An introduction to Unilateral Dynamics

Jean Jacques Moreau

An introduction to Unilateral Dynamics

Jean Jacques Moreau

Laboratoire de Mécanique et Génie Civil,
cc 048, Université Montpellier II,
F-34095 Montpellier Cedex 5, France

Abstract. The paper is devoted to mechanical systems with a finite number of degrees of freedom. After showing how inequality requirements in evolution problems can be handled through differential inclusions, one introduces dynamics by an elementary example of unilateral mechanical constraint. Then a general setting is constructed for multibody multicontact systems. The description of unilateral interaction at each possible contact point is formalized, with account of possible friction. This generates the numerical time-stepping policy called Contact Dynamics. The treatment of collisions or other frictional catastrophes in this framework leads to measure-differential inclusions, an essential tool in nonsmooth dynamics. The energy balance of nonsmooth evolutions is discussed. Two illustrations of the proposed numerical methods are presented. The former concerns the mechanisms of collapse of a bridge arch under local forcing. In the latter, the construction of a conical pile of grains is simulated, in order to investigate stresses in the bulk and the distribution of pressure on ground.

1 Purpose

1.1 Computation in multibody dynamics

The numerical dynamics of collections of bodies treated as perfectly indeformable, subject to the constraints of non-interpenetrability, with friction taken into account in the event of contact, currently is an active domain of research. Applications include the dynamics of machines, in particular robots, the dynamics of masonry works submitted to transient actions (earthquakes, gusts of wind or impacts), animated computer graphics and numerical simulation in granular mechanics. In all these domains, assuming the perfect indeformability of each part of the system leads to efficient numerical procedures which, in many circumstances, satisfy the needs. Possibly, some variables are added in order to also account for a certain deformability of these parts [2][37], without essentially changing the computational strategies. Of course, for the handling of non-interpenetrability and friction, much may be learned from the rich literature devoted to the treatment of contact between deformable media discretized through finite elements, even if in most cases only quasistatic evolutions are addressed (see e.g. [3][12][18][20][35] or, concerning a dynamical situation [75]).

The techniques used in multibody dynamics may be classified into the three following categories.

1.2 Event-driven methods

The methods so qualified (abbr. ED) are practical mainly when the concerned time-interval equals the union of not too many subintervals, a priori unknown, over which the *status* of the various contacts remain unchanged, i.e. no collision which would create new contacts occurs, no contact either gets loose, nor any critical situation needing a change in the analytical expression of the Coulomb law of dry friction is met. On each of these subintervals, the same numerical techniques as in the investigation of machines with classical bilateral, possibly frictional, constraints may be used [28]. As integration proceeds, there only is to watch the evolution of some indicators. In particular, the contact forces will be calculated. If, after a certain instant, some of the values computed for these forces are found to have directions incompatible with the unilaterality of the non-interpenetrability constraints (here we neglect adhesive, i.e. gluing effects), the programme decides that the motion has to be calculated otherwise. But one should keep in mind that the contacts which get loose after the critical instant are not necessarily those for which an unfeasible contact force has just been evaluated [22]. A popular approach to such discussions consists of reducing them to *complementarity problems*, similar to what is commonly met in constrained optimization.

More embarrassing is the calculation of the new velocity in case the critical instant is that of a *collision*. Contact forces take very large values during a very short time-interval and only poor phenomenological information is usually available about such physical circumstances. Even in the simplest case, that of the collision of two otherwise free members of the system, the traditional *coefficient of restitution* is known to depend not only on the materials these bodies are made of, but also on their shapes and relative orientations at the collision locus [72]; only the case of spherical beads appears relatively comfortable [25]. Furthermore, if some of the colliding bodies belong to clusters of already contacting ones, percussional reactions should be expected at all existing contact points. Though rigid body collisions currently are an active domain of research [9][13][16][27][73], computation has to rely on pragmatic rules whose validity has to be checked in each domain of application.

1.3 Smoothing methods

'Nonsmoothness' is the salient feature of the problems in view. In fact, after the set of the possible positions of the investigated system has been parametrized through an element q of $\mathbf{R}^n$, the geometric restriction that the non-interpenetrability constraints impose on q are expressed by a set of inequalities. Hence, instead of running in a smooth submanifold as in traditional analytical dynamics, the point q is confined in a region of $\mathbf{R}^n$ whose boundary is made of a lot of pieces of hypersurfaces (millions or billions of them in current applications to granular materials): this is *nonsmoothness in space*. Furthermore, collisions are expected to induce velocity jumps: this is

nonsmoothness in time. To end, the contact forces or 'reactions' associated with the non-interpenetrability constraints are governed by highly irregular laws. These forces vanish as soon as the corresponding contacts break while, if contact holds, the commonly stipulated mechanical conditions do not express them as functions of q. If, in addition, dry friction is taken into account (most usually in the form of Coulomb's law) it introduces some irregular relationships between contact forces and the sliding velocities. All this may be called *nonsmoothness in law.*

In such a state of the affairs, a natural move is to replace, approximately, the nonsmooth governing relationships by some *regularized* ones. First the non-interpenetrability constraints will be replaced by some stiff repulsion laws which take effect as soon as two members of the system come close to each other. This automatically handles the possible collisions, as far as one considers them as 'elastic', while the dissipativity of collisions may be accounted for by adding some damping actions or also by using different repulsion laws in the episodes of approach and of separation. Similarly, frictional contact may be somewhat regularized through the introduction of local elastic microdeformation and of viscosity-like effects. The dynamics of the approximate system is then governed by differential equations with sufficient regularity to be handled through standard numerical techniques. The drawback is that the need of precision requires the use of very stiff appproximate laws. Hence the time-stepping procedures applied have to resort to very small step-length and possibly also have to enforce numerical stability by introducing artificial damping or artificially increasing inertia. When treating dynamical applications, the effect of such an artificial alteration of the mechanical data may blur the picture. Significant simulations of loose (collisional) flows of granular materials have been obtained in that way, but when dense collections of bodies are concerned (pieces of masonry or compact granulates) the method is mainly applied to *quasi-static evolutions* in which only a succession of equilibrium states is looked for. Dynamical computation is then used only as a way of attaining each of these equilibrium states, a strategy referred to as 'Dynamic Relaxation'.

The pioneering work of P. Cundall [19] was precisely based on the regularization strategy, today implemented in the majority of commercial pieces of software intended to handle non-interpenetrability. Because such computation techniques are close to those applied in molecular simulations, they are commonly referred to as 'Molecular Dynamics' methods (abbr. MD), specially in the domain of granular mechanics [76].

1.4 Contact Dynamics

This is the technique (abbr. CD) advocated by the author, to which the present paper is meant to provide an introduction. It originated from [52], where the (unilateral) contact, possibly frictional and/or collisional, between rigid bodies received a formulation in terms of elementary convex analysis

which proves suitable for computation. Mathematically, the resulting evolution problems are governed in smooth cases by *differential inclusions*. General information on the latter concept may be found in [5][21], but the need of treating also non-smooth evolutions calls for its extension to that of a *measure differential inclusion*. The Sweeping Process introduced earlier by the author, with motivation in the quasi-static evolution of elastoplastic systems [48][50] seems to have provided the first occurence of measure differential inclusions in literature. This process and some related evolution problems are still today the objet of mathematical research; see references in [42]. In Sect.2 of the present paper a description of the process, of purely kinematical nature, has been estimated able to make a tutorial introduction to the handling of unilaterality in mechanics.

A directing idea of the Contact Dynamics approach is that the main object of computation is the *velocity function* $t \mapsto u \in \mathbf{R}^n$. Time-stepping algorithms essentially have to determine the evolution of this function, by applying the principles of dynamics and the specified force laws. The position function $t \mapsto q$ is only to be updated at each step through adequate integration. Possibly, some members of u are 'pseudo-parameters' such as the components of the spin vector of a solid, instead of the time-derivatives of position parameters. From the geometrical standpoint, one should observe that, in analytical dynamics, the position q ranges in some differential manifold, while for each q the possible velocities are elements of the tangent space to the manifold at this point. The latter is a *vector space* relatively to which such concepts as linearity, convexity, etc. make sense, while no algebraic concept of this sort is generally available in the position manifold.

Contact Dynamics procedures rest on drawing the balance of momentum of the investigated mechanical system over each time-step. No estimation of the acceleration is needed and the burden of calculating mathematically the curvatures of the involved surfaces is avoided. These curvatures are simply accounted for by the fact that, from one step to the other, the normal directions to the detected contacts vary.

The method results in time-stepping schemes which, at least in what concerns the velocity function, are of the *implicit* type. For this reason, a single computation step is needed in particular to decide whether *equilibrium* in a given position is a possible motion.

In principle, the Contact Dynamics algorithms are ready at each step to treat collisions on the same footing as persistent contacts but, of course, the need of physical information about such phenomena is not overcome.

1.5 Contents of the paper

The contents of Sect. 2 is purely kinematical, in other words it pertains to differential and integral calculus in the variables of time and space. There is shown how inequality requirements imposed on a moving point in $\mathbf{R}^n$ may be enforced through conditions involving its velocity vector. The example of

the Sweeping Process, which is governed by a differential inclusion of first order, provides an introduction to the mathematical aspects of Unilateral Dynamics.

In Sect. 3 is developed the elementary example of the dynamics of a punctual particle unilaterally confined by a frictionless material boundary with prescribed motion. The traditional aim of rational mechanics, that of eliminating the unknown reactions of frictionless constraints, is attained in the form of a differential inclusion. In this simple mechanical example, the decisive step leading to the Contact Dynamics method is introduced. It consists in replacing the familiar formulation of frictionless contact by an apparently more complicated relationship involving velocities. The advantage of this transformation is illustrated by its application to a time-stepping scheme for the numerical approximation of the motion.

The framework for treating multibody multicontact systems is constructed in Sect. 4. This starts, as usual, with the parametrization of the system positions by means of an element q of $\mathbf{R}^n$, after what the geometrical conditions of non-interpenetration of the system members and their possible confinement by external obstacles are assumed expressed by a finite set of inequalities. A contact corresponds to one of these inequalities being satisfied as equality. At every instant such that the derivative $dq/dt = u \in \mathbf{R}^n$ exists, every possible contact involves a vector $\mathcal{U}$ of local relative velocity of the contacting objects which, generically, is expressed as an affine function of u. As far as the contact actions consist of a simple force $\mathcal{R}$, calculating the element r of $\mathbf{R}^n$ made of its generalized components is only the matter of matrix transposition. If the traditional assumption of differentiability of $t \mapsto u$ is made, the equations of the system dynamics may be written down through Lagrange's formalism or any other method of stereodynamics.

Exploiting these equations requires some information about the possible contact forces. Such an information receives a codified form in Sect. 5 under the name of a *contact law*, generically a relationship between $\mathcal{R}$ and $\mathcal{U}$, usually depending on the actual configuration attained by the system. Here again, the special case of frictionless contact is taken as an introduction. It allows one to stress the importance of the concept of a 'prospective' contact law. Roughly speaking such a law, rather than connecting the actual values of $\mathcal{R}$ and $\mathcal{U}$, pertains to their limits on the right of the concerned instant. The Coulomb law of dry friction receives an adaptation to such a formalism.

Then comes in Sect. 6 the application of the preceding to the construction of a time-stepping numerical scheme for the numerical approximation of solutions, first developed in the traditional framework of smooth motions, i.e. the unknown function $t \mapsto u$ is assumed locally absolutely continuous. This scheme is of the implicit type with regard to u. Due to the form given to contact laws and to their discretized forms, this time-stepping procedure automatically secures the preservation of the non-interpenetration conditions and manages the possibility of contact breaking.

In contrast, the event of the sudden introduction of new contacts, i.e. collisions, leads in Sect. 7 to giving up the framework of smooth dynamics, since u is expected to exhibit jumps. Henceforth, the function u is assumed to be of (locally) bounded variation on the considered time-interval. With such a function, there classically is associated an $\mathbf{R}^n$-valued measure du on the time-interval, which may be called the differential measure of u. Smooth dynamics is retrieved as the special case where du possesses a density function with regard to the Lebesgue measure dt.

The measure du presents an atom at every instant of discontinuity of u, but this functional framework may also accomodate other sorts of frictional catastrophes, as referred to in 7.1. Instead of a classical differential equation, dynamics is now governed by a measure-differential equation. Contact actions are not necessarily represented as time-dependent forces but as impulsion measures. Their integrals on a time-interval constitute impulsions, a familiar concept in traditional dynamics.

Since the time-stepping numerical procedure of Sect. 6 consists of drawing the balance of impulsions on each time-step, it remains formally applicable to non-smooth evolutions. There does not seem objectionable to apply the precedingly identified contact laws so long as the measure du is diffuse, i.e. u continuous. The case of a collision occuring at some instant t_c is more embarrassing since each contact-impulsion throughout the system is liable to possess an atom at point t_c. The weight of such an atom is nothing but the vector traditionally called a percussion, say $\mathcal{P}$. In 7.5 an efficient computation trick is introduced, by which $\mathcal{P}$ is connected, through Coulomb's law in the prospective form, with some weighted mean of the pre- and post-collision values of the local velocity $\mathcal{U}$. The resulting calculation is definitely richer than the traditional laws of restitution in that all contacts present in the system at instant t_c are collectively taken into account, but its phenomenological validity has to be tested in each application context.

Nonsmooth dynamics does not allow one to draw energy balances as precise as those traditionally available for smooth evolutions. The reason is that some rules of the differential and integral calculus have to be replaced by a calculus of differential measures, shortly presented in Sect. 8. The thermodynamic correctness of collision models requires dissipativity, a property which is not always confirmed by energy balances, as drawn there.

As final illustrations, two examples of the numerical application of the foregoing are presented.

In Sect. 9 a two-dimensional model of a stone bridge is used to show how the collapse mechanism under some localized forcing depends on the value admitted for the interblock friction coefficient.

Section 10 is devoted to the numerical simulation of the construction of a conical pile of grains. Eager controversy took place in recent years about the experimental finding of a local minimum of ground pressure at the vertical of the pile apex. The numerical exploration of stress in a numerical simulation

conducted by the CD method appears to remove any paradoxical features from this subject.

2 The differential handling of inequality conditions

2.1 Expressing viability in terms of velocity

Let us understand by a *moving point* a mapping q of a (time) interval I into $\mathbf{R}^n$. We want $q(t)$ to comply for every t in I with some inequality

$$f(t, q(t)) \leq 0, \tag{1}$$

where $f : I \times \mathbf{R}^n \to \mathbf{R}$ denotes a $\mathcal{C}^1$ real function. In other words, $q(t)$ is required to belong for every t to $\Phi(t) := \{x \in \mathbf{R}^n \mid f(t,x) \leq 0\}$.

It is assumed that, for t in I and x in $\mathbf{R}^n$, the *gradient* $\nabla f(t,x) := (\partial f/\partial x^1, \ldots, \partial f/\partial x^n)$ is a nonzero n-vector.

Let some t be such that the *right-side derivative* $q'^+(t)$, called the *right-side velocity* of the moving point, exists. Through the chain rule, the real function $\tau \mapsto f(\tau, q(\tau))$ is found to possess at $\tau = t$ a right-side derivative equal to $f'_t(t, q(t)) + q'^+(t).\nabla f(t, q(t))$. This derivative should be ≤ 0 if (1), assumed to hold everywhere in I, is satisfied at t as *equality*. In contrast, if inequality holds strictly at t, no sign condition comes to restrain right-side derivatives.

For t in I and x in $\mathbf{R}^n$, put

$$\Gamma(t,x) := \begin{cases} \{v \in \mathbf{R}^n \mid f'_t(t,x) + v.\nabla f(t,x) \leq 0\} & \text{if} \quad f(t,x) \geq 0 \\ \mathbf{R}^n & \text{otherwise.} \end{cases}$$

so that the above observation means $q'^+(t) \in \Gamma(t, q(t))$.

What follows may be viewed as providing a *converse*.

Suppose that the interval I, nonnecessarily compact, contains its origin t_0 and that q is *locally absolutely continuous* on I. Equivalently, the (two-side) derivative dq/dt exists almost everywhere in I and equals a function $u : I \to \mathbf{R}^n$, the *velocity function*, which is locally integrable with regard to the Lebesgue measure on I; notation : $u \in \mathcal{L}^1_{\text{loc}}(I; \mathbf{R}^n)$, meaning that u is Lebesgue-integrable over every compact subinterval of I. And, for every t in I, one has

$$q(t) = q(t_0) + \int_{t_0}^{t} u(s)\, ds. \tag{2}$$

The following is established in [58]:

Viability Lemma. *Assume that q is locally absolutely continuous on I and that the inclusion*

$$\frac{dq}{dt} \in \Gamma(t, q(t)) \tag{3}$$

holds for almost every t in I. If inequality (1) *is verified at the initial instant t_0, it is verified for every subsequent t.*

The term *viability* is used in system theory and control, in particular when economic systems are concerned, to express that the trajectory of a process should remain in a specified set (one may refer to [4], a book actually developed in a too sophisticated topological context for such an elementary statement as the above Lemma to find place in it).

