the ones generally used in multibody dynamics (Shampine and Gordon 1975, Gear 1981), include an error control that supports the acceptance of rejection of a particular time step. Such decision is based on numerical issues, related to the dynamic response of the system, rather than on any physical reason. The idea behind this procedure, to handle the control on the initial penetration, is to devise a complementary control for the selection of the integration time step based on physical reasoning only. Say that at a given time, during the integration of the equations of motion of multibody system, the internal numerical control of the integration algorithm tests a time step Δt_{trial} and decides to accept it. Before it is definitely accepted, the following physical condition must be met by all new contacts detected in the system:

$$\delta^{(-)}\left(\mathbf{q},\ t + \Delta t_{trial}\right) < \delta_{max} \tag{3.24}$$

If the condition described by (3.24) is met by all new contacts, the integration continues without any further interference. If (3.24) is not met the integration algorithm takes it as an indication to reject the time step and attempts a smaller time step. Generally such action corresponds to halving the attempted time step, but particular

integration error controls may take different actions. When a smaller new time step is attempted the condition defined by (3.24) is checked again and a decision is made. Eventually a suitable time step that ensures the fulfillment of (3.24) for all new contacts is identified. The integrators available in math libraries include features to inform to the user if the error control intends to accept or reject a time step before doing it. When such features are available the procedure just described is easily implemented.

3.6 Summary

In this chapter, different continuous contact-impact force models for both spherical and cylindrical shape surface collisions in multibody mechanical systems were reviewed. In addition, various types of friction force models based on the Coulomb's law were also listed and discussed. Because modeling contact forces plays a crucial role in the analysis of multibody mechanical systems that experience impacts, the contact force model must be computed using suitable constitutive laws that take into account material properties of the colliding bodies, geometric characteristics of the impacting surfaces and, eventually, the impact velocity. Additionally the numerical method for the calculation of the contact force should be stable enough to allow for the integration of the mechanical equations of motion with acceptable efficiency. These characteristics are ensured by using a continuous contact force model in which the force and penetration vary in a continuous manner and for which some energy dissipation is included. This approach has the extra benefit of leading to a behavior of the variable time-step integrator that is more stable.

Some important conclusions can be drawn from the study presented in this chapter. Among the spherical-shaped contact areas, the linear Kelvin–Voigt contact model does not represent the overall nonlinear nature of impact phenomenon. The Hertz relation does not account for the energy dissipation during the impact process. Therefore the Hertz relation, along with the modification to represent the energy dissipation, in the form of internal damping, can be adopted for modeling contact forces in a multibody system. This model is straightforward and easy to implement in a computational program.

The cylindrical models are nonlinear and implicit functions, and therefore, they require a numerical iterative procedure to be performed. Furthermore these models have been proposed as purely elastic, not being able to explain the energy dissipation during the impact process. From the comparison between the spherical and cylindrical contact force models, it can be concluded that the spherical and cylindrical force–displacement relations are reasonably close. Therefore, the straightforward force–penetration relation proposed by Lankarani and Nikravesh (1990), with a modification of the pseudo-stiffness parameter in the case of cylindrical contact, is largely used for mechanical contacts owing to its simplicity and easiness of implementation in a computational program. Aiding to those advantages, this is the only model that accounts for energy dissipation during the impact process.

In dynamic analysis of multibody systems, the deformation/indentation is known at every time step from the configuration of the system, the forces evaluated being based on the state variables. With the variation of the contact force during the contact period, the dynamic system response is obtained by simply including updated forces into the equations of motion. Since the equations of motion are integrated over the period of contact, this approach results in a rather accurate response. This procedure was further improved by including in the time integration scheme a procedure that controls the time step in order to prevent large penetrations to develop in the initial contact. Furthermore this methodology accounts for the changes in the system's configuration during the contact period. This approach is employed in the forthcoming chapters to describe the impact between the elements that compose the joint clearances.

