Multibody Syst Dyn (2013) 29:77–117
DOI 10.1007/s11044-012-9316-9

Energetic consistency conditions for standard impacts

Part I: Newton-type inequality impact laws and Kane's example

Christoph Glocker

Received: 24 January 2012 / Accepted: 23 April 2012 / Published online: 13 June 2012

Abstract The paper studies the conditions under which Newtonian impacts lead to an energetically consistent post-impact state, and identifies the mechanisms responsible for a potential violation. Unilateral and bilateral geometric and kinematic constraints, as well as various Coulomb type friction elements are equipped with a kinematic restitution coefficient and are taken into account in the impact problem as normal cone inclusions. Based on the condition number of the Delassus operator, bounds on the impact coefficients are derived that ensure energetic consistency. As counter-examples, a slide-push mechanism is presented, and Kane's example of a frictional impact at a double pendulum is analyzed.

Keywords Impact · Collision · Friction · Energy · Newton · Coulomb · Kane

1 Introduction

In mechanics, just two classical collision laws are available: Newton's kinematic impact law which inverts the normal relative velocity, and Poisson's impact law, which splits the impact into two phases and determines the post-impact velocity via the ratio of the impulsive forces. For frictional collisions, these impact laws are normally supplemented by a Coulomb friction element in the tangential direction. This combination of only collision and friction elements seems to be very limited when compared with the variety of force elements that are available and used in impact-free motion. Further elements which may transfer impulsive forces and have therefore to be considered in an extended impact dynamics are, e.g. bilateral geometric and kinematic constraints, but also the unilateral kinematic constraint with its technical realization as sprag clutches, and various extensions of the friction element to spatial configurations.

Against this background, a common concept for all the mentioned force elements is needed to put them uniformly into the framework of either Newton's or Poisson's law. The contribution at hand is divided into two parts. The first part is devoted to Newtonian impacts,

C. Glocker (✉)
IMES—Center of Mechanics, ETH Zurich, CLA J23.1, 8092 Zurich, Switzerland
e-mail: glocker@imes.mavt.ethz.ch

which are more extensively used than Poisson impacts, in particular within the non-smooth dynamics community. In the second part, Poisson impacts are treated, which lack to date a structurally sound formulation of the expansion phase for non-zero impact coefficients in friction elements.

Even the more basic concept of Newtonian impacts is not yet developed to a stage at which it could be regarded as completed and fully understood. Although known to exist, restitution coefficients in the friction elements are normally not taken into account when frictional impacts are analyzed, but regarded as a peculiarity because of their limited technical relevance. However, by banishing them from the friction laws, one immediately loses the opportunity to state the friction laws in the same mathematical form as the impact contacts, and one therefore destroys a common structure needed to set up a unifying theory. The better idea is to formally admit tangential restitution coefficients and set them later equal to zero if not needed. Even worse is the situation for sprag clutches. Their usage is acknowledged for certain application problems, but they rarely are mentioned when it comes to impacts. This constitutes a serious deficit, as sprag clutches are the technical realization of kinematic unilateral constraints, and as such they are as fundamental as the geometric unilateral constraints representing the impact contacts. It will be shown at the end of the paper that the friction element can be decomposed into kinematic unilateral constraints, demonstrating this way how powerful they are. In addition, since bilateral constraints can always be represented by pairs of unilateral constraints, it turns out that every and each element used for impact modeling can be traced back to the geometric and kinematic unilateral constraint. These two elements have therefore to be placed in the center of any and all questions on impacts.

Among the various conditions required by the principles of mechanics to hold for any force element, are three of particular interest for impact elements: Kinematic, kinetic, and energetic consistency. For Newtonian impacts, kinetic consistency is explicitly taken into account by the inequality restrictions on the impulsive forces as part of the constitutive laws and cannot be violated. Kinematic consistency poses conditions on the pre- and post-impact relative velocities and, therefore, restrictions on the impact coefficients. In contrast to kinematic consistency which can be ensured by looking at each impact element individually, energetic consistency requires the full evaluation of the entire system, which is much more demanding and treated in detail in this paper. Our view of energetic consistency is that the overall kinetic energy of the system must not increase at the impact, whereas we do not exclude a local increase in one of the impact elements needed, e.g. to realize an energy transport from one impact element to another. Whether Newtonian impact laws in general are capable for providing such a local increase is another question, and not treated in this paper.