2.2 Selectors

A condition of the form (3) is called a *differential inclusion* [5][21]. Since the right-hand side is set-valued, uniqueness is a priori not expected to hold for the solutions (if any) to initial value problems. By a *selector* of the 'multifunction' $(t,x) \mapsto \Gamma(t,x)$, one means a single-valued function, say $(t,x) \mapsto \gamma(t,x)$, such that $\gamma(t,x) \in \Gamma(t,x)$ for every t and x. Then

$$\frac{dq}{dt} = \gamma(t, q(t)) \tag{4}$$

is a differential equation whose (locally absolutely continuous) solutions, if any, consequent to some initial condition verifying (1), meet the assumptions of the Viability Lemma, making $q(t)$ belong to $\Phi(t)$ for every subsequent t.

A basic example is provided by taking as γ the 'lazy selector' of Γ, i.e. by defining $\gamma(t,x)$ as *the element of $\Gamma(t,x)$ with minimal Euclidean norm.* When $f(t,x) < 0$, i.e. when x belongs to the interior of $\Phi(t)$, the set $\Gamma(t,x)$ consists of the whole of $\mathbf{R}^n$, so that $\gamma(t,x)$ equals $\mathbf{0}$, the zero of $\mathbf{R}^n$. Otherwise, $\Gamma(t,x)$ equals a half-space which contains $\mathbf{0}$ if $f'_t(t,x) \leq 0$, in which case $\gamma = \mathbf{0}$ again. If $f'_t(t,x) > 0$, one finds $\gamma = -(f'_t/\|\nabla f\|^2)\nabla f$, a vector oriented in the direction of decreasing f, normal at x to the hypersurface $f(t,.) =$ Const drawn through this point.

For such a choice of γ, a solution to (4) consequent to some initial position $q(t_0)$ in $\Phi(t_0)$ may be described as follows. The point $q(t)$ belongs for every t to the moving region $\Phi(t)$. As long as it lies in the interior of $\Phi(t)$, q stays at rest. It is only when the boundary of $\Phi(t)$, i.e. the hypersurface with equation $f(t,.) = 0$, moves inward and reaches q that the point takes on a velocity in inward normal direction, so as to go on belonging to $\Phi(t)$. The magnitude of the velocity vector equals the 'normal speed' of the hypersurface.

We have proposed to call *Sweeping Process* the above kinematical association of point motions to the given motion of a set (in $\mathbf{R}^n$ or in a real Hilbert space). See [42][59] for references and recent developments of the subject.

2.3 Characterizing the process by a differential inclusion

If, at time t, a point x lies on the hypersurface $f(t,.) = 0$, the vector $\nabla f(t,x)$ (we have assumed it nonzero) normal to this hypersurface is directed outward

of the region $\Phi(t)$. The half-line emanating from the origin of $\mathbf{R}^n$, generated by $\nabla f(t,x)$, is said to constitute the (outward) *normal cone* to $\Phi(t)$ at point x; notation: $\mathrm{N}_{\Phi(t)}(x)$. The definition of a normal cone for less regular sets than Φ is a vast subject; some other cases will be met in the sequel. For x in the interior of $\Phi(t)$, it proves consistent to view $\mathrm{N}_{\Phi(t)}(x)$ as reduced to the zero of $\mathbf{R}^n$, while the cone shall be defined as empty if $x \notin \Phi(t)$. By discussing the various cases occuring in the calculation of the lazy selector, one sees that if $\gamma(t,x)$ equals this selector, every solution q to the differential equation (4) verifies, for almost every t, the differential inclusion

$$-\frac{dq}{dt} \in \mathrm{N}_{\Phi(t)}(q(t)). \tag{5}$$

Unexpectedly *the converse is true*, i.e.(5) in spite of its multivalued right-hand side actually is *equivalent* to the differential equation (4), as far as locally absolutely continuous solution are concerned.

In fact let $q : I \to \mathbf{R}^n$, be such a solution to (5). For almost every t, the two-side derivative $q' = dq/dt$ exists, so that the right-hand member is non-empty and therefore $q(t) \in \Phi(t)$; the same is true for every t, by continuity. For t such that $q(t)$ happens to lie in the interior of $\Phi(t)$, (5) implies $q' = \mathbf{0}$, which makes that (4) is also satisfied. Otherwise, suppose that $q(t)$ belongs to the boundary, i.e. the function $\tau \mapsto f(\tau, q(\tau))$ vanishes at $\tau = t$. Then the right-derivative $f'_t(t,q(t)) + q'^{+}(t).\nabla f(t,q(t))$, if it exists, is ≤ 0 while, symmetrically, the left-derivative is ≥ 0. Therefore $q'(t)$, when it exists, satisfies $f'_t(t,q(t)) + q'(t).\nabla f(t,q(t)) = 0$, i.e. it belongs to the boundary of the half-space $\Gamma(t,q(t))$. Furthermore, (5) entails that $q'(t)$ is directed along the inward normal to the half-space. All this elementarily characterizes $q'(t)$ as the proximal point to $\mathbf{0}$ in $\Gamma(t,q(t))$, namely $\gamma(t,q(t))$.

It was under the formulation (5) that the Sweeping Process was primitively introduced [47][50], with $\Phi(t)$ denoting a nonempty closed convex subset of a real Hilbert space H. The motivation then was in the quasi-static evolution of elastoplastic systems [48][49]. The convexity assumption allows one to establish the existence of solutions under rather mild conditions concerning the evolution of $\Phi(t)$, even discontinuous. Another consequence of this convexity is that the multifunction $x \mapsto \mathrm{N}_{\Phi(t)}(x)$ is *monotone* in the following sense (see e.g. [8]) : whichever are x_1, x_2 in H, y_1 in $\mathrm{N}_{\Phi(t)}(x_1)$, y_2 in $\mathrm{N}_{\Phi(t)}(x_2)$, one has $(x_1 - x_2).(y_1 - y_2) \geq 0$, with the dot denoting the scalar product of H. By elementary calculation, this inequality entails that, if $t \mapsto q_1(t)$ and $t \mapsto q_2(t)$ are two solutions to (5), the Hilbert distance $\|q_1 - q_2\|$ is a non-increasing function of t. From this *non expansion* property, it follows that at most one solution to (5) can agree with some initial position $q(t_0)$.

Another source of interest of the formulation (5) is to render evident that the successive positions of the point q are connected with those of the given region Φ in a *rate-independent* way. In fact, because the right-hand member is a cone, the differential inclusion is found invariant under any non-decreasing differentiable change of variable.

2.4 Implicit versus explicit time-stepping

Coming to the numerical approximation of solutions through time-stepping schemes, let us denote by $[t_\mathrm{i}, t_\mathrm{f}]$, with length h, a time-step ('i' as in *initial*, 'f' as in *final*). From an estimate q_i of $q(t_\mathrm{i})$, obtained as the result of the antecedent time-step, computation has to deliver an estimate q_f of $q(t_\mathrm{f})$.

The formulation (4) naturally leads to take $u_\mathrm{i} = \gamma(t_\mathrm{i}, q_\mathrm{i})$ as an estimate of the velocity throughout the time-step, generating the prediction $q_\mathrm{f} = q_\mathrm{i} + hu_\mathrm{i}$. This is a computation scheme of the *explicit* type.

If (5) is discretized by viewing $(q_\mathrm{f} - q_\mathrm{i})/h$ as a representative of the velocity, a strategy of the explicit type would not allow one to express q_f, since the right-hand member is multivalued. In contrast, the *implicit* strategy consists in invoking the value that this right-hand member would take at the unknown point, so one has to solve

$$q_\mathrm{i} - q_\mathrm{f} \in \mathrm{N}_{\Phi(t_\mathrm{f})}(q_\mathrm{f}) \tag{6}$$

(the positive factor h has been dropped since $\mathrm{N}_{\Phi(t_\mathrm{f})}$ is a cone). This qualifies q_f as an *orthogonal projection* of q_i onto $\Phi(t_\mathrm{f})$. In the case where $\Phi(t_\mathrm{f})$ is convex, the projection is unique and this characterizes q_f as the nearest point to q_i in $\Phi(t_\mathrm{f})$. In particular $q_\mathrm{f} = q_\mathrm{i}$ when q_i happens to belong to $\Phi(t_\mathrm{f})$. We have proposed to call this procedure the *catching-up algorithm* [50].

2.5 Complementarity

From the description made of the Sweeping Process in 2.2 it is clear that velocity may be discontinuous. As for the explicit prediction $q_\mathrm{f} = q_\mathrm{i} + hu_\mathrm{i}$, it only requires of u_i to be the derivative of the function q *on the right* of t_i. The following observation makes an introduction to analogous, but more complicated, situations we are to meet in Dynamics.

Let q, associated with u through (2), verify (5) almost everywhere in I. Let $t_1 \in I$ and assume that the function u possesses a limit on the right of t_1, say u_1^+; in view of (2) this limit also provides the right derivative $\dot{q}^+(t_1)$. As a stronger assertion than (4), which pertained to bilateral derivative and was only declared to hold almost everywhere, let us prove that $u_1^+ = \gamma(t_1, q(t_1))$.

Trivially, both members of this equality are $\mathbf{0}$ if $f_1 := f(t_1, q(t_1)) < 0$. Otherwise, i.e. when $f_1 = 0$, it was seen that

$$\dot{f}_1^+ = f_t'(t_1, q(t_1)) + u_1^+ . \nabla f(t_1, q(t_1)) \leq 0. \tag{7}$$

Inclusion (5) means the existence of a function $t \mapsto \lambda(t) \leq 0$ such that $u(t) = \lambda(t) \nabla f(t, q(t))$. Since ∇f is continuous and nonzero, the assumed existence of u_1^+ secures that of the right-limit λ_1^+ and

$$u_1^+ = \lambda_1^+ \nabla f(t_1, q(t_1)). \tag{8}$$

If $\dot{f}_1^+ < 0$, instant t_1 is followed by an interval throughout which $f < 0$. This has been observed to imply $u = 0$, so that λ vanishes on this interval and consequently also its right-limit λ_1^+. Summing up, one has

$$\dot{f}_1^+ \leq 0, \quad \lambda_1^+ \leq 0, \quad \dot{f}_1^+ \lambda_1^+ = 0, \tag{9}$$

a system of *complementarity conditions*. This is a popular formalism in many research domains where inequality requirements are faced. Solving (8) (9), with $\dot{f}_1^+$ defined as in (7), constitutes a *linear complementarity problem*. Through arguments from Convex Analysis, such problems are shown to be equivalent to finding the critical points of some quadratic functions over polyhedral convex sets. In the present setting, where inequality $\dot{q}_1^+ \leq 0$ simply expresses that u_1^+ belongs to the half-space $\Gamma(t_1, q(t_1))$, one readily checks that the above system of conditions characterizes u_1^+ as the minimizing point of the function $x \mapsto \|x\|^2/2$ over $\Gamma(t_1, q(t_1))$, namely $\gamma(t_1, q(t_1))$ as announced.

2.6 A hydromechanical illustration

Assume that $n = 2$ and that t, x^1, x^2 are Cartesian coordinates in physical space, with the t-axis vertical and directed downward. Picture the region $f(t, x^1, x^2) \leq 0$ as an underground cavity and the curve $x^1 = q^1(t), x^2 = q^2(t)$ as a stationary waterstream dripping down into it.

The differential equation (4), on account of the diverse circumstances met in the definition of the lazy selector, expresses that: i) any part of this stream which happens to be detached from the cavity wall is rectilinear and vertical; ii) when water runs on the wall, it follows a *line of steepest descent* (this agrees with hydrodynamics under the simplifying assumption that inertia effects are negligible with regard to gravity and to liquid/wall friction); iii) the dependence of γ on the sign of f_t' makes that the stream can run only on a part of the wall exposed upward: when it reaches the rim of a possible overhang, water gets loose and falls vertically down as described in i).

In this example, under the complication typically added by unilaterality, the comparison of (4) and (5) merely reflects the classical equivalence between the two standard properties of the lines of steepest descent in a surface: at each point on such a line i) the slope is maximal; ii) the direction is orthogonal to the level curve of the surface.

3 Frictionless confinement of a particle

3.1 Primary formulation

Notations are the same as in Sect.2, with $n = 3$. The element $q := (q^1, q^2, q^3)$ now consists of the *orthonormal* coordinates of a material point Q, with mass m, moving under the action of a given force field $(t, x) \mapsto X(t, x)$ and constrained in the region $\Phi(t)$ by the impenetrability of its boundary, assumed

to be realized as a material surface with imposed motion. This is the occasion of recalling that, in Mechanics, the description of a *constraint* never reduces – as it does, for instance, in Optimization or in Abstract System Theory – to imposing a geometrical restriction on positions. Some information is needed about the mechanical process through which this restriction is enforced. For instance, using some *servomechanism* in order to secure (1) could result in quite different motions than those obtained under the present assumption of confinement by the *contact action* of a given material boundary.

On a time interval I throughout which the motion is smooth enough for the *velocity function* u in (2) to be locally absolutely continuous (this precludes collisions, events to which we shall come back later), the motion of the particle Q obeys, almost everywhere in I,

$$m\frac{du}{dt} = X(t, q(t)) + r(t), \tag{10}$$

where the force $r := (r^1, r^2, r^3)$ denotes the unknown *reaction* possibly exerted by the confining boundary. In this Section, the confinement process is assumed to comply at every instant with the following model:

- this is a *contact* process

$$f(t, q) < 0 \quad \Rightarrow \quad r = 0, \tag{11}$$

- the possible contact is *frictionless*

$$f(t, q) = 0 \quad \Rightarrow \quad \exists \lambda \in \mathbf{R} : r = \lambda \nabla f(t, q), \tag{12}$$

- without adhesion

$$\lambda \leq 0. \tag{13}$$

If compared with the definition given in 2.3 for the normal cone $\mathrm{N}_{\Phi(t)}(x)$ at a point x, the above system of three conditions is found equivalent to

$$-r \in \mathrm{N}_{\Phi(t)}(q). \tag{14}$$

Therefore, the traditional aim of Analytical Dynamics, namely the elimination of the unknown reactions of the so-called perfect constraints is attained by rapproaching conditions (10) and (14)

$$X(t, q(t)) - m\frac{du}{dt} \in \mathrm{N}_{\Phi(t)}(q(t)), \tag{15}$$

an integro-differential inclusion, as the unknown functions q and u are essentially connected by (2). Formally, this inclusion implies $q(t) \in \Phi(t)$ for almost every t in I, since otherwise the right-hand member would be empty. By continuity $q(t) \in \Phi(t)$ holds for every t.

3.2 The 'Contact Dynamics' approach

From the theoretical standpoint as well as in the elaboration of approximation schemes, the differential inclusion (15) proves difficult to handle (see however [64], under the assumption of convexity for the function f). In what concerns the position function q, the problem at hand is of differential order 2, as expected in a dynamical context, while in the purely kinematical setting of Sect. 2, q was the unknown of a differential problem of order 1. However, some of the arguments used then will be transposed in what follows.

Let Γ be defined as in 2.1. *Provided the initial position $q(t_0)$ lie in $\Phi(t_0)$, the solutions of* (15) *are the same as those of*

$$X(t,q(t)) - m\frac{du}{dt} \in \mathrm{N}_{\Gamma(t,q(t))}(u(t)). \tag{16}$$

In fact (16), assumed to hold for almost every t in I, implies $u(t) \in \Gamma(t,q(t))$. Since $q(t_0) \in \Phi(t_0)$, the Viability Lemma secures $q(t) \in \Phi(t)$ for every t in I. The definition of Γ then makes that the normal cone $\mathrm{N}_{\Gamma(t,q(t))}(u)$ is contained in $\mathrm{N}_{\Phi(t)}(q(t))$ whatever is u, so that (15) is a fortiori satisfied.