References

Ahmed S, Lankarani HM, Pereira MFOS (1999) Frictional impact analysis in open loop multibody mechanical system. Journal of Mechanical Design 121:119–127.

Ambrósio J (2000) Rigid and flexible multibody dynamics tools for the simulation of systems subjected to contact and impact conditions. European Journal of Solids A/Solids 19:S23–S44.

Ambrósio J (2002) Impact of rigid and flexible multibody systems: deformation description and contact models. Virtual nonlinear multibody systems, NATO Advanced Study Institute, Prague, Czech Republic, June 23–July 3, edited by W Schiehlen and M Valásek, Vol. II, pp. 15–33.

Bagci C (1975) Dynamic motion analysis of plane mechanisms with Coulomb and viscous damping via the joint force analysis. Journal of Engineering for Industry, Series B 97(2):551–560.

Baumgarte J (1972) Stabilization of constraints and integrals of motion in dynamical systems. Computer Methods in Applied Mechanics and Engineering 1:1–16.

Bottasso CL, Citelli P, Taldo A, Franchi CG (1999) Unilateral contact modeling with adams. International ADAMS user's conference, Berlin, Germany, November 17–18, 11pp.

Brach RM (1991) Mechanical impact dynamics, rigid body collisions. Wiley, New York.

Dietl P (1997) Damping and stiffness characteristics of rolling element bearings—theory and experiment. Ph.D. Dissertation, Technical University of Vienna, Austria.

Dubowsky S (1974) On predicting the dynamic effects of clearances in planar mechanisms. Journal of Engineering for Industry, Series B 96(1):317–323.

Dubowsky S, Freudenstein F (1971) Dynamic analysis of mechanical systems with clearances, part 1: formulation of dynamic model. Journal of Engineering for Industry, Series B 93(1):305–309.

ESDU 78035 Tribology Series (1978) Contact phenomena. I: stresses, deflections and contact dimensions for normally loaded unlubricated elastic components. Engineering Sciences Data Unit, London, England.

Flores P, Ambrósio J, Claro JCP, Lankarani HM (2006) Influence of the contact-impact force model on the dynamic response of multibody systems. Proceedings of the Institution of Mechanical Engineers, Part-K Journal of Multi-body Dynamics 220:21–34.

Gear CW (1981) Numerical solution of differential-algebraic equations. IEEE Transactions on Circuit Theory CT-18:89–95.

Goldsmith W (1960) Impact—the theory and physical behaviour of colliding solids. Edward Arnold, London, England.

Greenwood DT (1965) Principles of dynamics. Prentice Hall, Englewood Cliffs, NJ.

Han I, Gilmore BJ (1993) Multi body impact motion with friction analysis, simulation, and validation. Journal of Mechanical Design 115:412–422.

Haug EJ, Wu SC, Yang SM (1986) Dynamics of mechanical systems with Coulomb friction, stiction, impact and constraint addition deletion—I theory. Mechanism and Machine Theory 21(5):401–406.

Hertz H (1896) On the contact of solids—on the contact of rigid elastic solids and on hardness. Miscellaneous papers (Translated by DE Jones and GA Schott), pp. 146–183. Macmillan, London, England.

Hunt KH, Crossley FR (1975) Coefficient of restitution interpreted as damping in vibroimpact. Journal of Applied Mechanics 7:440–445.

Keller JB (1986) Impact with friction. Journal of Applied Mechanics 53:1–4.

Khulief YA, Shabana AA (1986) Dynamic analysis of constrained system of rigid and flexible bodies with intermittent motion. Journal of Mechanisms, Transmissions, and Automation in Design 108:38–45.

Lankarani HM (1988) Canonical equations of motion and estimation of parameters in the analysis of impact problems. Ph.D. Dissertation, University of Arizona, Tucson, AZ.

Lankarani HM (2000) A poisson based formulation for frictional impact analysis of multibody mechanical systems with open or closed kinematic chains. Journal of Mechanical Design 115:489–497.