It is known that Newtonian impacts may lead to an increase in kinetic energy when combined with friction. The most prominent example is Kane's double pendulum [13] which strikes the environment by a frictional impact, and which is analyzed in detail in this paper. It turns out that the mechanism of energy increase can easily be revealed as soon as the friction element is decomposed into its basic elements, the sprag clutches. Even more, one may easily design other mechanisms that show the same type of inconsistency by only using one geometric and one kinematic unilateral constraint.

The paper is organized as follows: In Sect. 2, the normal cone of convex analysis is reviewed, as it is later used to state the impact laws of various impact elements in the form of normal cone inclusions. It is further shown how the set-valued sign function, later representing the planar friction element, can be obtained through the normal cone on the interval $[-1, 1]$, and how it can be decomposed into unilateral primitives. This decomposition will be the key to represent the friction curve as an arrangement of sprag clutches, leading to an explanation of the energy gain in Kane's example.

In Sect. 3, the structure of the impact problem, as far as it concerns this paper, is set up. Investigated are autonomous Lagrangian systems, i.e. mechanical systems driven by a symmetric and positive definite mass matrix without any explicit dependency on time. The impact laws are supposed to be using a kinematic restitution coefficient and meeting the structure of normal cone inclusions, which defines what we understand by a Newton-type inequality impact law. It is further reviewed how the impact equations together with the impact laws can be brought into a form suitable for numerical evaluation, i.e. by introducing the Delassus operator and switching to local velocities.

Section 4 discusses various impact elements when equipped with a Newton-type inequality impact law. The following cases are treated: Geometric unilateral constraints, kinematic unilateral constraints, geometric bilateral constraints, kinematic bilateral constraints, kinematic step constraints of Coulomb type, isotropic Coulomb-type friction, and orthotropic Coulomb type friction. For each impact element, the physical meaning of the associated Newton-type inequality impact law is discussed in detail. We further state first restrictions on the restitution coefficients, which are necessary to ensure kinematic consistency.

In Sect. 5, the energetic consistency of the impact problem is investigated. In a first step, the classical restriction $0 \leq \varepsilon \leq 1$ on the restitution coefficient for geometric unilateral constraints is verified and extended to the other impact elements by assuming that the various impact elements do not interfere with each other, mathematically realized by assuming a block-diagonal Delassus operator. In a second step, fully coupled systems are investigated. Based on the condition number of the Delassus operator, two configurations are presented which necessarily lead to energetic consistency when the impact coefficients are small enough or similar to each other. As special cases of the above, equal impact coefficients as well as completely elastic and inelastic impacts are discussed. In a third step, it is shown that impact configurations consisting of only geometric unilateral, geometric bilateral, and kinematic bilateral constraints are always energetically consistent if the associated kinematic restrictions on the impact coefficients are satisfied.

Based on the energy proofs of Sect. 5, a slide-push mechanism is presented in Sect. 6 as the most basic example that contradicts energetic consistency. The system consists of two masses which interact with each other by a sprag clutch, and which may impact against the environment. It turns out that the combination of geometric and kinematic unilateral constraints is utmost disastrous for energetic consistency when equipped with different impact coefficients.

Section 7 provides a full analysis of the two cases addressed by Kane in [13], for which an increase in the kinetic energy is observed at the frictional impact of a double pendulum for post-impact slip and post-impact stick. Kane's problem is treated within the framework of non-smooth dynamics, for which the impact elements as introduced in Sect. 4 are used. In view of a possible explanation of the paradox, special emphasis is placed on the discussion of various impact elements that could alternatively represent the tangential impact.

In order to identify the mechanism leading to the energy increase in Kane's example, a system is designed which shows the same topology as the slide-push mechanism of Sect. 6. The parameter values of this system are chosen such that the equations for the impact coincide with those of the double pendulum. In this way, Kane's paradox can be identified to be caused by the same mechanism as the energy increase in the slide-push mechanism.

In Sect. 9, it is shown that the one-dimensional friction element can be represented by an arrangement of sprag clutches. This decomposition, originally introduced in [5] for impact-free motion, is basically the key to understand Kane's paradox, as it finally reduces the problem of a frictional impact to the multi-impact of a combination of geometric and kinematic unilateral constraints.