Conversely, let q be a solution to (15), hence $q(t) \in \Phi(t)$ for every t in I. When $f(t,q(t)) < 0$, the set $\mathrm{N}_{\Phi(t)}(q(t))$ reduces to the singleton $\{\mathbf{0}\}$ and the same is true for the right-hand member of (16). In contrast, for t such that $f(t,q(t)) = 0$, the set $\Gamma(t,q(t))$ is a half-space and the argument used in 2.3 proves that the element $u(t) = u^+(t) = u^-(t)$ belongs to its boundary plane. Consequently, $\mathrm{N}_{\Gamma(t,q(t))}(u)$ consists of the cone generated by $\nabla f(t,q(t))$ hence equals $\mathrm{N}_{\Phi(t)}(q(t))$.

3.3 First example of a *CD* numerical scheme

As before, let us denote by $[t_\mathrm{i}, t_\mathrm{f}]$, with length h, a time-step. From the approximate values $q_\mathrm{i}, u_\mathrm{i}$ obtained for q and u at t_i, one has to calculate $q_\mathrm{f}, u_\mathrm{f}$, pertaining to t_f. The given force field X is assumed to depend smoothly on its arguments, so that one chooses to approximate it throughout the time-step by the value it takes at $t_\mathrm{m} := t_\mathrm{i} + h/2$ and $q_\mathrm{m} := q_\mathrm{i} + hu_\mathrm{i}/2$. It is also at the point $(t_\mathrm{m}, q_\mathrm{m})$ that f is calculated, in order to decide whether boundary contact is in effect or not and to determine the set Γ accordingly. Depending on the sign of $f(t_\mathrm{m}, q_\mathrm{m})$, the latter equals the whole of $\mathbf{R}^3$ or a half-space with $\nabla f(t_\mathrm{m}, q_\mathrm{m})$ as normal vector. Inclusion (16) is thus discretized in the form

$$X(t_\mathrm{m}, q_\mathrm{m}) - \frac{m}{h}(u_\mathrm{f} - u_\mathrm{i}) \in \mathrm{N}_{\Gamma(t_\mathrm{m},q_\mathrm{m})}(u_\mathrm{f}), \tag{17}$$

i.e. in view of the right-hand member being a cone,

$$u_\mathrm{i} + \frac{h}{m}X(t_\mathrm{m}, q_\mathrm{m}) - u_\mathrm{f} \in \mathrm{N}_{\Gamma(t_\mathrm{m},q_\mathrm{m})}(u_\mathrm{f}).$$

This classically characterizes u_f as the *proximal point* to the known element $u_i + hX(t_m, q_m)/m$ in $\Gamma(t_m, q_m)$. One finishes the calculation with

$$q_f = q_m + \frac{h}{2} u_f.$$

This algorithm automatically handles the possible breaking of contact : this happens if $u_i + hX(t_m, q_m)/m$ falls into the interior of $\Gamma(t_m, q_m)$.

Remark 1.– Provided $q(t_0) \in \Phi(t_0)$, inclusion (16) has been precedingly observed to secure $f(t, q(t)) \leq 0$ at every consequent t. But, in the above time-stepping procedure, it is only u_f which, at each step, is constructed as an element of $\Gamma(t_m, q_m)$, so that the Viability Lemma is just involved through time-discretization. One thus may fear that some violation of the inequality would build up from step to step. Actually, if the step-length is not too large, numerical experiments show some self-corrective effect which, on the contrary, tends to reduce violations. This effect seems related to $\Phi(t)$ possessing a nonempty interior. In contrast, in industrial softwares devoted to the dynamics of machines, some *bilateral* constraints are commonly introduced, leaving a set of feasible positions with empty interior. Their treatment in terms of velocities [28] then requires corrective procedures to prevent cumulative errors.

Remark 2.– The calculation of u_f from u_i is based on mechanical elements evaluated at the mid-position q_m; in turn, u_f is used to calculate from q_m the final position of the current step and, from there, the mid-position of the subsequent time-step. This interleaving makes the above time-stepping procedure resemble the policy called 'leapfrog' in Molecular Dynamics simulations. In order to figure out how it improves precision, compared with an ordinary Euler explicit time-stepping scheme, one may apply it to calculating the parabolic unconstrained motion of a particle in a uniform gravity field. If h is constant, the values found for q at the successive steps coincide with the exact solution, while the Euler scheme generates cumulative errors. Of course, when an algorithm of the above sort is implemented with constant h, one may calculate each q_m from the antecedent one by a single incrementation. The proper output of the computation, namely q_f, may not be needed at each step.

Remark 3.– In contrast with the approach of the motion of a point in a surface through the traditional methods of dynamics, the above time-stepping procedure does not require calculating the curvature of the boundary. This curvature is implicitely accounted for by the fact that the direction of ∇f evolves from one step to the other. The question of existence of solutions to (15) or (16) is not addressed in this paper; one naturally expects that f has to be twice dfferentiable.

4 Multicontact systems

4.1 Parametrization

Let the possible configurations of a body collection be parametrized (at least locally) through generalized coordinates, say $q := (q^1, \ldots, q^n)$. For the sake of reducing the number n, this parametrization may be constructed with account of possible permanent, frictionless linkages imposed on the members of the system. After that, the constraints of *non-interpenetrability* are additionally considered. The geometric restriction consequently imparted on the system positions is assumed expressed by a finite set of inequalities

$$f_\alpha(t, q) \leq 0, \quad \alpha \in \{1, \ldots, \kappa\}, \tag{18}$$

where $f_1, \ldots, f_\kappa$ are given functions. Through the presence of t as an argument of f_α, provision is made for the inequality to describe the confinement of a member of the system by some external obstacle or boundary with prescribed motion. Under this parametrization, a motion of the system consists of a mapping $t \in I \mapsto q(t) \in \mathbf{R}^n$ and, as in the foregoing, this mapping is assumed locally absolutely continuous, i.e. there exists a locally integrable velocity function $u : I \to \mathbf{R}^n$ from which q may be retrieved in the form (2).

As an example, one may consider a pair of members of the system whose positions in a chosen reference frame are well located as soon as the value of the element q of $\mathbf{R}^n$ is known (together with the time t in case of a time-dependent parametrization). Then, one may take as f_α the expression, as a function of (t, q), of some measure of the *overlap* of the two bodies. This overlap should be understood as a directed quantity, so that it becomes negative in case the bodies lie apart from each other. The convention applied in (18) of characterizing the permitted configurations through the ≤ 0 inequality, comes from Convex Optimization theory, where such a sign convention offers technical advantages. No convexity hypothesis is made here concerning the functions f_α: since such an assumption would not be preserved under a change of parametrization, it cannot in general have any mechanical meaning. If one prefers to deal with the ≥ 0 symbol, there only is to consider, instead of the overlap the opposite quantity, usually called the *gap* between the considered bodies.

The above formalism is not limited to collections of strictly rigid bodies, since q may also include parameters accounting for a finite-freedom approximation of deformability. Such additional parameters possibly arise from some modal representation of the deformation dynamics or from the Finite Element discretization of deformable parts [37].

4.2 Contact kinematics

Suppose that inequality $f_\alpha \leq 0$ expresses the local non-interpenetration of some pair of members of the system, say $\mathcal{B}$ and $\mathcal{B}'$, so that equality $f_\alpha =$

0 corresponds to these bodies touching each other at some point of space denoted by M_α. This we shall assume here to be an isolated contact point, but other contacts, associated with different values of α, may also be in effect between the same bodies at the same instant. For every imagined motion $t \mapsto q(t)$ bringing the system through the considered contacting position for some t, with a definite value of $u = dq/dt \in \mathbf{R}^n$, the velocity vectors $\mathcal{V}_\alpha$ and $\mathcal{V}'_\alpha$, relative to the chosen reference frame, of the respective particles of $\mathcal{B}$ and $\mathcal{B}'$ passing at point M_α let themselves be expressed as affine functions of u. The same is thus true for the *relative velocity* $\mathcal{U}_\alpha = \mathcal{V}_\alpha - \mathcal{V}'_\alpha$ of body $\mathcal{B}$ with respect to body $\mathcal{B}'$ at this point, say

$$\mathcal{U}_\alpha = G_\alpha \, u + \mathcal{W}_\alpha, \tag{19}$$

where $G_\alpha : \mathbf{R}^n \to \mathbf{E}^3$ (the space of the vectors of physical space) denotes a linear mapping, depending on t and q. No attention is paid at this stage to the imagined motion preserving contact or not. The term $\mathcal{W}_\alpha \in \mathbf{E}^3$, a known function of t and q, vanishes in the familiar case of a time-independent parametrization.

Similar formula holds if inequality $f_\alpha \leq 0$ expresses the confinement of a member $\mathcal{B}$ of the system by some external material boundary with prescribed motion. Assume that equality $f_\alpha = 0$ corresponds to contact taking place at some point, here again denoted by M_α. The local velocity, at this point, of body $\mathcal{B}$ with respect to the boundary has the same form as $\mathcal{U}_\alpha$ in (19), where $\mathcal{W}_\alpha$ now reflects the known velocity of the boundary (for a time-independent parametrization, $\mathcal{W}_\alpha$ equals the negative of this velocity vector).

At the contact point M_α, we assume that a *common tangent plane* to the respective surfaces of the concerned bodies has been defined. This does not require of both surfaces to be smooth; for instance, contact may take place between a smooth body and some corner point or sharp asperity of the other. Let $\mathbf{n}^\alpha$ denote the unit vector normal to this plane, *directed toward* $\mathcal{B}$. In computation, as well as in existential studies, it proves useful that the definition of the above elements would be conventionally extended to a neighbourhood of the concerned value of (t, q) in $\mathbf{R} \times \mathbf{R}^n$. This allows one to express as a function of (t, q) the *normal gap*, say $g_\alpha(t, q)$, between $\mathcal{B}$ and $\mathcal{B}'$, *counted as negative in the case of overlap*. Classically, the derivative of the function $t \mapsto g_\alpha(t, q(t))$ is found equal to $\mathcal{U}_\alpha . \mathbf{n}^\alpha$, the normal component of the relative velocity of the contacting bodies at point M_α.

Sometimes, in Computation literature, the second time-derivative of the gap is improperly referred to as the 'normal relative acceleration'. Actually, since the material particles involved in the definition of $\mathcal{U}_\alpha$ are not the same from an instant to the other, this second derivative has in general nothing to do with the relative acceleration vector. As an example, one may consider a body of circular or spherical shape: a variety of rotations may be imparted to such a body without altering its overall location, so yielding the same gap while the normal acceleration is changed.

Remark. – The representation of non-interpenetration through of a finite set of inequalities is operative in most practical situations. However, non-interpenetration cannot be described in that way in the neighbourhood of a configuration where two sharp asperities or corners come into contact by their points. This is evidenced by the fact that, in this case, the set of the values of the local right-velocity $\mathcal{U}_\alpha^+$ which are compatible with non-interpenetration is not a convex cone anymore. In contexts where the probability of such an event cannot be treated as negligible, numerical techniques have to resort to adequate procedures (possibly involving $\mathcal{U}_\alpha$) for the identification of a mechanically plausible contact plane.

4.3 Contact forces

Assume that the contact actions that body $\mathcal{B}$ experiences at point M_α from body $\mathcal{B}'$ are described as a simple force $\mathcal{R}^\alpha$ (there would be no conceptual difficulty in adding to this description some local torque, accounting for a resistance to rolling). Then $\mathcal{B}'$ experiences from $\mathcal{B}$ the force $-\mathcal{R}^\alpha$. The standard machinery of Analytical Dynamics needs a representation of this pair of forces, in regard to the chosen parametrization, through its *covariant components* (or 'generalized components'), namely the element r^α of $\mathbf{R}^n$ expressed as

$$r^\alpha = G_\alpha^* \mathcal{R}^\alpha, \tag{20}$$

with $G_\alpha^* : \mathbf{E}^3 \to \mathbf{R}^n$ denoting the *transpose* of G_α.

The convention of implicit summation will never be applied to Greek indices.

If inequality $f_\alpha \leq 0$, expresses the confinement of a member $\mathcal{B}$ of the system by some external obstacle with prescribed motion, (19) still holds with $\mathcal{U}_\alpha$ denoting the local velocity of $\mathcal{B}$ relative to this obstacle. Then it is found that r^α in (20) consist of the covariant components of the force $\mathcal{R}^\alpha$ alone, acting on $\mathcal{B}$ at the contact point. Its counterpart, exerted by $\mathcal{B}$ upon the obstacle, is not in this case a force experienced by the system. Incidentally observe that the term $\mathcal{W}_\alpha$ does not appear in (20).

4.4 The equation of Dynamics

As before, the context here is that of standard dynamics, involving the second derivative of the function $t \mapsto q(t)$. Therefore the velocity function $t \mapsto u(t)$ is required to be locally absolutely continuous. Using Lagrange's technique or any other tools from classical solid dynamics, one obtains a differential equation, to be read as an equality of elements of $\mathbf{R}^n$

$$A(t,q)\frac{du}{dt} = F(t,q,u) + \sum_\alpha r^\alpha, \tag{21}$$

where A denotes the $n \times n$ *inertia matrix*. The expression F comprises certain standard terms (commonly referred to as 'centrifugal' and 'gyroscopic') and also the covariant components of some applied forces, supposed given as functions of time, the position of the system and its velocity. The elements $r^\alpha, \alpha \in \{1, 2, \ldots, \kappa\}$, are made of the covariant components of the respective contact forces, as expressed in (20).

The same formalism remains more generally valid with q related to some velocity function u by other kinematical relations than (2). For instance, when dealing with 3-dimensional rigid bodies, it is usual to attach to each of them a frame of principal axes of inertia emanating from its center of mass. Then one may choose to enter, among the constituents of the $\mathbf{R}^n$-valued function u, the three components relative to these axes of the *spin vector* of the rigid body, instead of the time-derivatives of some directional parameters. This offers the considerable advantage of generating a contribution in the matrix A which is diagonal and constant with regard to t and q. Retrieving from these spin components the evolution of some directional parameters of the concerned body is only the matter of integrating adequate kinematical formulas. Correlatively, if forces are applied to the rigid body, the total moments of these forces about the same axes should be entered as covariant components into the corresponding lines of the right-hand side of (21).

Since contact forces vanish when contact is not in effect, the summation in the right-hand side of (21) may be restricted to the values of α belonging to

$$J(t, q) := \{\alpha \in \{1, \ldots, \kappa\} \mid f_\alpha(t, q) \geq 0\}. \tag{22}$$

The geometric conditions (18) of non-interpenetrability, joined to the differential equation of dynamics (21), clearly are not enough for determining the motion consequent to initial data. Some phenomenological information should be added, concerning the contact forces. Since the contact phenomenon takes place in physical space, this information is expected to involve the vectors $\mathcal{R}^\alpha$, $\mathcal{U}_\alpha$, as well as the values of t and q specifying the actual configuration of the system. Hence, for every α labelling a possible contact, a relationship of the form

$$\mathrm{law}_\alpha(t, q, \mathcal{U}_\alpha, \mathcal{R}^\alpha) = \mathrm{true}, \tag{23}$$

called a *contact law*, should be available.

5 Contact laws

5.1 Frictionless contact

Under the present notations we are to meet the same circumstances as in Sect. 3. If the contact labelled α concerns two bodies denoted by $\mathcal{B}$ and $\mathcal{B}'$,

with common normal unit $\mathbf{n}^\alpha$ directed toward $\mathcal{B}$, the assumptions of no-friction an no-adhesion mean $\exists \rho_\alpha \geq 0 : \mathcal{R}^\alpha = \rho_\alpha \mathbf{n}^\alpha$. It has been agreed in the foregoing to extend the definition of $\mathbf{n}^\alpha$, at least in a neighbourhood of the concerned values of t and q, to cases where g_α, the *normal gap*, takes nonzero values and to state $\mathcal{R}^\alpha$=0 if $g_\alpha > 0$. Define

$$\mathcal{K}_\alpha(t,q) := \begin{cases} \{\mathcal{V} \in \mathbf{E}^3 \mid \mathcal{V}.\mathbf{n}^\alpha \geq 0\} & \text{if} \quad g_\alpha(t,q) \leq 0 \\ \mathbf{E}^3 & \text{otherwise.} \end{cases}$$

This is the set of the values of the local *right-velocity* of $\mathcal{B}$ relatively to $\mathcal{B}'$ (the latter may be a member of the system or an external obstacle with prescribed motion) which are compatible with non-interpenetration. In the first line, $\mathcal{K}_\alpha$ equals a half-space, hence the *normal cone* $\mathrm{N}_{\mathcal{K}_\alpha}$, evaluated at the origin $\mathbf{0}$ of $\mathbf{E}^3$, equals the half-line generated in this vector space by $-\mathbf{n}^\alpha$. Otherwise $\mathcal{K}_\alpha = \mathbf{E}^3$, so that the cone $\mathrm{N}_{\mathcal{K}_\alpha}(\mathbf{0})$ reduces to the set $\{\mathbf{0}\}$.