Lankarani HM, Nikravesh PE (1990) A contact force model with hysteresis damping for impact analysis of multibody systems. Journal of Mechanical Design 112:369–376.

Lankarani HM, Nikravesh PE (1994) Continuous contact force models for impact analysis in multibody systems. Nonlinear Dynamics 5:193–207.

Love AEH (1944) A treatise on the mathematical theory of elasticity, Fourth edition. Dover Publications, New York.

Maw N, Barber JR, Fawcett JN (1976) The oblique impact of elastic spheres. Wear 38:101–114.

Nikravesh PE (1988) Computer-aided analysis of mechanical systems. Prentice Hall, Englewood Cliffs, NJ.

PC-Crash (2002) A simulation program for vehicle accidents, Technical manual, Version 6.2.

Pedersen S, Hansen J, Ambrósio J (2002) A novel roller-chain drive model using multibody dynamics analysis tools. Virtual nonlinear multibody systems, NATO Advanced Study Institute, Prague, Czech Republic, June 23–July 3, edited by W Schiehlen and M. Valásek, Vol. II, pp. 180–185.

Pedersen SL (2001) Chain vibrations. MS Dissertation, Department of Mechanical Engineering, Solid Mechanics, Technical University of Denmark, Lyngby, Denmark.

Pfeiffer F, Glocker C (1996) Multibody dynamics with unilateral constraints. Wiley, New York.

Ravn P (1998) A continuous analysis method for planar multibody systems with joint clearance. Multibody System Dynamics 2:1–24.

Rooney GT, Deravi P (1982) Coulomb friction in mechanism sliding joints. Mechanism and Machine Theory 17:207–211.

Ryan RR (1990) ADAMS—multibody system analysis software. Multibody systems handbook. Springer, Berlin Heidelberg New York.

Shampine L, Gordon M (1975) Computer solution of ordinary differential equations: the initial value problem. Freeman, San Francisco, CA.

Shivaswamy S (1997) Modeling contact forces and energy dissipation during impact in multibody mechanical systems. Ph.D. Dissertation, Wichita State University, Wichita, KS.

Silva MPT, Ambrósio J (2004) Human motion analysis using multibody dynamics and optimization tools. Technical report IDMEC/CPM—2004/001, Instituto Superior Técnico of the Technical University of Lisbon, Lisbon, Portugal.

Smith RC, Haug EJ (1990) DADS—dynamic analysis and design system. Multibody systems handbook. Springer, Berlin Heidelberg New York.

Stolarski TA (1990) Tribology in machine design. Butterworth-Heinemann, Oxford, England.

Threlfall DC (1978) The inclusion of Coulomb friction in mechanisms programs with particular reference to DRAM. Mechanism and Machine Theory 13:475–483.

Timoshenko SP, Goodier JN (1970) Theory of elasticity. McGraw-Hill, New York.

Wang Y, Manson M (1992) Two dimensional rigid-body collisions with friction. Journal of Applied Mechanics 59:635–642.

Wilson R, Fawcett JN (1974) Dynamics of slider–crank mechanism with clearance in the sliding bearing. Mechanism and Machine Theory 9:61–80.

Wu SC, Yang SM, Haug EJ (1986a) Dynamics of mechanical systems with Coulomb friction, stiction, impact and constraint addition deletion—II planar systems. Mechanism and Machine Theory 21(5):407–416.

Wu SC, Yang SM, Haug EJ (1986b) Dynamics of mechanical systems with Coulomb friction, stiction, impact and constraint addition deletion—III spatial systems. Mechanism and Machine Theory 21(5):417–425.

Zhu SH, Zwiebel S, Bernhardt G (1999) A theoretical formula for calculating damping in the impact of two bodies in a multibody system. Proceedings of the Institution of Mechanical Engineers 213(C3):211–216.

Zukas JA, Nicholas T, Greszczuk LB, Curran DR (1982) Impact dynamics. Wiley, New York.