2 Normal cones, and the set-valued maps Upr and Sgn

Normal cones [23] are quintessential for a proper formulation of non-smooth dynamics [1, 15, 18, 24]. This applies for impact-free motion [7], but for the impacts as well [8]. The set-valued maps associated with the normal cones are called normal cone inclusions. They constitute a special case of the so-called generalized equations, which are used in optimization theory to express various types of constraints in the form of inclusions. In mechanics, normal cone inclusions can be used to achieve a mathematically sound formulation of a rather general class of force interactions with switching behavior, i.e. force elements that switch in a certain way between impressed and constraint forces. Even perfect bilateral constraints on geometric and kinematic level are covered by this approach. The classical definitions, commonly called the principles of d'Alembert/Lagrange and Jourdain, are nothing else than normal cone inclusions to specific sets $\mathscr{C}$; see [7].

The *normal cone* $\mathscr{N}_{\mathscr{C}}(\mathbf{y})$ to a convex subset $\mathscr{C}$ of $\mathbb{R}^n$ at a point $\mathbf{y} \in \mathscr{C}$ consists by definition [23] of all vectors $\mathbf{x} \in \mathbb{R}^n$ which do not form an acute "angle" with any line segment starting at $\mathbf{y}$ and ending at $\mathbf{y}^*$ somewhere in $\mathscr{C}$,

$$\mathscr{N}_{\mathscr{C}}(\mathbf{y}) = \left\{ \mathbf{x} \,|\, \mathbf{x}^{\mathrm{T}}\left(\mathbf{y}^* - \mathbf{y}\right) \leq 0 \text{ for } \mathbf{y} \in \mathscr{C},\ \forall \mathbf{y}^* \in \mathscr{C} \right\}. \tag{1}$$

If, for example, $\mathscr{C}$ is a convex subset of $\mathbb{R}^2$ as shown in Fig. 1, and $\mathbf{y}_1$ is a point on a smooth section of the boundary of $\mathscr{C}$, then the normal cone $\mathscr{N}_{\mathscr{C}}(\mathbf{y}_1)$ is a half-line perpendicular to the boundary of $\mathscr{C}$ at $\mathbf{y}_1$ with the 0-element placed at its end, i.e. at the location where $\mathbf{y}_1$ is. In the case that the boundary of $\mathscr{C}$ forms a corner as in $\mathbf{y}_2$, the normal cone is generated by the two half-lines of the adjoining smooth boundary sections. Finally, and most important, the normal cone to a point in the interior of $\mathscr{C}$ as, e.g. for $\mathbf{y}_3$, reduces by (1) to a single element which is the 0-element, $\mathscr{N}_{\mathscr{C}}(\mathbf{y}_3) = \{0\}$. Also note that the normal cone is indeed a cone, since $r\mathbf{x} \in \mathscr{N}_{\mathscr{C}}(\mathbf{y})$ whenever $\mathbf{x} \in \mathscr{N}_{\mathscr{C}}(\mathbf{y})$ and $r \geq 0$.

An additional, important property is gained from (1) if the convex set $\mathscr{C}$ contains the 0-element. In this case, it holds for any $\mathbf{x}$ from the normal cone $\mathscr{N}_{\mathscr{C}}(\mathbf{y})$ at a chosen point $\mathbf{y}$ that $-\mathbf{x}^{\mathrm{T}}\mathbf{y} \leq 0$,

$$\mathbf{x} \in \mathscr{N}_{\mathscr{C}}(\mathbf{y}) \quad \wedge \quad 0 \in \mathscr{C} \quad \Longrightarrow \quad -\mathbf{x}^{\mathrm{T}}\mathbf{y} \leq 0. \tag{2}$$

To verify this, simply set $\mathbf{y}^* = 0$ in (1). Applied to mechanical systems, this property will turn out as to indicate the dissipative character of certain impact relations and will be used in Sect. 5 when the energetic consistency conditions of the impact laws are discussed.

If $\mathscr{C}$ is represented by the level set of a differentiable convex function $f : \mathbb{R}^n \to \mathbb{R}$ as

$$\mathscr{C} = \left\{ \mathbf{y} \,|\, f(\mathbf{y}) \leq 0 \right\}, \tag{3}$$

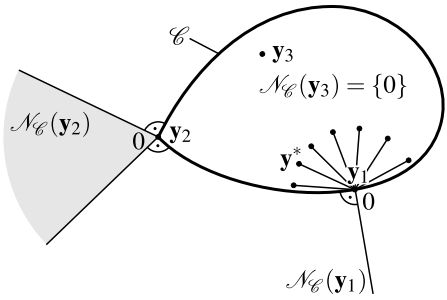

Fig. 1 The normal cone $\mathscr{N}_{\mathscr{C}}(\mathbf{y})$ to the convex set $\mathscr{C}$ at various points $\mathbf{y}_i \in \mathscr{C}$