Therefore, at time t, the no-friction and no-gluing assumptions (including the case of no-contact) are equivalent to assert

$$-\mathcal{R}^\alpha \in \mathrm{N}_{\mathcal{K}_\alpha(t,q)}(\mathbf{0}). \tag{24}$$

The move made in Sect.3 of replacing inclusion (15) by inclusion (16) admits as a counterpart here the replacement of inclusion (24) by

$$-\mathcal{R}^\alpha \in \mathrm{N}_{\mathcal{K}_\alpha(t,q)}(\mathcal{U}_\alpha), \tag{25}$$

In fact, in this context where u is continuous, the same argument as in 2.3 shows that $\mathcal{U}_\alpha$ belongs to the boundary plane of the half-space $\mathcal{K}_\alpha$, hence $\mathrm{N}_{\mathcal{K}_\alpha}(\mathcal{U}_\alpha) = \mathrm{N}_{\mathcal{K}_\alpha}(\mathbf{0})$, while in case of no-contact $\mathrm{N}_{\mathcal{K}_\alpha}(\mathcal{U}_\alpha) = \{\mathbf{0}\}$ whatever is $\mathcal{U}_\alpha$. In short, (25) contains all the stipulations implied when a contact is declared frictionless.

But, *in addition*, (25) entails $\mathcal{U}_\alpha \in \mathcal{K}_\alpha(t,q)$, since otherwise $\mathrm{N}_{\mathcal{K}_\alpha}(\mathcal{U}_\alpha)$ would be empty. If $g_\alpha(t,q) > 0$ this actually imparts no restriction on $\mathcal{U}_\alpha$ while if $g_\alpha(t,q) \leq 0$, i.e. $\alpha \in J(t,q)$, this implies $\mathcal{U}_\alpha.\mathbf{n}^\alpha \geq 0$. It has been precedingly recalled that $\mathcal{U}_\alpha.\mathbf{n}^\alpha$ equals the derivative of the function $t \mapsto g_\alpha(t,q(t))$. This allows one to invoke the Viability Lemma (2.1), with $f = -g_\alpha$, in order to prove that the assumption of (25) being verified for almost every t in I entails : *non-interpenetration holds for every* $t > t_0$, provided it holds at t_0.

The latter statement applies more generally to any contact law which, among other phenomenological stipulations, secures the following

- in all cases $\mathcal{U}_\alpha \in \mathcal{K}_\alpha$,
- if $\mathcal{U}_\alpha \in$ interior $\mathcal{K}_\alpha$, then $\mathcal{R}^\alpha = 0$.

In other words, one has the implications

$$g_\alpha(t,q) \leq 0 \Rightarrow \mathbf{n}^\alpha.\mathcal{U}_\alpha \geq 0 \quad . \tag{26}$$

$$\mathbf{n}^\alpha . \mathcal{U}_\alpha > 0 \Rightarrow \mathcal{R}^\alpha = 0. \tag{27}$$

We propose to say that a package of information, concerning the possible contact labelled α, if it possesses these two properties, is a *contact law of prospective type* (or, in the terminology precedingly used by the author [58][59], a *complete* contact law). The underlying idea is that such a law does not properly govern the values of $\mathcal{U}_\alpha$ and $\mathcal{R}^\alpha$ at the actual instant, but their limits on the right of this instant, assumed to exist. In fact, if $\mathcal{U}_\alpha . \mathbf{n}^\alpha > 0$, the concerned instant is followed by a contactless time-interval. Since $\mathcal{R}^\alpha$ vanishes over this interval, the same is true for its right-limit.

5.2 Multicontact frictionless Dynamics

We are now to see how the observations made in Sect.3 let themselves be transposed into the present setting. Even the simple case then considered, of a particle confined by a frictionless material boundary, may exhibit multicontact features if the boundary consists of several parts with smooth equations, each of them enforcing an inequality of the form (18). If two of these smooth surfaces meet to form an edge, the particle, when lying on this edge, experiences contact forces from both parts, the resultant of which may take any value in the convex cone generated by the respective inward normals. Therefore, the writing in (14) remains valid provided $\mathrm{N}_{\Phi(t)}(q)$ is defined as the convex cone generated by the outward normals to the surfaces which form the edge.

When coming to general multicontact systems, one has to connect normality in the linear Euclidean space $\mathbf{E}^3$ of the vectors of physical space with normality in the space $\mathbf{R}^n$ of the abstract components. The following relationship is found to hold [52] between the element ∇f_α of $\mathbf{R}^n$ and the normal unit vector $\mathbf{n}^\alpha$ at point M_α to the contacting bodies, directed toward $\mathcal{B}$

$$\exists \lambda_\alpha \geq 0 \quad \text{such that} \quad G_\alpha^* \, \mathbf{n}^\alpha = -\lambda_\alpha \, \nabla f_\alpha. \tag{28}$$

The proof of this rests on a unilateral version of the algebraic theorem of Lagrange multipliers, known in Convex Analysis as *Farkas' lemma* [69].

In all the sequel, we shall assume that the mapping G_α is *surjective* of $\mathbf{R}^n$ onto $\mathbf{E}^3$; equivalently, its transpose G_α^* is *injective* of $\mathbf{E}^3$ into $\mathbf{R}^n$. Then λ_α in (28) is nonzero. Some special positions of a multibody system may give rise to 'wedging' effects which contradict this assumption.

In view of (19) and (20) this allows one to replace the laws of frictionless contact, either (24) or (25), by equivalent relationships involving only the abstract components u and r^α instead of $\mathcal{U}_\alpha$ and $\mathcal{R}^\alpha$. Under the definition (22) of J, put

$$W(t,q) := \{ v \in \mathbf{R}^n \,|\, \forall \alpha \in J(t,q) \, : \, \frac{\partial f_\alpha}{\partial t} + v . \nabla f_\alpha \leq 0 \}, \tag{29}$$

a polyhedral closed convex set. One then finds [59] that a value r of the sum $\sum_\alpha r^\alpha$ is compatible with the contact law (25) holding for all α, if and only if $-r \in \mathrm{N}_{W(t,q)}(u)$. Consequently, the elimination of frictionless reactions from the dynamical equation (21) is achieved in the writing

$$F(t,q,u) - A(t,q)\,\frac{du}{dt} \in \mathrm{N}_{W(t,q)}(u). \tag{30}$$

This allows one to derive a time-stepping scheme for the numerical approximation of solutions quite similar to the one presented in 3.3. The occurrence of the matrix A in (30) at the place occupied in (16) by the scalar factor m does not constitute an essential complication. It only means that, instead of the standard Euclidean metric of $\mathbf{R}^3$, one is using in $\mathbf{R}^n$ the Euclidean metric defined by the positive definite matrix A.

5.3 Handling inequality conditions in terms of acceleration

The differential inclusion (30) formally resembles (5) which has been found to characterize the Sweeping Process. In fact, in the special case where $F = 0$ and $A = \mathbf{1}$, (30) makes the function $t \mapsto u$ appear as a solution to the Sweeping Process by the moving (closed, convex) set $t \mapsto W(t, q(t))$. Of course, the latter is not given, since it depends on q which itself is connected to u by (2), but one may infer from this analogy that, similarly to the equivalence of (5) to (4), the inclusion (30) could be replaced by a differential *equation* whith right-hand member defined by a minimization property. This is the object of the forthcoming.

A time-stepping scheme of the *implicit* type, rests on predicting the velocity whithout resorting to any expression of the acceleration. In contrast, what follows is aimed at determining the *right-acceleration* $\dot{u}^+$, which is needed when a scheme of the *explicit* type is being planned, as well as an Event Driven calculation [1][65].

Let t_1 denote an instant preceded by some time-interval throughout which the motion, with absolutely continuous u, satisfies (30). Assume that u remains continuous at instant t_1, i.e. *no collision occurs*. Non-interpenetration entails that, for every α in $J(t_1, q(t_1))$, the right-derivative of the function $t \mapsto f_\alpha(t, q(t))$ at t_1 is ≤ 0, while the left-derivative is ≥ 0. Hence this function has zero derivative of order one at t_1.

Let us assume in this Subsection that the functions f_α are $\mathcal{C}^2$. The investigation of the motion by explicit time-stepping or by an ED policy rests on the assumption that t_1 is followed by a nonzero interval throughout which (30) is verified again. It just may happen that some contacts break at t_1, inducing a change of $J(t,q)$. If u possesses a derivative on the right of t_1, the function $t \mapsto f_\alpha(t, q(t))$ possesses a second derivative on the right of t_1, expressed through the chain rule with a certain a_α, in the form $a_\alpha(t_1, q_1, u_1) + \dot{u}_1^+ . \nabla f_\alpha(t_1, q_1)$. Since the function is zero at t_1, as well as its

first derivative, non-interpenetration requires of this second derivative to be ≤ 0. If it is strictly negative, $f_\alpha(t, q(t))$ becomes strictly negative on a subsequent interval, making the corresponding quantities $r^\alpha = \mu_\alpha \nabla f_\alpha$ vanish on this interval, as well as their limits for $t \downarrow t_1$, assumed to exist. One thus obtains a set of complementarity conditions

$$a_\alpha + \dot{u}_1^+ . \nabla f_\alpha \leq 0, \quad \mu_\alpha \leq 0, \quad (a_\alpha + \dot{u}_1^+ . \nabla f_\alpha)\, \mu_\alpha = 0.$$

By joining them to the equation of dynamics

$$A\, \dot{u}_1^+ - F = \sum_\alpha \mu_\alpha \nabla f_\alpha$$

one reduces the determination of $\dot{u}_1^+$ and of the multipliers μ_α to a *Linear Complementarity Problem* in standard form.

Because the matrix A is positive definite, such a problem is classically equivalent to minimizing a convex quadratic function in a closed convex polyhedral subset of $\mathbf{R}^n$. It has been shown in [45][46] that this extremal characterization of the acceleration may be viewed as extending Gauss' *Principle of the Least Deviation* to mechanical systems subject to unilateral frictionless constraints. Some *dual* minimization property is also found to characterize the contact forces.

5.4 Coulomb friction

The presence of dry friction, governed by the law of Coulomb, at the possible contact with label α, is expressed by a relationship of the form (23). Traditionally, the law of Coulomb is only invoked for persistent contact, but when devising numerical schemes, making it meaningful for $\mathcal{U}_\alpha . \mathbf{n}^\alpha \geq 0$ and securing that a contact law of the prospective type is so stated, is just the matter of writing the code adequately.

Apart from the numerical success, the consistency of the concept of prospective type is illustrated by the following feature.

Dropping the label α for brevity, one defines the (non adhesive) Coulomb friction at some contact point by giving the *Coulomb cone* $\mathcal{C}$, a closed convex conical region of $\mathbf{E}^3$ to which the contact force $\mathcal{R}$ exerted by $\mathcal{B}'$ upon $\mathcal{B}$ should belong in any circumstance. In the standard case, $\mathcal{C}$ is rotationally symmetric about the normal vector $\mathbf{n}$ and contains it, but more general situations, accounting for anisotropic friction, are possible. The law consists in a relationship between the force $\mathcal{R}$ and the local velocity $\mathcal{U}$ of $\mathcal{B}$ relative to $\mathcal{B}'$ which resembles a plasticity law in that the values of $\mathcal{R}$ lying in the interior of $\mathcal{C}$ are compatible with $\mathcal{U} = 0$ only. But the 'flow rule' which characterizes the values of $\mathcal{U}$ compatible with a value of $\mathcal{R}$ lying on the boundary of $\mathcal{C}$ does not involve the normality of $\mathcal{U}$ to this boundary. In short, this is not an 'associated' flow rule (except in the case of zero friction coefficient).

The concept of *bipotential* has been introduced by De Saxcé and Feng [23] as a tool for handling non-associated force/velocity (or stress/strain-rate) laws, from the theoretical standpoint, as well as in numerical techniques. Here is the aspect this concept takes in the case of Coulomb friction.

As usual, one decomposes the vectors $\mathcal{R}$ and $\mathcal{U}$ into their normal and tangential components. Let $\mathbf{T}$ denote the subspace of the vectors in $\mathbf{E}^3$ which are orthogonal to $\mathbf{n}$.

$$\mathcal{R} = \mathcal{R}_{\mathrm{T}} + \mathcal{R}_{\mathrm{N}}\,\mathbf{n}, \quad \mathcal{R}_{\mathrm{T}} \in \mathbf{T}, \quad \mathcal{R}_{\mathrm{N}} \in \mathbf{R},$$

$$\mathcal{U} = \mathcal{U}_{\mathrm{T}} + \mathcal{U}_{\mathrm{N}}\,\mathbf{n}, \quad \mathcal{U}_{\mathrm{T}} \in \mathbf{T}, \quad \mathcal{U}_{\mathrm{N}} \in \mathbf{R}.$$

Classically, if the Coulomb law is restricted to situations where $\mathcal{R}_{\mathrm{N}}$ is known, say $\mathcal{R}_{\mathrm{N}} = 1$, the law turns out to reduce to a relationship between $\mathcal{R}_{\mathrm{T}}$ and $\mathcal{U}_{\mathrm{T}}$ of the associated type. Let $D_1 := \{\mathcal{R}_{\mathrm{T}} \in \mathbf{T} \mid \mathcal{R}_{\mathrm{T}} + \mathbf{n} \in \mathcal{C}\}$ (the 'unit section' of the cone $\mathcal{C}$) and define in $\mathbf{T}$ the real function (the 'dissipation function' of the said restricted law)

$$\mathcal{T} \in \mathbf{T} \mapsto \varphi_1(\mathcal{T}) := \sup\{\mathcal{S}.\mathcal{T} \mid \mathcal{S} \in -D_1\}.$$

In the traditional case of isotropic friction with coefficient γ, one simply has $\varphi_1(\mathcal{T}) = \gamma\|\mathcal{T}\|$.

Similarly to 5.1, define

$$\mathcal{K}(t,q) := \begin{cases} \{\mathcal{V} \in \mathbf{E}^3 \mid \mathcal{V}.\mathbf{n} \geq 0\} & \text{in case of contact or overlap} \\ \mathbf{E}^3 & \text{in case of no-contact.} \end{cases}$$

The Coulomb cone depends on t and q; put $\mathcal{C} = \{\mathbf{0}\}$ in case of no-contact. Using arguments from Convex Analysis, one establishes that the relation between the elements $\mathcal{U}$ and $\mathcal{R}$ of $\mathbf{E}^3$ consisting of the system of conditions

$$\mathcal{U} \in \mathcal{K},\ \mathcal{R} \in \mathcal{C},\ -\mathcal{U}.\mathcal{R} = \varphi_1(\mathcal{U}_{\mathrm{T}})\mathcal{R}_{\mathrm{N}} \tag{31}$$

is a contact law of the prospective type which, in the standard situation, reduces to the law of Coulomb.

Furthermore, one may prove

$$\forall\,\mathcal{V} \text{ in } \mathcal{K}, \quad \forall\,\mathcal{S} \text{ in } \mathcal{C}: \quad \mathcal{V}.\mathcal{S} + \varphi_1(\mathcal{V}_{\mathrm{T}})\,\mathcal{S}_{\mathrm{N}} \geq 0$$

so that (31) expresses that the real function $(\mathcal{V}, \mathcal{S}) \mapsto \mathcal{V}.\mathcal{S} + \varphi_1(\mathcal{V}_{\mathrm{T}})\,\mathcal{S}_{\mathrm{N}}$, separately convex with regard to $\mathcal{V}$ and $\mathcal{S}$, attains at point $(\mathcal{U}, \mathcal{R})$ its minimal value relative to the product set $\mathcal{K} \times \mathcal{C}$ and that *this minimal value is zero.*

6 Time-stepping

6.1 Numerical dynamics of multicontact systems

As before, denote by $[t_{\mathrm{i}}, t_{\mathrm{f}}]$, with length h, a time-step. From the approximate values $q_{\mathrm{i}}, u_{\mathrm{i}}$ obtained for q and u at t_{i} as the result of antecedent computation, one has to calculate $q_{\mathrm{f}}, u_{\mathrm{f}}$, corresponding to t_{f}.

The inertia matrix $A(t,q)$ is assumed to depend smoothly on t and q, so that one chooses to approximate it throughout the time-step by the value A_{m} it takes at the 'midpoint' $t_{\mathrm{m}} := t_{\mathrm{i}} + h/2$, $q_{\mathrm{m}} := q_{\mathrm{i}} + hu_{\mathrm{i}}/2$. Similarly, the term $F(t,q,u)$ in (21) is assumed to vary slowly enough for allowing one to replace it by the value F_{m} it takes at $t = t_{\mathrm{m}}$, $q = q_{\mathrm{m}}$, while the value u_{i} of u is used. Thereby it is understood that the force fields or pull-back actions which, among other terms, are compounded into the function F, do not depend on q in too steep a way. If such is not the case, a trick to overcome this difficulty is presented in [36]; it results in replacing A by an artificial inertia matrix in the construction of which the possible stiff elasticity coefficients are entered.

It is also at the point $(t_{\mathrm{m}}, q_{\mathrm{m}})$ that the functions f_α are calculated, in order to decide which contacts are to be treated as effective. The set $J(t,q(t))$ is thus estimated to equal $J_{\mathrm{m}} := J(t_{\mathrm{m}}, q_{\mathrm{m}})$ throughout the time-step.

The dynamical equation (21) is then discretized in the form

$$A_{\mathrm{m}}(u_{\mathrm{f}} - u_{\mathrm{i}}) = hF_{\mathrm{m}} + \sum_{\alpha \in J_{\mathrm{m}}} p^\alpha, \tag{32}$$

where the element p^α of $\mathbf{R}^n$ is made of the covariant components of the *impulsion* at contact α, i.e. the integral $\mathcal{P}^\alpha$ over $[t_{\mathrm{i}}, t_{\mathrm{f}}]$ of the contact force $\mathcal{R}^\alpha$. Throughout the time-step, one ascribes to the linear mapping G_α^* of (20) its value computed at $(t_{\mathrm{m}}, q_{\mathrm{m}})$, hence

$$p^\alpha = G_\alpha^* \, \mathcal{P}^\alpha. \tag{33}$$

One does the same with the linear mapping G_α and with the rheonomic term $\mathcal{W}_\alpha$ of (19). Since a discretization scheme of the implicit type is being planned, the *final* velocity u_{f} is invoked, so as to calculate

$$\mathcal{U}_{\alpha,\mathrm{f}} := G_\alpha \, u_{\mathrm{f}} + \mathcal{W}_\alpha. \tag{34}$$

In the exact problem, the contact law with label α should connect $\mathcal{U}_\alpha$ to the contact force $\mathcal{R}^\alpha$ at every instant. Our approximation scheme of the implicit type consists in connecting, through the same relationship, the final value $\mathcal{U}_{\alpha,\mathrm{f}}$ of $\mathcal{U}_\alpha$ to the average of this force, namely $\mathcal{P}^\alpha/h$. In usual situations such as the no-friction case described by (25) or also the case of Coulomb friction, the contact law happens to be *positively homogeneous* with degree zero relatively to the contact force, so the discretized law reads

$$\mathrm{law}_\alpha(t_{\mathrm{m}}, q_{\mathrm{m}}, \mathcal{U}_{\alpha,\mathrm{f}}, \mathcal{P}^\alpha) = \mathrm{true}. \tag{35}$$

6.2 Nonlinear Gauss-Seidel iterations

Solving the system of conditions (32) to (35) is the hard part of the computation. From (32) and (33), one obtains

$$u_{\mathrm{f}} = u_{\mathrm{r}} + \tilde{A}_{\mathrm{m}}^{-1} \sum_{\beta \in J_{\mathrm{m}}} G_{\beta}^{*} \mathcal{P}^{\beta}, \tag{36}$$

where $u_{\mathrm{r}} := u_{\mathrm{i}} + h A_{\mathrm{m}}^{-1} F_{\mathrm{m}}$ may be called the *relaxed velocity*. Here is an iteration technique à la Gauss-Seidel which amounts to treating a succession of single-contact problems.

Let an estimated solution $(u_{\mathrm{f}}^{\mathrm{esti}}, \mathcal{P}_{\mathrm{esti}}^{\beta})$, β running through J_{m}, be obtained with (36) satisfied. One expects to obtain a corrected estimate, say $(u_{\mathrm{f}}^{\mathrm{corr}}, \mathcal{P}_{\mathrm{corr}}^{\beta})$, by selecting a label α and altering only $\mathcal{P}^{\alpha}$, i.e. by making $\mathcal{P}_{\mathrm{corr}}^{\beta} = \mathcal{P}_{\mathrm{esti}}^{\beta}$ for $\beta \neq \alpha$. The new estimate is astrained to verify (36); equivalently, since the old estimate satisfies the same,

$$u_{\mathrm{f}}^{\mathrm{corr}} = u_{\mathrm{f}}^{\mathrm{esti}} + A_{\mathrm{m}}^{-1} G_{\alpha}^{*} (\mathcal{P}_{\mathrm{corr}}^{\alpha} - \mathcal{P}_{\mathrm{esti}}^{\alpha}) \tag{37}$$

and to satisfy the discretized contact law (35). By applying G_{α} to both members of (37), one gives to (35) the form

$$\mathrm{law}_{\alpha}(t_{\mathrm{m}}, q_{\mathrm{m}}, G_{\alpha}\, u_{\mathrm{f}}^{\mathrm{esti}} + \mathcal{W}_{\alpha} + H_{\alpha}(\mathcal{P}_{\mathrm{corr}}^{\alpha} - \mathcal{P}_{\mathrm{esti}}^{\alpha}), \mathcal{P}_{\mathrm{corr}}^{\alpha}) = \mathrm{true} \tag{38}$$

where $H_{\alpha} := G_{\alpha} A_{\mathrm{m}}^{-1} G_{\alpha}^{*}$ is a symmetric positive definite 3×3 matrix.

Solving (38) in the unknown $\mathcal{P}_{\mathrm{corr}}^{\alpha}$ is easy in some significant cases [52], such as two-dimensional Coulomb contact or the case where H_{α} is axissymmetric about $\mathbf{n}^{\alpha}$. Generally, some iterative procedures may be applied, in which the normal and tangential components of the contact force are alternatively treated as known. Anyway, when a solver has been devised, the above procedure of correcting successive estimates is iterated, with α ranging cyclically in J_{m}. The decision of stopping iterations may be made on observing the magnitude of the vector $\mathcal{P}_{\mathrm{corr}}^{\alpha} - \mathcal{P}_{\mathrm{esti}}^{\alpha}$; this actually is equivalent to checking the precision at which each pair $(u_{\mathrm{f}}^{\mathrm{esti}}, \mathcal{P}_{\mathrm{esti}}^{\alpha})$ satisfies the corresponding contact law.

Observe that, provided such a numerical convergence check is made, the linear operator H_{α} in (38) may be replaced by any other mapping of $\mathbf{E}^3$ into itself with zero limit at the origin, with the possible advantage of making resolution easier. This replacement is also used in tricks for accelerating convergence.

The mathematical convergence of algorithms of this sort, in the case of Coulomb contact, as well as the very existence of solutions to the problems addressed, has only been established in special situations [44][40]; uniqueness cannot be expected in general.

If all the contact laws invoked are of the prospective type, one sees that, for each α, the final relative velocity $\mathcal{U}_{\alpha \mathrm{f}}$ belongs to $\mathcal{K}_{\alpha}(t_{\mathrm{m}}, q_{\mathrm{m}})$. Similarly to 3.3, Remark 1, this ensures non-interpenetration with a welcome self-corrective

effect if the step-length is not too large. Contact laws of the prospective type also manage automatically the possible breaking of contacts, without resorting to any analysis of complementarity conditions.

To start iterations, one needs an estimate $(u_{\mathrm{f}}^{\mathrm{esti}}, \mathcal{P}_{\mathrm{esti}}^{\beta})$ verifying (36). One may take $u_{\mathrm{f}}^{\mathrm{esti}} = u_r$ and all $\mathcal{P}_{\mathrm{esti}}^{\beta}$ equal to zero. In cases where the set J_{in} of the active contacts does not change much from one step to the other, it could be much more efficient to take as first guess for the $\mathcal{P}_{\mathrm{esti}}^{\beta}$ the values calculated at the antecedent time-step for the contacts already present and to adjust $u_{\mathrm{f}}^{\mathrm{esti}}$ by means of (36).

Anyway, iterations do not require the handling of large matrices since, in the representation of each G_α for a rigid body, only a 3×6 submatrix is nonzero (2×3 in the case of a two-dimensional multibody system).

Also observe that, due to the computation being of the implicit type relative to velocity, a single computation step is needed to check whether a given position of the system is that of a possible equilibrium. One just have to launch iterations with $u_{\mathrm{i}} = 0$; if u_{f} is found zero, the corresponding set of computed values for the contact forces is compatible with equilibrium. In this sort of application, the inertia matrix A only provides the backbone of the computation and its value may be chosen so as to ease convergence.

7 Nonsmooth dynamics and collisions

7.1 Discontinuous velocity functions

The preceding Sections were developed in the context of usual 'smooth' dynamics, in which the velocity function u is assumed locally absolutely continuous on the time interval I. The properly 'unilateral' feature was only the possibility offered to contacts of breaking at any instant.

In contrast, the sudden occurrence at some instant t_c of a new contact, i.e. a *collision* is expected to generate a velocity jump. Even without such an event, the presence of dry friction at a contact point has long been known, in some cases, to forbid the existence of a smooth solution beyond some catastrophic instant, on the left of which the contact force, as well as the derivative u', may become unbounded [24]. This is a dynamic analogue to the *locking* situations familiarly met in the statics of frictional systems. Around year 1900, such a lack of smooth solution for an apparently well set problem seemed inadmissible to P. Painlevé and induced him to question the very concept of a contact force. Hence the inadequate denomination of 'Painlevé's paradox' which has been given to the observation. Today, one is accustomed to see models reaching the limits of their validity domain. Such is the case for smooth dynamics at a catastrophic instant of any sort, so that one is induced to enlarge the function space where u is looked for [52][71].

In order to develop *Nonsmooth Dynamics*, the widely accepted extension of the functional framework is that of $\mathbf{R}^n$-valued functions with *bounded*

variation. More precisely, since the time interval I has not been supposed compact, we consider the linear space of the functions with *locally* bounded variation, i.e. they have bounded variation over every compact subinterval ; notation $\mathrm{lbv}(I, \mathbf{R}^n)$. Any u in this space is locally Lebesgue-integrable, so that (2) remains meaningful. The classical property of a function with bounded variation to possess a derivative almost everywhere has no relevance in the present situation, for such a pointwise defined derivative of u (generating values for the accelerations of the system elements in physical space) does not allow one to retrieve the function through integration. In fact, the possible discontinuity points of u, which make at most a countable, hence Lebesgue-negligible, subset of I may be responsible in an essential way for the evolution of this function. For instance, a step-function possesses almost everywhere a derivative equal to zero without having to be a constant.

The role of the function u' is now played by an $\mathbf{R}^n$-valued *measure* on I, called the Stieltjes measure or *differential measure* of u, that we shall denote by du. In the smooth case, where u is locally absolutely continuous, one has $du = u'\,dt$, with dt denoting the Lebesgue measure (this is nothing but the differential measure of the real function $t \mapsto t$).

Classically, a function u in $\mathrm{lbv}(I, \mathbf{R}^n)$ possesses a left- and a right-limit at every point t of I, say $u^-(t)$ and $u^+(t)$ (with the convention $u^-(t_0) = u(t_0)$ and symmetrically at the possible other end of I) and one has for every compact subinterval $[\sigma, \tau]$ of I

$$\int_{[\sigma,\tau]} du = u^+(\tau) - u^-(\sigma).$$

In particular, the integral of du over the singleton $\{\sigma\}$ equals the possible *jump* of u at point σ.

7.2 A measure-differential equation

Since one cannot rely on the second derivative of the function $t \mapsto q(t)$ anymore, the accelerations of the elements of the system in physical space are also missing, as well as forces. The latter have to be replaced by $\mathbf{E}^3$-valued measures, the integrals of which over every compact subinterval of I constitute *impulsions.* In smooth dynamics, the impulsion measures admit, relatively to dt, some *density functions* whose values, at a given instant, are properly forces, so that impulsions, in accordance with the traditional definition, equal their time-integrals.

In particular, instead of a contact force $\mathcal{R}^\alpha$, there will more generally be invoked a *contact impulsion*, an $\mathbf{E}^3$-valued measure on the time interval I. For the homogeneity of notations, let us denote by $d\mathcal{S}^\alpha$ this measure, without having to pay attention to any function $\mathcal{S}^\alpha$ admitting it as differential measure. The *covariant components* of $d\mathcal{S}_\alpha$, relative to the parametrization

in use, are $\mathbf{R}^n$-valued measure on I, say ds_α, which, similarly to (20), are defined as

$$ds^\alpha = G_\alpha^* \, d\mathcal{S}^\alpha. \tag{39}$$

This expression makes sense if the linear mapping G_α^* depends continuously on t and q, an assumption usually verified in multibody dynamics.

The differential equation of dynamics (21) has then to be replaced by

$$A(t,q)\, du = F(t,q,u)\, dt + \sum_\alpha ds^\alpha, \tag{40}$$

an equality of $\mathbf{R}^n$-valued measures on I, called a *measure-differential equation.*

Concerning the connexion of this writing with a general formulation of classical dynamics, refer to [54]. It may be asked why (40) is given the restricted form of an equality of measures, instead of an equality of distributions of arbitrary order on the interval I. Our answer, in the present context, is that nonsmoothness comes from the contact actions which, due to unilaterality are subject to inequality requirements (even in the more general case of adhesive contact [26][68]). This is a basic fact that distributions verifying inequalities are necessarily measures.

Such an equality of measures may equivalently be exploited as an equality of functions, thanks to the following observation. Given a finite collection of $\mathbf{R}$-valued or $\mathbf{R}^n$-valued measures, such are dt, du and ds^α above, there exists (non uniquely) a nonnegative real measure on I, say $d\mu$, relatively to which these measures respectively admit density functions $t'_\mu \in \mathcal{L}^1_{\mathrm{loc}}(I, d\mu; \mathbf{R})$, $u'_\mu \in \mathcal{L}^1_{\mathrm{loc}}(I, d\mu; \mathbf{R}^n)$, $(s^\alpha)'_\mu \in \mathcal{L}^1_{\mathrm{loc}}(I, d\mu; \mathbf{R}^n)$. Then (40) is equivalent to the equality of $\mathbf{R}^n$-valued functions

$$A(t,q)\, u'_\mu = F(t,q,u)\, t'_\mu + \sum_\alpha (s^\alpha)'_\mu, \tag{41}$$

holding $d\mu$-almost everywhere in I. Replacing the *base* measure $d\mu$ by another admissible one amounts to multiply densities by positive real functions [53].

There remains to precise how contact laws may be entered into this formalism. As observed before, in the case of Coulomb friction as well as in the no-friction case, contact laws in the form (23) are conic, i.e. positively homogeneous of degree zero with regard to the argument $\mathcal{R}^\alpha$. If the density of contact impulsion $(\mathcal{S}^\alpha)'_\mu$ is introduced at the place of $\mathcal{R}^\alpha$ in the contact law, the relationship so stated between the contact impulsion measure and the local velocity function $\mathcal{U}_\alpha$ is thus indifferent to the change of the base measure. Such a policy does not seem questionable as long as the measure $d\mu$ and, consequently, the measure $d\mathcal{S}^\alpha$ are *diffuse* on the considered time-interval. In this case u is continuous, so that the local velocity $\mathcal{U}_\alpha$ is well defined for every t.

7.3 Collisions

In contrast, assume that, at some instant t_c, a velocity jump occurs as the result of a collision.

The measure du then presents an *atom* at the point t_c of I and the same should be expected from the other measures involved in the calculation. Necessarily the base measure $d\mu$ in use presents an atom at t_c and $d\mu$ may be adjusted in such a way that this atom has unit mass, i.e. the said atom is a Dirac measure located at t_c. With $d\mu$ so specified, the value $\mathcal{P}^\alpha = (\mathcal{S}^\alpha)'_\mu(t_c)$ of the density of contact impulsion for the contact labelled α, is nothing else than the vector called *percussion* in the traditional formalization of shocks. Consequently, the measure equation (40) readily applies to the standard question of determining the post-collision velocity $u^+(t_c)$, knowing the pre-collision value $u^-(t_c)$, *provided some information is available about contact impulsions.*

An information of this sort is implicitely included in the various pragmatic approaches proposed in the form of *collision laws*, since the very times of Descartes or Newton. These laws usually involve empirical parameters meant to account for the 'bounciness' of the collision (coefficients of restitution) and for the possible role of friction during the process. It is today recognized that such parameters can only be identified in narrowly delimited situations [72] such as the collision of two, otherwise free, spherical beads. The coefficients possibly determined for a pair of free bodies are not valid anymore if any of them is subject to additional constraints [32]. It should be kept in mind that if one of the bodies involved in a collision is part of a cluster of contacting objects, nonzero contact percussions should be expected at all contacts in the cluster.