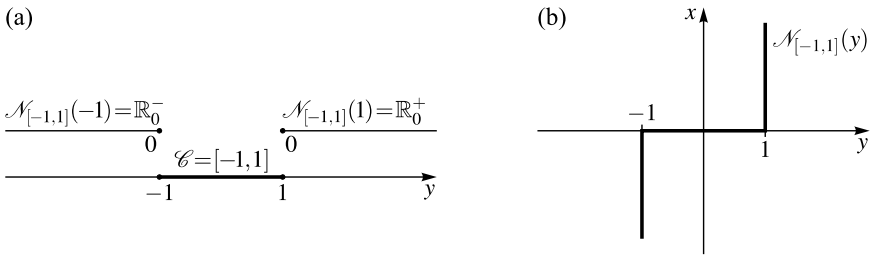

Fig. 2 (**a**) The normal cone $\mathcal{N}_{[-1,1]}(y)$ to the interval $[-1, 1] \subset \mathbb{R}$ at the points $y = -1$ and $y = 1$, together with (**b**) the graph of the resulting normal cone inclusion $x \in \mathcal{N}_{[-1,1]}(y)$

then it holds that

$$\mathbf{x} \in \mathcal{N}_{\mathscr{C}}(\mathbf{y}) \quad \Longleftrightarrow \quad \mathbf{x}^{\mathrm{T}} = \alpha \frac{\partial f}{\partial \mathbf{y}}, \quad f(\mathbf{y}) \leq 0, \quad \alpha \geq 0, \quad \alpha f(\mathbf{y}) = 0 \tag{4}$$

with $\alpha \in \mathbb{R}$ as shown in [23]; see also [4]. If f is not differentiable, the sub-differential has to be used in (4) instead.

As an illustrative example, we will now determine the normal cone $\mathcal{N}_{[-1,1]}(y)$ to the closed interval $[-1, 1] \subset \mathbb{R}$ at the points $y \in [-1, 1]$, and the graph of the associated normal cone inclusion $x \in \mathcal{N}_{[-1,1]}(y)$ by direct evaluation of (1). The normal cone for a chosen $y \in [-1, 1]$ is in this case described by the variational inequality $\mathcal{N}_{[-1,1]}(y) = \{x \,|\, x(y^* - y) \leq 0\}$, which has to hold for any and all $y^* \in [-1, 1]$. For its evaluation, one has to distinguish between the three structurally different cases that y is either in the interior or at the left or at the right boundary of the interval. We start with the case $y = -1$, for which the normal cone becomes $\mathcal{N}_{[-1,1]}(-1) = \{x \,|\, x(y^* + 1) \leq 0\}$. This expression has to hold for any and all $y^* \in [-1, 1]$. By testing all these elements on the term $(y^* + 1)$, we observe that $(y^* + 1) \geq 0$. With the abbreviation $\alpha := (y^* + 1)$, the normal cone at $y = -1$ therefore becomes $\mathcal{N}_{[-1,1]}(-1) = \{x \,|\, x\alpha \leq 0\}$ for any and all $\alpha \geq 0$. This is apparently only true if $x \leq 0$. Thus, $\mathcal{N}_{[-1,1]}(-1) = \{x \,|\, x \leq 0\} = (-\infty, 0]$; see also Fig. 2(a). The other two cases, i.e. y in the interior or at the right end of $[-1, 1]$ have to be processed in the same way, which altogether gives

$$\mathcal{N}_{[-1,1]}(y) = \begin{cases} (-\infty, 0] & \text{for } y = -1 \\ \{0\} & \text{for } -1 < y < 1 \\ [0, +\infty) & \text{for } y = 1. \end{cases} \tag{5}$$

Once the normal cone has been determined, the associated inclusion $x \in \mathcal{N}_{[-1,1]}(y)$ may be formulated and graphically displayed; see Fig. 2(b). We observe that the graph is maximal monotone [23] and describes an inverted set-valued sign function. We further see that this graph only meets the first and the third quadrant, but does not cross the second or the fourth quadrant. As a consequence, $xy \geq 0$ holds for each point (y, x) on this graph or, in other words, for each pair (y, x) fulfilling $x \in \mathcal{N}_{[-1,1]}(y)$. This property is nothing else than the realization of (2) for the set $\mathscr{C} = [-1, 1]$, which of course contains the 0-element.