Viewing a collision as a strictly instantaneous process is a schematic model, pertaining to the dynamics of strictly indeformable bodies. In contrast, if some slight deformability of the colliding bodies is evoked, a nonzero time-interval, say $[t_c, t_c + \theta]$, should be ascribed to the interaction. Large values of the contact forces are expected, the time-integral of which over this interval makes the contact percussion vector, but at this scale smooth dynamics remains applicable.

Studies of the latter sort are usually developed through a *multiple scaling* approach : a micro-time is introduced, in order to investigate the variation of the velocity function, while the variation of the position q is neglected. The direction of the sliding velocity $\mathcal{U}_\alpha$ at some of the involved contacts may be found to exhibit large variations as a function of the micro-time, preventing the identification of any representative value which could, phenomenologically, be connected with the total impulsion $\mathcal{P}^\alpha$. The situation is better if $\mathcal{U}_\alpha$ remains zero : because the Coulomb cone, closed and convex, is constant with regard to the micro-time (see however the discussion in 8.1), the condition of $\mathcal{R}^\alpha$ belonging to it commutes with the integration invoked in defining $\mathcal{P}^\alpha$.

Deeper investigation of collision processes has also been conducted, by taking into account the whole deformation of the interacting bodies, either

analytically [43] or numerically thanks to the finite element representation of deformability [76]. This also comes to confirm that collision laws can only roughly approximate the complexity of the process.

7.4 Frictionless inelastic collision

In the time-stepping scheme of Sect. 6, contact forces were involved only through their total impulsions on each time-interval, so that it remains applicable to problems formulated in the nonsmooth framework. This numerical procedure handles collisions in an automatic way. It only happens that, in the course of computation, a time-step exhibits some contacts not detected at the preceding step. Since the velocity u_{f} has to comply with some non-interpenetration conditions from which u_{i} is exempt, the contact impulsions p^{α} involved in (32) are expected to take values of larger magnitude than in the case of constant J_{m}. But calculation remains the same and delivers kinematically admissible post-collision velocities. In short, because the contact laws (35) have been assumed *positively homogeneous* with degree zero in their last arguments, the algorithm treats collisions on the same footing as permanent contacts.

There remains to precise which collision model the algorithm turns out to approximate. Assume for simplicity that a single contact has come to increment J_{m}, i.e. there occured a single impact at some unknown instant t_{c} of $[t_{\mathrm{i}}, t_{\mathrm{f}}]$. The endpoint values u_{i} and u_{f} are viewed as approximating $u^{-}(t_{\mathrm{c}})$ and $u^{+}(t_{\mathrm{c}})$ respectively.

We first have a look at the calculation to be made in the special frictionless setting of 5.2. The polyhedral set W defined in (29), namely the set of the values of u^{+} compatible with all non-interpenetration conditions, including the new one, is approximated by

$$W_{\mathrm{m}} := \{v \in \mathbf{R}^{n} \,|\, \forall \alpha \in J_{\mathrm{m}} \,:\, \frac{\partial f_{\alpha}}{\partial t}(t_{\mathrm{m}}, q_{\mathrm{m}}) + v.\nabla f_{\alpha}(t_{\mathrm{m}}, q_{\mathrm{m}}) \leq 0\}. \qquad (42)$$

As in 6.2, one constructs the relaxed velocity $u_{\mathrm{r}} := u_{\mathrm{i}} + hA_{\mathrm{m}}^{-1}F_{\mathrm{m}}$. If the positive definite matrix A_{m} is used to define in $\mathbf{R}^{n}$ a Euclidean metric, the core of the computation simply consists in constructing u_{f} as *the nearest point to* u_{r} *in* W_{m}. By imagining the step-length h arbitrarily close to zero, one sees that the calculation approximates the following *collision law* :

The post-collision velocity $u^{+}(t_{\mathrm{c}})$ *equals, in the sense of the kinetic metric* $A(t_{\mathrm{c}}, q(t_{\mathrm{c}}))$, *the nearest point to* $u^{-}(t_{\mathrm{c}})$ *in* $W(t_{\mathrm{c}}, q(t_{\mathrm{c}}))$.

Let us refer now to the contact locus, in the notations of 5.1. The above calculation amounts to admit that, for each α, the contact impulsion $\mathcal{P}^{\alpha}$ is connected, through the law of frictionless contact (25), with the *post-collision* velocity $\mathcal{U}_{\alpha}^{+}$. Due to this law being of prospective type, the implication $\mathcal{P}^{\alpha} \neq \mathbf{0} \Rightarrow \mathcal{U}_{\alpha}^{+}.\mathbf{n}^{\alpha} = 0$ holds, meaning that, if the contact labelled α takes an effective part in the nonsmooth process, it exhibits the feature traditionally

formulated by saying that *the Newton restitution coefficient is zero.* But, depending on the circumstances met in constructing the projection, one may also have $\mathcal{P}^\alpha = \mathbf{0}$, with $\mathcal{U}_\alpha^+ . \mathbf{n}^\alpha$ only astrained to be nonnegative.

The present model (it was introduced in [51] under the name of Standard Inelastic Shock) is somewhat more realistic than the traditional Newton statement, in that *all the contacts present at instant t_c are treated collectively.* This produces plausible results for instance in the case, popular in the Earthquake Engineering literature, of a rectangular block *rocking* on a horizontal plane: if the traditional formulation, with zero restitution coefficients, was applied to all contacts, no rocking could be found. In contrast, the above collision law allows one to discuss the outcome with regard to the aspect ratio of the block. A slender block presents a succession of oscillations, progressively damped by the inelasticity of collisions, while, below some critical aspect ratio, a single episode of motion leads to permanent contact.

7.5 A three-parameter collision law

In the same line, there remains to show how non-zero friction and non-zero restitution can be entered into a computationally efficient collision model. At an instant where the system experiences a velocity jump we propose, for every contact α, to relate through the Coulomb law of friction (stated in prospective form) with friction coefficient γ_α, the contact percussion $\mathcal{P}^\alpha$ with some *average value* $\mathcal{U}_\alpha^a$ of the local velocity. This artificial value is defined as a weighted mean of the (known) pre-collision velocity $\mathcal{U}_\alpha^-$ and of the (unknown) post-collision velocities $\mathcal{U}_\alpha^+$. Introducing as before the normal and tangential components of the concerned vectors, put

$$\mathcal{U}_{\alpha\mathrm{N}}^a = \frac{\rho_\alpha}{1+\rho_\alpha}\mathcal{U}_{\alpha\mathrm{N}}^- + \frac{1}{1+\rho_\alpha}\mathcal{U}_{\alpha\mathrm{N}}^+ \tag{43}$$

$$\mathcal{U}_{\alpha\mathrm{T}}^a = \frac{\tau_\alpha}{1+\tau_\alpha}\mathcal{U}_{\alpha\mathrm{T}}^- + \frac{1}{1+\tau_\alpha}\mathcal{U}_{\alpha\mathrm{T}}^+. \tag{44}$$

The empirical parameters ρ_α and τ_α will be called *the normal coefficient of restitution* and *the tangential coefficient of restitution* at the contact labelled α, denominations justified by what follows.

Since the invoked contact law is of the prospective type, implications (26) and (27) show that $\mathcal{P}^\alpha$ can be nonzero only if $\mathcal{U}_{\alpha\mathrm{N}}^a = 0$, i.e. $\mathcal{U}_{\alpha\mathrm{N}}^+ = -\rho_\alpha \mathcal{U}_{\alpha\mathrm{N}}^-$, which formally is Newton's restitution law. But the present formulation is richer than stipulating a normal restitution law separately for each contact ; it also allows $\mathcal{P}^\alpha = 0$, in which case only the inequality $\mathcal{U}_{\alpha\mathrm{N}}^a \geq 0$ happens to be asserted. It is the global calculation, involving all the contacts together through the equation of dynamics, which decides between these two alternatives.

Similarly, the global calculation, if friction is large enough, may end in the zero sliding case of Coulomb's law at contact α. Then $\mathcal{U}_{\alpha \mathrm{T}}^{+} = -\tau_{\alpha} \mathcal{U}_{\alpha \mathrm{T}}^{-}$, which is a *law of tangential restitution.*

The construction of the average velocity may readily be incorporated into the time-stepping computation scheme of Sect.6, so that the algorithm is ready to face possible collisions at every instant. At the price of a few extra lines of code at each time-step before launching the Gauss-Seidel iterations, no computation cost at all is added to the iterations themselves

Of course the same collision model may be used in conjonction with other solvers [38][39].

The time-stepping scheme also yields plausible results when some of the frictional catastrophes referred to in 7.1 are met [52].

It should however be admitted that relating the percussions $\mathcal{P}^{\alpha}$ to the respective average velocities $\mathcal{U}_{\alpha}^{a}$ is nothing but a pragmatic trick, generating a collision law of the sort which precisely was criticized in 7.3. Before relying on it in a specific context, one has to check its practical value against any available experimental measurement.

In the simplest case of all, that of the collision of two otherwise free spherical beads, the resulting law turns out to be mathematically equivalent to another three-parameter collision model anteriorly proposed on the basis of quite different arguments [76] and which, experimentally, has been found in fairly good agreement with reality [25].

The various circumstances met when a ball bounces against a fixed plane are also convincingly reproduced, in particular the exotic behaviour of a 'superball' [55][56].

The rocking of a slender block supported by an oscillating table has provided another occasion of testing the model. Some sensitive features of the rocking regimes have been correctly predicted by CD computation [66].

It is in the domain of the numerical simulation of granular materials that the computational simplicity of the procedure proves invaluable. In fact, one may then have several ten thousands of contacts to investigate at each time-step.

Some satisfactory comparisons of computation with experiments are presented in [55][56], pertaining in particular to the circulatory currents observed in shaken granular samples and their connection with the segregation of grain sizes [41].

The study of the flow of grains over a rough incline has also benefited from joining experimental investigation with CD numerical simulations [14].

The efficiency of the method in the dynamics of granular materials mainly stems from that, once a time-discretization has been chosen, all the collisions which have been detected as occurring on a given time-step are treated together. This of course entails some trade-offs which have to be assessed. It does not seem to matter that the ordering of collisions, which mechanically should be successive, is only internal to the algorithm. In fact the problems

in view are physically *undeterministic*, since a slight change in the initial conditions is liable to generate after a short time a completely different sequence of positions and collisions. A more critical issue is that each contact point is only counted once in the analysis. An accumulation of bounces, such as those of a ping-pong ball coming to rest, is thus viewed as a single collision as soon as the successive bounces are all comprised in the time-step. This could result in underestimating the total energy loss arising from inelastic collisions. This source of error may be checked by repeating computation with step-length reduced to half : if no appreciable change is found in the energy versus time curve, one may conclude that no harm was done.

It has been stressed in the foregoing that, in compact assemblies, the bodies involved in a collision may be part of clusters of already contacting objects, between which percussions are likely to be induced. The described procedure of global dynamical calculation does take this into account. Physically however, the transmission of impulses in clusters should involve elasticities in a way analogous to sound propagation, an effect which cannot be analyzed through the present model of strictly rigid objects.

Anticipating on the question of energy balance, to be studied in the next Section, one may finally rise another objection. Like other popular collision laws, the trick of average velocities offers no security against the possible violation of the laws of Thermodynamics. In fact, for colliding objects of not too usual shapes, the calculated outcome may entail some energy creation. This is a physically inacceptable conclusion, except of course for artificially boosted collisions as in electric billiard games. A systematic thermodynamic formalism has been proposed by M. Frémond [27], securing that the empirical collision laws one may insert in it are dissipative [15][16].

8 Energy balance and calculus for differential measures

8.1 Departing from smooth dynamics

Drawing the energy balance of a dynamical evolution is essentially relevant to the case where the external obstacles or boundaries are fixed in the (Galilean) reference frame in use. In the smooth dynamical context of Sect.4 it was alluded to the possibility of constructing the parametrization (q) with account of linkages or permanent hinges through which members of the system would be connected with some external supports. Also these external supports have to be assumed fixed, in order to obtain a 'scleronomic', i.e. time-independent parametrization. This makes the inertia matrix $A(t,q)$ constant with regard to its first argument and gives to the kinetic energy of the system an expression which, for every q, is a positive definite quadratic form in u, namely $\mathcal{E}_{\mathrm{k}} = A_{ij}u^i u^j/2$. In this framework of smooth dynamics, one classically establishes that the time-derivative of $\mathcal{E}_{\mathrm{k}}$ equals the total power of the forces experienced by the system, including the contact forces between members and the possible action of external supports. If contacts are frictionless as

well as the (fixed) possible external hinges the corresponding terms in the power vanish. In the familiar case where, additionally, the other forces in presence derive from a potential energy, the integration of derivatives yields the well known property of *energy conservation.*

In contrast, the elementary example of a particle hitting 'inelastically' a fixed frictionless plane shows that, in nonsmooth situations, *the absence of friction does not secure energy conservation anymore.* Mathematically, the reason is that the formula for the time-derivative of a product of (locally absolutely continuous) functions, used in establishing the energy balance in the smooth case, has to be replaced by less precise relationships to be presented in 8.2 below. The physical background needs some comments.

The discussion sketched in 7.3, in which a nonzero time-interval $[t_c, t_c+\theta]$ was, at least qualitatively, assigned to the process, makes one understand that the small deformations that the contacting bodies undergo during the collision may generate values of the local velocity vector $\mathcal{U}_\alpha$ disagreeing with condition $\mathcal{U}_\alpha.\mathbf{n}^\alpha = 0$, while the vector $\mathbf{n}^\alpha$ itself may exhibit transient variations. The instant power of contact forces can then be nonzero, yielding for their total work over the time-interval a nonzero, commonly negative, value. The same remark applies to the external linkages asserted to connect the system with fixed supports : even frictionless, they may transmit work to the outside world. This induces one to question, when nonsmooth circumstances are met, the legitimity of involving these linkages in the construction of the parametrization : the fact that such linkages are perfect in the sense of the traditional smooth dynamics does not secure that the percussions they impart in a collisional situation should have zero generalized components.

8.2 Nonsmooth differential calculus

Some rules of a calculus for l.b.v. vector functions of a real variable t are established in [53]. In particular, one finds expressions for the differential measures of functions constructed from elements of l.b.v. through multilinear operations. For instance one has

$$d(u^i u^j) = u^{i+} du^j + u^{j-} du^i \tag{45}$$

and various similar formulas where left- and right-velocities are differently combined. If the scalar product in $\mathbf{R}^n$ is denoted by a dot, this yields

$$d\|u\|^2 = (u^+ + u^-).\, du. \tag{46}$$

Furthermore, thanks to the quadratic form $u \mapsto u.u = \|u\|^2$ being nonnegative, one establishes the following inequality, in the sense of the ordering of real measures on I

$$2u^-.du \leq d\|u\|^2 \leq 2u^+.du. \tag{47}$$

This holds true, more generally, in any Euclidean or real Hilbert linear space.

By integrating the concerned measures on the singleton $\{t_c\}$, one may derive from these rules of extended differential calculus a formula for the jump of kinetic energy at the instant of a possible collision

$$\mathcal{E}_k^+ - \mathcal{E}_k^- = \frac{1}{2}(u^+ + u^-).A(u^+ - u^-). \tag{48}$$

This result, sometimes called Kelvin's theorem, actually is of purely algebraic nature and one may recognize in it the elementary identity $(a+b).(a-b) = \|a\|^2 - \|b\|^2$ applied to the Euclidean metric associated with the inertia matrix $A = A(q(t_c))$. If (40) is used here to express $A(u^+ - u^-)$ (recall that the measure dt has zero integral on a singleton), this yields

$$\mathcal{E}_k^+ - \mathcal{E}_k^- = \frac{1}{2}(u^+ + u^-).A(u^+ - u^-) = \frac{1}{2}(u^+ + u^-).\sum_\alpha p^\alpha. \tag{49}$$

where $p^\alpha = G_\alpha^* \mathcal{P}^\alpha$ is the n-vector made of the generalized component of the percussion $\mathcal{P}^\alpha$ occurring at the contact labelled α. The right-hand member equivalently writes down as

$$\begin{aligned}\frac{1}{2}\sum_\alpha (u^+ + u^-).G_\alpha^* \mathcal{P}^\alpha &= \sum_\alpha \frac{1}{2}(G_\alpha u^+ + G_\alpha u^-).\mathcal{P}^\alpha \\ &= \sum_\alpha \frac{1}{2}(\mathcal{U}_\alpha^- + \mathcal{U}_\alpha^+).\mathcal{P}^\alpha. \end{aligned} \tag{50}$$

Therefrom the jump of $\mathcal{E}_k$ is expressed as a sum of terms corresponding to the respective contacts effective in physical space. Mathematically, it may be convenient to view each of these terms as defining the *energy* or *work* of the corresponding percussion. But one should keep in mind that nothing in this analysis can justify the conception that such terms could, each for its part, express a separate energy balance for some physical phenomenon occurring at the corresponding contact locus.

As an illustration, one may come back to the frictionless collisions presented in 7.4. Since each percussion $\mathcal{P}^\alpha$ was then assumed connected with the corresponding $\mathcal{U}_\alpha^+$ through the law of frictionless contact (25), one has $\mathcal{U}_\alpha^+.\mathcal{P}^\alpha = 0$. From the right-side inequality in (47), one concludes that such collisions essentially involve some energy loss. In contrast, one might stipulate that the same law of frictionless contact should hold between $\mathcal{P}^\alpha$ and the arithmetic mean $(\mathcal{U}_\alpha^- + \mathcal{U}_\alpha^+)/2$. Then (50) would entail energy preservation. Note that if some of the contacts involved take place between a member of the system and some external obstacle, the latter must be fixed.

When invoked in the space $\mathbf{R}^n$ of the parameters, the norm and scalar product should be understood in the sense of the kinetic metric. More generally than in the above investigation of an instant phenomenon, the energy

balance of the possibly nonsmooth motion of the system over a time interval may be stated as an equality of measures, in which the evolution of the matrix A is taken into account [52][53].

9 Behaviour of a masonry structure

The dynamics of block assemblies has been submitted to CD computation, in order to simulate their response to seismic actions [34]. In this domain of application, it may be necessary to take also into account a slight deformability of the blocks and the presence of mortar joints [2].

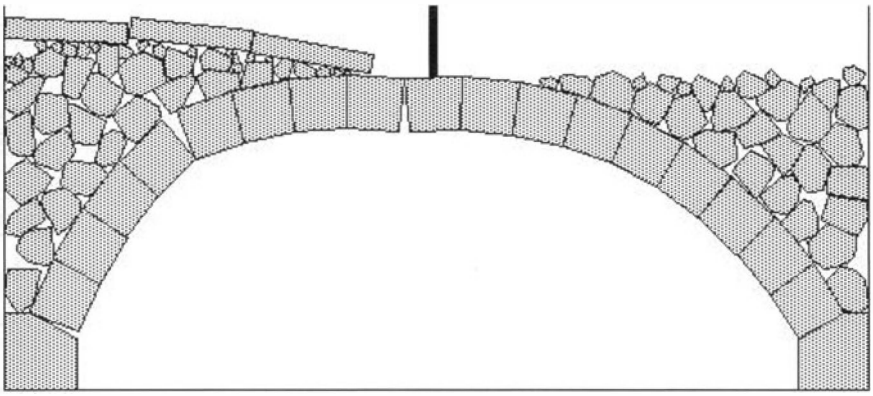

Fig. 1. Friction coefficient 0.6

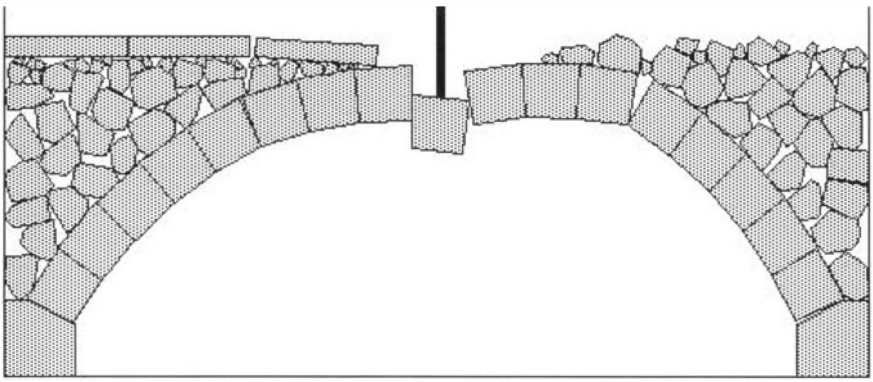

Fig. 2. Friction coefficient 0.3

Numerical simulation is applied here to a two-dimensional model of stone bridge in order to display its collapse under some localized forcing. The pur-

pose is pedagogical, with a view to make clear that the ruin is not the result of the crushing strength of stone beeing exceded, but has to be discussed in terms of geometry and of the unilaterality of non-interpenetrability constraints [29][30]. If mortar joints are present between stones, their strength is neglected and the contact described as obeying the Coulomb law of friction. Coulomb himself, when investigating this issue, observed that, in practical instances, the friction coefficient was large enough for no sliding to occur between archstones so that the evolution toward collapse consisted of the opening of some joints through the relative rotation of the corresponding archstones about *hinge points*.

The structure, initially in equilibrium under its own weight, is submitted to the action of an external object with imposed motion, forcing the central archstone down. Computation of the subsequent motion is performed twice, with the respective values of 0.6 and 0.3 for the friction coefficient at all contacts.

10 Stress distribution in a conical grain pile

10.1 Constructing a pile

Numerical simulation is a popular mode of investigation in granular mechanics. In particular, dry cohesionless granular materials are naturally modelled as collections of bodies which interact by contacts affected with Coulomb friction. Provided the model is sufficiently validated by comparing some measurements, feasible in laboratory on physical experiments, with the corresponding numerical results, computer simulation may bring precious informations about intimate mechanisms and quantities otherwise inaccessible.

The following observation has, in recent years, started a lot of speculation and controversy [11][70]. If a conical pile is created by pouring grains from a point source onto a rough rigid horizontal ground, some experimentalists have been surprised to find that the distribution of pressure at various ground contact points was not proportional to the height of the material above and even that some local minimum of ground pressure could be present at the vertical of the apex.

The distribution of stress in a granular material not only depends on the forces actually exerted on it, but strongly also on the way the granular mass has been prepared. Numerical simulation thus has to reproduce the preparation process.

Figure 3 displays a view of the following numerical experiment [61], performed with the Contact Dynamics algorithm of Sect. 6. A pile of about 14000 spherical grains is constructed by depositing grains, one by one, at the contact of already existing ones. Grains diameters are distributed at random, uniformly from 0.25 cm to 0.5 cm. The fixed horizontal ground roughness is simulated by a random pavement of grains with the same distribution of

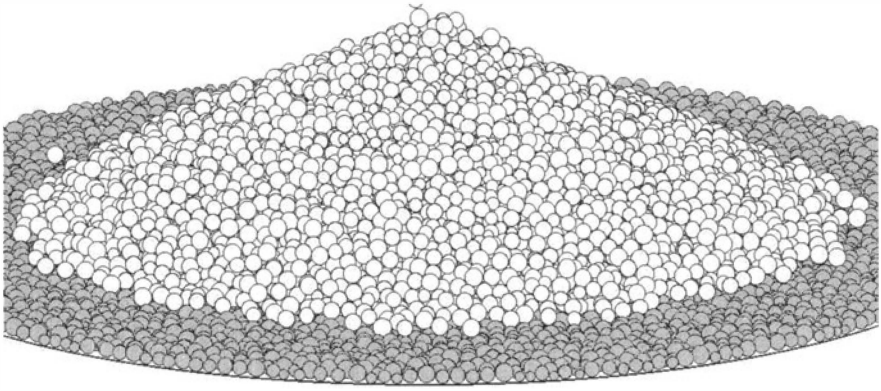

Fig. 3. View of the pile

sizes. Friction coefficient : 0.4 everywhere. Restitution coefficients : 0. Gravity : 981 cm/s^2. The vertical of a grain center at the time of its deposition is chosen at random in the neighbourhood of the Oz axis at a maximal distance of 0.8 cm. Deposition frequency : 200 per second.

A large part of the deposited grains run down the pile slope before stopping, sometimes triggering avalanches. The velocity of running grains may reach 24 cm/s. For sufficient precision in calculating these motions, including the avalanches, the step length is fixed at 2×10^{-4} s. The duration of the simulated phenomenon equals 72 s, so that 360 000 steps have been needed. In the final stages the number of contact points is about 33 000. The Gauss-Seidel procedure remains practical on a microcomputer (but a matter of weeks) for a system of such a size because the largest part of the pile stays in quasi-equilibrium : the contact impulsions calculated in the antecedent time-step for the contacts already effective are used as first guess in iterations.

Figure 4 shows, in four stages of the pile building, a slice 2 cm thick containing Oz. The grains deposited between $t = 15.5$ and $t = 18.7$ and which, at the considered stage, have their centers in this slice are represented in black. One observes that a noticable proportion of the deposited grains do not run out on the slopes, but accumulate in the central region so as to cause some plastic deformation of the existing granular mass. This deformation may be assessed by visualizing the distribution of these black grains at the successive stages. Let us call a *fossile layer* such a collection of grains observed at a certain date and deposited during a specified anterior period.

10.2 Stress distribution

The programme allows one to choose a line segment in a plane drawn through Oz. By rotation about Oz, this segment generates a curved strip in the shape of a truncated cone or of a cylinder, across which force transmission will be analyzed. To this end, the contacts occurring in grain pairs with centers

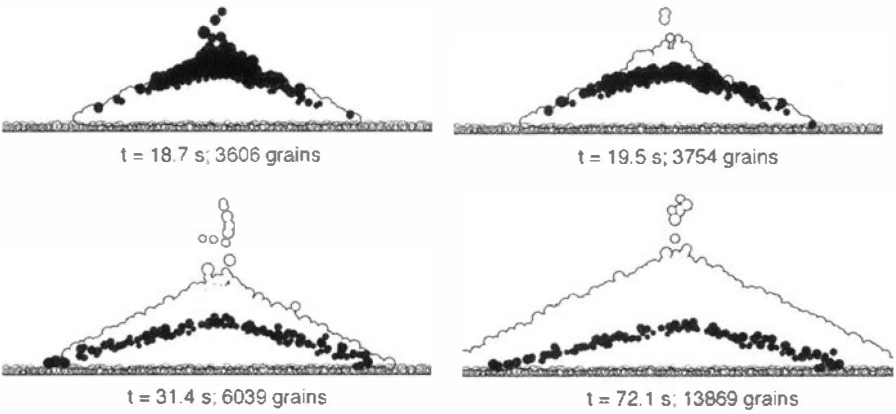

Fig. 4. Fossilelayers

separated by the conical or cylindrical surfaces are reviewed. A contact force $\mathcal{R}$ is retained in the list if the corresponding contact point M has a projection onto the separating surface which belongs exactly to the strip. The meridian half-plane Π containing M is used, in order to decompose $\mathcal{R}$ into a component $\mathcal{R}_z$ in the Oz direction and a component $\mathcal{R}_r$ in the axifugal direction (the component of $\mathcal{R}$ orthogonal to Π is not investigated). The sum of all $\mathcal{R}_z$ in the list, as well as the sum of all $\mathcal{R}_r$, are divided by the area of the strip. This yields the components in Π of a vector $\mathcal{T}_\Pi$ which may be viewed as the *average density of meridian force* transmitted across the strip.

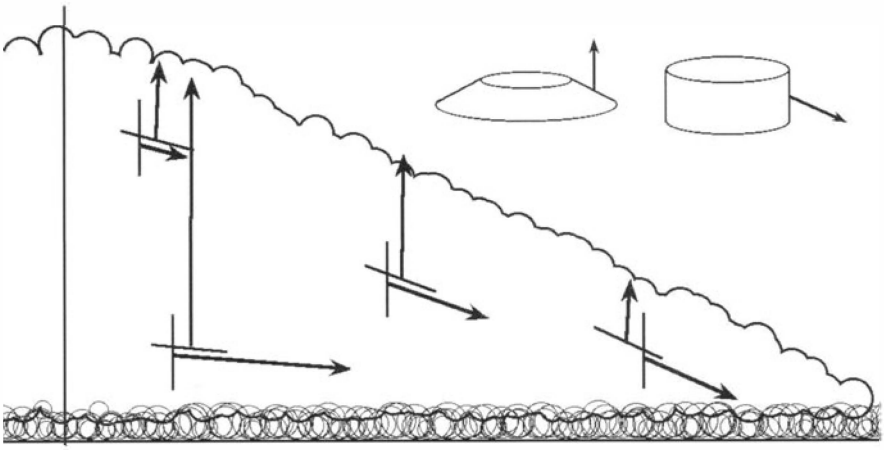

Fig. 5. Reciprocal cuts

In a classical continuous material, possessing a Cauchy stress tensor field, the similar construction would deliver an average density of meridian force

related to the strip normal unit ν in the form $\mathcal{T}_\Pi = \sigma_\Pi \nu$, where σ_Π denotes a *symmetric* two-dimensional tensor. Symmetry implies a *reciprocity* property for a pair of (conical or cylindrical) strips drawn through a given point, with respective normal units ν and ν' and force densities $\mathcal{T}_\Pi$ and $\mathcal{T}'_\Pi$: one readily finds $\nu.\mathcal{T}'_\Pi = \nu'.\mathcal{T}_\Pi$. For a common value equal to zero, there comes out that the strip with normal ν is parallel to $\mathcal{T}'_\Pi$ if and only if the strip with normal ν' is parallel to $\mathcal{T}_\Pi$.

That a granular material, observed at large scale, admits a Cauchy stress field like any continuous medium is a familiar fact, for instance in Soil Mechanics. When some numerical simulation has revealed the values of intergranular forces, the assessement of the average stress over a delimited sample is also a classical matter but the question of the representativity of the sample, in regard to its size, needs investigation. In fact, photoelastic experiments as well as numerical simulations reveal that the transmission of forces in a granular mass in equilibrium exhibits a certain inhomogeneity. Loads happen to be carried by privileged grains forming *force chains*. A sample has naturally to exceed the size of these chains in order to yield a representative value of stress. For an investigation of this subject through Contact Dynamics simulations, see [57][60].

The purpose of Fig. 5 is to check that, in the present numerical experiment, the numbers of grains is large enough for the above reciprocity property to hold reasonably true, in spite of sampling aleas (a two-dimensional example, with irregular polygonal grains is presented in [60]).

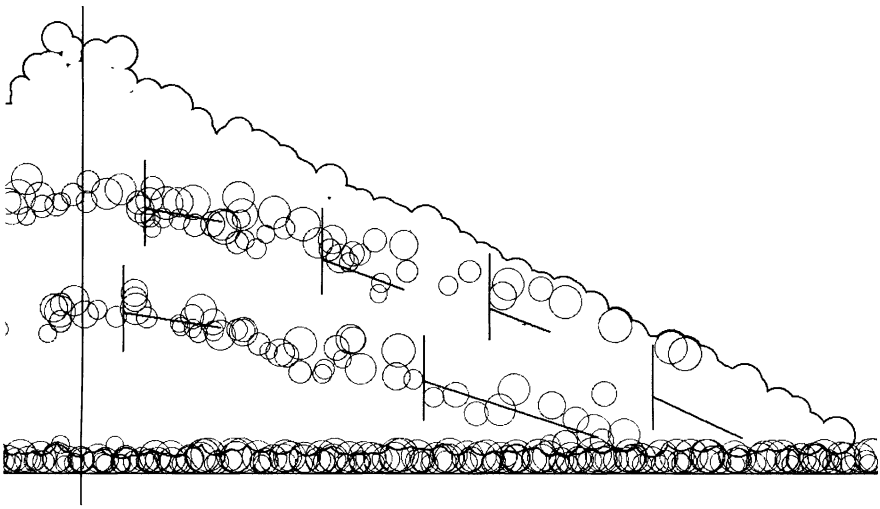

Fig. 6. Reference to fossile layers

Figure 6 shows the same 2 cm thick slice as before, with two fossile layers outlined. This suggests the following approximate law :

In each point of the pile, the average density of meridian force $\mathcal{T}_\Pi$ corresponding to a cylindrical strip with axis Oz has the same direction as the meridian section of a fossile layer containing this point.

This distribution of meridian forces expresses an *arching* effect, present even in the central region of the pile.