We introduce now the *set-valued sign function* or filled-in relay function $\mathrm{Sgn}(x)$ as just announced, i.e. by inversion of the normal cone inclusion $x \in \mathcal{N}_{[-1,1]}(y)$ or equivalently, by mirroring the graph in Fig. 2(b) over the diagonal. The latter gives the graph displayed in

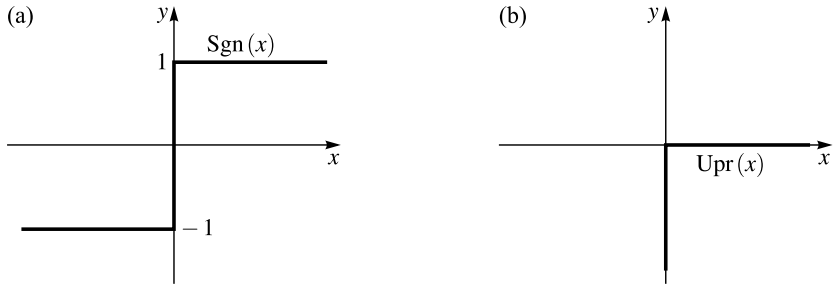

Fig. 3 The graphs of (**a**) the set-valued sign function $x \mapsto \mathrm{Sgn}(x)$ and of (**b**) the unilateral primitive $x \mapsto \mathrm{Upr}(x)$

Fig. 3(a) which is maximal monotone again, and which can mathematically be expressed by the inclusion $y \in \mathrm{Sgn}(x)$ with

$$\mathrm{Sgn}(x) := \begin{cases} \{+1\} & \text{if } x > 0 \\ [-1,+1] & \text{if } x = 0 \\ \{-1\} & \text{if } x < 0. \end{cases} \tag{6}$$

Note the difference at $x = 0$ to the classical sgn-function, which is defined as $\mathrm{sgn}(x = 0) = 0$. Just for completeness, the aforementioned inversion process is fundamentally anchored in convex analysis [23] and can mathematically be done in the following way: Normal cones are sub-differentials of indicators, which enables us to write $x \in \mathcal{N}_{[-1,1]}(y) = \partial I_{[-1,1]}(y)$. Take now the Legendre conjugate of $I_{[-1,1]}(y)$ which is a support function $I^*_{[-1,1]}(x)$, in this case equal to the absolute value $I^*_{[-1,1]}(x) = |x|$, and perform sub-differentiation with respect to x. This yields $\partial I^*_{[-1,1]}(x) = \partial |x| = \mathrm{Sgn}(x)$ which completes the inversion process. This approach is general and can be performed for any normal cone inclusion in precisely the same way.

There is another set-valued function, even more fundamental than the set-valued sign function (6). It has been named in [11] the *unilateral primitive* $\mathrm{Upr}(x)$ and is the most important multifunction related to complementarity. It is defined as

$$\mathrm{Upr}(x) := \begin{cases} \{0\} & \text{if } x > 0 \\ (-\infty, 0] & \text{if } x = 0, \end{cases} \tag{7}$$

and can also be related to normal cone inclusions in the way that $y \in \mathrm{Upr}(x)$ is the inverse of $x \in \mathcal{N}_{\mathbb{R}^-_0}(y)$. The graph of $y \in \mathrm{Upr}(x)$ is again maximal monotone and depicted in Fig. 3(b). It is not hard to see that $y \in \mathrm{Upr}(x)$ may alternatively be expressed by the so-called inequality-complementarity conditions

$$y \in \mathrm{Upr}(x) \quad \Longleftrightarrow \quad y \leq 0, \quad x \geq 0, \quad xy = 0, \tag{8}$$

which reveal linear complementarity problems as special cases of normal cone inclusion problems. The complementarity condition $xy = 0$ is again a manifestation of (2), which holds here even as an equality, because $0 \in \mathbb{R}^-_0$ is a boundary point and the *only* boundary point of $\mathbb{R}^-_0$.