In the peripheral bank, which has been constructed by the deposition of avalanching grains, the fossile layers are nearly parallel to the free surface. Through the reciprocity property, this is equivalent to saying that, on a conical cut parallel to the free surface, the average density of transmitted force is vertical. This feature looks natural since the bank is made of layers successively deposited by avalanches. When such a layer stops, the supporting forces it experiences from the material beneath should equilibrate its weight, and this verticality of transmitted forces is expected to persist after other layers possessing the same property have been superimposed.

10.3 Ground pressure

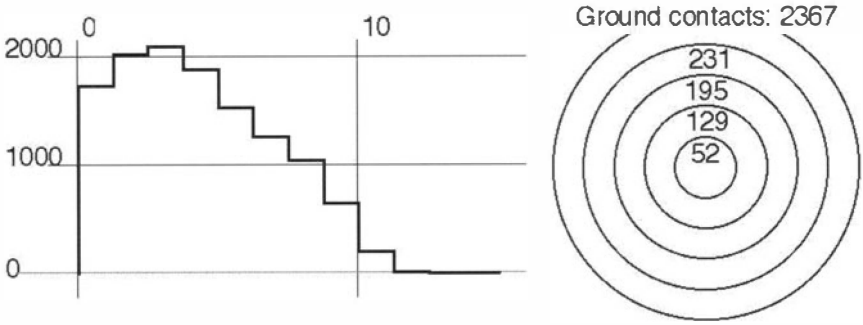

Fig. 7. Average ground pressure in annular regions

The central core of the pile clearly has part of its weight supported by the arching effect of the peripheral bank and this explains that in some experiments, a local minimum of ground pressure has been found at the vertical of the pile apex. Figure 7 displays the results obtained in the present numerical simulation. The supporting ground surface has been divided into annular regions with axis Oz and uniform width. For all the contact points between the pile and the rough gound surface wich project in one of the annular regions, the total vertical component of the contact forces is divided by the area of the region, yielding the corresponding average pressure. The graph shows the respective values of this pressure (in deciPascal) plotted against the distance to Oz. Regions of small radii are naturally more subject to sampling aleas

than the larger ones, but a central local minimum is in evidence. The pressure at center is smaller than the hydrostatic pressure which would be found in a liquid column of the same height and the same average volume mass, namely 3320 dPa corresponding to the height of 5.5 cm and a volume mass of 0.615 (in computation, the volume mass of the material the grains are made of has been taken equal to 1).

Anyway, the statics of a pile of dry grains has nothing to do with hydrostatics. This is evidenced by the fact that the free surface of the pile at rest is not horizontal. The merit of the reciprocity property disclosed in the foregoing is to establish a mathematical connexion between the arching effect and the inclination of the free surface.

References

1. Abadie M. (2000) Dynamic simulation of rigid bodies : modelling of frictional contact. In: Brogliato B. (Ed.) Impacts in Mechanical Systems. Analysis and Modelling. Springer, Berlin Heidelberg, 61–144
2. Acary V. (2001) Contribution à la modélisation mécanique et numérique des édifices maçonnés, Thesis, Université Aix-Marseille II
3. Alart P, Curnier A. (1991) A mixed formulation for frictional contact problems prone to Newton like methods. Comput. Meth. Appl. Mech. Engrg. **92**, 353-375
4. Aubin J. P. (1990) Viability Theory. Birkhäuser-Verlag, Basel Boston
5. Aubin J. P., Cellina A. (1984) Differential inclusions. Springer-Verlag, Berlin
6. Ballard P. (2001) Formulation and well-posedness of the dynamics of rigid body systems with unilateral or frictional constraints. Phil. Trans. Roy. Soc., Ser. A, to appear
7. Baraff D. (1993) Issues in computing contact forces for non-penetrating rigid bodies. Algorithmica **10**, 292–352
8. Brézis H. (1973) Opérateurs maximaux monotones et semi-groupes de contraction dans les espaces de Hilbert. North-Holland, Amsterdam
9. Brogliato B. (1999) Nonsmooth Mechanics, 2d. edition. Springer, London
10. Brogliato B., ten Dam A. A., Paoli L., Génot F., Abadie M. (2001) Numerical simulation of finite dimensional multibody nonsmooth mechanical systems. ASME Applied Mechanics Reviews, to appear.
11. Cates M. E., Wittmer J. P., Bouchaud J.-P., Claudin P. (1998) Development of stresses in cohesionless poured sand. Phil. Trans. Roy. Soc. London, Ser. A, **356**, 2535–2560
12. Chabrand P., Dubois F., Raous M. (1998) Various numerical methods for solving unilateral contact problems with friction. Mathl. Comput. Modelling **28**, 97–108
13. Chatterjee A., Ruina A. (1998) Two interpretations of rigidity in rigid-body collisions. ASME J. Appl. Mech. **65**, 894–900
14. Chevoir F., Prochnow M., Jenkins J., Mills P. (2001) Dense granular flows down an inclined plane. Powders and Grains 2001 (Sendai, Japan). Balkema, Rotterdam Brookfield, to appear
15. Cholet C. (1998) Chocs de solides rigides. Thesis, Univ. Paris 6

16. Cholet C., Dimnet E., Chevoir F., Moucheront P., Frémond M. (1997) Experimental study of collisions of angular particles. In: Behringer R., Jenkins J. (Eds.) Powders and grains 97. Balkema, Rotterdam Brookfield, 543–546
17. Cottle R. W., Pang J.-S., Stone R. E. (1992) The Linear Complementarity Problem. Academic Press, Boston, etc.
18. Christensen P. W., Klarbring A., Pang J. S., Strömberg N. (1998) Formulation and comparison of algorithms for frictional contact problems. Int. J. Num. Meth. Engng. **42**, 145–173
19. Cundall P. A. (1971) A computer model for simulating progressive large scale movements of blocky rock systems, Proceedings of the Symposium of the International Society of Rock Mechanics, Vol.1, Nancy, France, 132–150
20. Curnier A. (1984) A theory of friction. Int. J. Solids Struct. **20**, 637–647
21. Deimling K. (1992) Multivalued Differential Equations. De Gruyter, Berlin
22. Delassus E. (1917) Mémoire sur la théorie des liaisons finies unilatérales. Ann. Sci. Ecole Norm. Sup. **34**, 95–179
23. De Saxcé G., Feng Z. Q. (1991) New inequation and functional for contact with friction. J. Mech. of Struct. and Machines **19**, 301–325
24. Génot F., Brogliato B. (1999) New results on Painlevé paradoxes. European Journal of Mechanics, A/Solids, **18**, 653–677
25. Foerster S., Louge M., Chang H., Allia K. (1994) Measurements of the collision properties of small spheres. Phys. Fluids **6**, 1108-1115.
26. Frémond M. (1987) Adhérence des solides. J. Mécanique Théorique et Appliquée **6**, 383–407
27. Frémond M. (1995) Rigid body collisions. Physics Letters A **204**, 33–41
28. Haug, E. J. (1989) Computer Aided Kinematics and Dynamics, vol. 1. Allyn and Bacon, Boston
29. Heyman J. (1966) The stone skeleton. Int. J. Solids Structures **2**, 249–279
30. Heyman J. (1995) The stone skeleton, structural engineering of masonry architecture. Cambridge University Press, Cambridge U. K.
31. Isac G. (1992) Complementarity Problems. Lecture Notes in Math., Vol. **1528**. Springer-Verlag, Berlin
32. Ivanov A. P. (1999) On the problem of constrained collision. In: Pfeiffer F., Glocker Ch. (Eds.) Unilateral Multibody Contacts. Kluwer, Dordrecht Boston London, 107–116
33. Jean M., Moreau J. J. (1987) Dynamics in the presence of unilateral contact and dry friction. In: Del Piero G., Maceri F. (Eds.) Unilateral Problems in Structural Analysis 2, CISM Courses and Lectures Vol. **304**. Springer-Verlag, Wien, New York, 151–196
34. Jean M., Moreau J. J. (1992) Unilaterality and dry friction in the dynamics of rigid body collections. In: Curnier A. (Ed.) Contact Mechanics International Symposium. Presses Polytechn. et Universit. Romandes, Lausanne, 31–48.
35. Jean M., Touzot G. (1988) Implementation of unilateral contact and dry friction in computer codes dealing with large deformations problems. J. Mécanique Théorique et Appliquée **7** (suppl. Nr. 1), 145–160
36. Jean M. (1995) Frictional contact in collections of rigid or deformable bodies: numerical simulation of geomaterials. In: Selvadurai A.P.S., Boulon M.J. (Eds.) Mechanics of Geomaterial Interfaces. Elsevier, Amsterdam, 453–486
37. Jean M. (1999) The Non Smooth Contact Dynamics method. In: Martins J. A. C., Klarbring A. (Eds.) Computational Modeling of Contact and Friction, special issue of Computer Meth. in Appl. Mech. and Engng. **177**, 235–257

38. Johansson L. (1999) A linear complementarity algorithm for rigid body impact with friction. European Journal of Mechanics, A/Solids, **18**, 703–717
39. Johansson L., Klarbring A. (2000) Study of frictional impact using a nonsmooth equation solver. ASME J. Appl. Mech. **67**, 267–273
40. Jourdan, F., Alart, P., Jean, M. (1998) A Gauss-Seidel-like algorithm to solve frictional contact problems. Computer Meth. Appl. Mech. Engng. **155**, 31–47
41. Knight J. B., Jaeger H. M., Nagel S. R. (1993) Vibration-induced size separation in granular media : the convection connection. Phys. Rev. Lett. **70**, 3728–3731
42. Kunze M., Monteiro Marques M. D. P. (2000) An introduction to Moreau's sweeping process. In: Brogliato B. (Ed.) Impacts in Mechanical Systems. Analysis and Modelling. Springer, Berlin Heidelberg, 1–60
43. Maw N., Barber J. R., Fawcett J. N. (1977) The rebound of elastic bodies in oblique impact. Mech. Res. Comm. **4**, 17–21
44. Monteiro Marques M. D. P. (1993) Differential Inclusions in Nonsmooth Mechanical Problems: Shocks and Dry Friction. Birkhäuser, Basel, Boston, Berlin
45. Moreau J. J. (1963) Les liaisons unilatérales et le principe de Gauss. Comptes Rendus Acad. Sci. Paris **256**, 871–874
46. Moreau J. J. (1966) Quadratic programming in mechanics: dynamics of one-sided constraints. SIAM J. Control **4**, 153–158.
47. Moreau J. J. (1973) Problème d'évolution associé à un convexe mobile d'un espace hilbertien. Comptes Rendus Acad. Sci. Paris **276**, 791–794
48. Moreau J. J. (1974) On unilateral constraints, friction and plasticity. In Capriz G., Stampacchia G. (Eds.) New variational techniques in mathematical physics, C.I.M.E. II ciclo 1973. Edizioni Cremonese, Roma,173–322.
49. Moreau J. J. (1976) Application of convex analysis to the treatment of elastoplastic systems. In: Germain P., Nayroles B. (Eds.) Applications of functional analysis to problems of mechanics, Lecture Notes in Mathematics **503**, Springer-Verlag, Berlin, Heidelberg, New York, 56-89.
50. Moreau J. J. (1977) Evolution problem associated with a moving convex set in a Hilbert space. J. Diff. Equ. **26**, 347–374
51. Moreau J. J. (1985) Standard inelastic shocks and the dynamics of unilateral constraints. In: Del Piero G., Maceri F. (Eds.) Unilateral Problems in Structural Analysis, CISM Courses and Lectures Vol. **288**. Springer-Verlag, Wien New York, 173–221
52. Moreau J. J. (1988) Unilateral contact and dry friction in finite freedom dynamics. In: Moreau J. J., Panagiotopoulos P.D. (Eds.) Nonsmooth Mechanics and Applications, CISM Courses and Lectures, Vol. **302**. Springer-Verlag, Wien New York, 1–82
53. Moreau J. J. (1988) Bounded variation in time. In: Moreau J. J. , Panagiotopoulos P. D., Strang G. (Eds.) Topics in Nonsmooth Mechanics. Birkhäuser, Basel Boston Berlin, 1–74
54. Moreau J. J. (1989) An expression of classical dynamics, Ann. Inst. H. Poincaré Anal. Non Linéaire. **6** (suppl.), 1-48. Volume also available as: Attouch H., Aubin J.-P., Clarke F., Ekeland I. (Eds.) Analyse Non Linéaire, Gauthier-Villars, Paris
55. Moreau J. J. (1994) Some numerical methods in multibody dynamics : application to granular materials. Eur. J. Mech., A/Solids **13**, n°4 - suppl., 93–114
56. Moreau J. J. (1995) Numerical experiments in granular dynamics: vibration-induced size segregation. In: Raous M. et al. (Eds.) Contact Mechanics. Plenum Press, New York, 347–358

57. Moreau J. J. (1997) Numerical investigation of shear zones in granular materials. In: Grassberger P., Wolf D. (Eds.) Proc. HLRZ-Workshop on Friction, Arching, Contact Dynamics. World Scientific, Singapore, 233–247
58. Moreau, J. J. (1999) Some basics of unilateral dynamics. In: Pfeiffer F., Glocker Ch. (Eds.) Unilateral Multibody Contacts. Kluwer, Dordrecht Boston London, 1–14
59. Moreau J. J. (1999) Numerical aspects of the sweeping process. In: Martins J. A. C., Klarbring A. (Eds.) Computational Modeling of Contact and Friction, special issue of Computer Meth. in Appl. Mech. and Engng. **177**, 329–349
60. Moreau, J. J. (1999) Évolutions en présence de liaisons unilatérales : notions de base. In: Guédra-Degeorges D., Ladevèze P., Raous M. (Eds.) Actes du 4e Colloque National en Calcul des Structures. Teknea, Toulouse, 25–40.
61. Moreau, J. J. (1999) Application des algorithmes "Contact Dynamics" aux milieux granulaires, Actes du 14e Congrès Français de Mécanique, Toulouse (CD ROM, ISBN 2-84088-040-7)
62. Moreau J. J. (2000) Contact et frottement en dynamique des systèmes de corps rigides. Rev. Europ. des Eléments Finis **9**, 9-28
63. Nouguier C., Bohatier C., Moreau J. J., Radjai F. (2000) Force fluctuations in a pushed granular material. Granular matter **2**, 171–178
64. Paoli L., Schatzman M. (1993) Mouvements à un nombre fini de degrés de liberté avec contraintes unilatérales: cas avec perte d'énergie. Math. Modelling and Num. Anal. **27**, 673–717
65. Pfeiffer F., Glocker Ch. (1996) Multibody Dynamics with Unilateral Contacts. John Wiley and Sons, New York
66. Raous M. (1993) Experimental analysis of the rocking of a rigid block. 3rd Pan American Congress of Applied Mechanics (PACAM III), Sao Paulo, Brazil
67. Raous M. (1999) Quasistatic Signorini problem with Coulomb friction and coupling to adhesion. In: Wriggers P., Panagiotopoulos P. (Eds) New developments in contact problems, CISM Courses and Lectures, Vol. **384**, Springer Verlag, Wien New York, 101–178
68. Raous M., Cangémi L., Cocu M. (1999) A consistent model coupling adhesion, friction and unilateral contact. Computer Meth. in Appl. Mech. and Engng. **177**, 383–399
69. Rockafellar R. T. (1970) Convex Analysis. Princeton Univ. Press, Princeton
70. Savage S. B. (1998) Modeling and granular material boundary value problems. In: Herrmann H. J. et al. (Eds.) Physics of dry ganular media. Kluwer, Dordrecht Boston London, 25–96
71. Stewart D. E. (1998) Convergence of a time-stepping scheme for rigid body dynamics and resolution of Painlevé's problem. Arch. Rational Mech. Anal. **145**, 215–260
72. Stoianovici S.P., Hurmuzlu Y. (1996) A critical study of the concepts of rigid body collision theory. J. Appl. Mech. **63**, 307–316
73. Stronge W. J. (2000) Contact problems for elasto-plastic impact in multi-body systems. In: Brogliato B. (Ed.) Impacts in Mechanical Systems. Analysis and Modelling. Springer, Berlin Heidelberg, 189–234
74. Trinkle J., Pang J.-S, Sudarski S., Lo G. (1997) On dynamic multirigid-body-contact problems with Coulomb friction. Zeitschr. Angew. Math. Mech. **77**, 267–279

75. Vola D., Pratt E., Jean M., Raous M. (1998) Consistent time discretization for a dynamical frictional contact problem and complementarity techniques. Rev. Europ. des Eléments Finis **7**, 149–162
76. Walton O. R. (1993) Numerical simulation of inelastic, frictional particle-particle interactions. In: Roco M. C. (Ed.) Particulate two-phase flow. Butterworth-Heinemann, Boston, 884-910