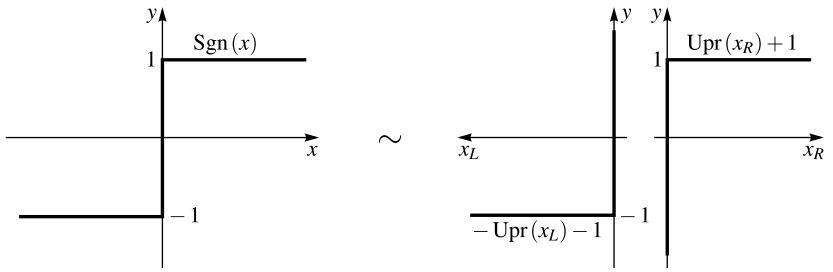

Fig. 4 Decomposition of Sgn into two unilateral primitives Upr, where $x = x_L - x_R$

The set-valued sign function (6) can be decomposed into two unilateral primitives as depicted in Fig. 4, which yields in terms of inclusions

$$y \in \mathrm{Sgn}(x) \quad \Longleftrightarrow \quad \exists x_R, x_L \text{ such that } \begin{cases} y \in +\mathrm{Upr}(x_R) + 1 \text{ and} \\ y \in -\mathrm{Upr}(x_L) - 1 \text{ and} \\ x = x_R - x_L. \end{cases} \tag{9}$$

Additional comments on this decomposition can be found in [7] together with various applications of set-valued interaction laws and their representations via unilateral primitives. We will later need this decomposition, because it allows us to interpret friction as a serial connection of sprag clutches. In this way, it will be possible to understand why frictional Newtonian impacts sometimes lead to an energy increase, as in the example of Kane [13].

3 The standard impact problem for inequality impact laws of Newton type

We define now the mathematical structure of the impact problem that will be presupposed for the discussion of energetic consistency in Sect. 5. A detailed description on how to obtain all the terms that are required to set up this structure for impact problems in multibody dynamics is found in [8]. As an impact we will understand a velocity jump which occurs at a discrete point t in time, and which is associated with infinite, i.e. impulsive forces as a consequence of finite, non-disappearing masses in the system. Processes with rapidly changing velocities but without discontinuities will not be understood as impacts as, e.g. from models that use local stiffness to represent the interactions in the contact zone.

Considered is an autonomous mechanical system within Lagrangian dynamics, described by a set of generalized coordinates $\mathbf{q} \in \mathbb{R}^f$. The generalized velocities, being assumed as functions of specially bounded variations and, therefore, allowing for a countable number of finite discontinuities [19], are denoted by $\mathbf{u}$ with $\mathbf{u} = \dot{\mathbf{q}}$ almost everywhere. After integration over the impact time $\{t\}$, we assume the resulting structure of the problem to be

$$\mathbf{M}\left(\mathbf{u}^+ - \mathbf{u}^-\right) = \sum_{i=1}^{n} \sum_{j=1}^{m(i)} \mathbf{w}_{ji} \Lambda_{ji}, \tag{10}$$

$$\gamma_{ji} = \mathbf{w}_{ji}^{\mathrm{T}} \mathbf{u}, \tag{11}$$

$$\xi_{ji} = \gamma_{ji}^{+} + \varepsilon_i \gamma_{ji}^{-}, \tag{12}$$

$$\text{impact law } (\xi_j, \Lambda_j)_i. \tag{13}$$

Equation (10) is the balance law for the impact which we call the *impact equation*, in contrast to the equation of motion for the impact-free case. The left-hand side of (10) consists of the instantaneous changes in generalized momentum, where $\mathbf{M}(\mathbf{q})$ denotes the symmetric and positive definite mass matrix of the system, and $\mathbf{u}^+$ and $\mathbf{u}^-$ are the post- and pre-impact generalized velocities. All generalized forces that may be of impulsive nature are collected in the right-hand side of (10). Finite forces, as from springs or dashpots, are non-impulsive and do therefore not contribute to the impact. We consider a total of n impact force elements, such as collision elements, Coulomb type friction elements, and the like. These impact elements are in detail described in Sect. 4. Each impact element i is characterized by $m(i)$ linearly independent generalized force directions $\mathbf{w}_{ji}(\mathbf{q})$, in which the impulsive force components Λ_{ji} act. Linear independence of the whole set $\mathbf{w}_{ji}$, however, is *not* required. Impulsive forces applied from the outside that can be regarded as external impact excitation, e.g. as the impulsive force from the queue on the ball when playing billiard, will not be examined in this context.

We denote by γ_{ji} the local velocity component in the impact element that is associated with the impulsive force Λ_{ji}. It is measured in the same direction as the impulsive force Λ_{ji} acts, and it therefore takes the form (11), with the same $\mathbf{w}_{ji}$ as in (10). This can easily be verified by the invariance of the associated virtual power expression under coordinate transformation: With $\delta P_{ji} = \Lambda_{ji}\delta\gamma_{ji} = \mathbf{R}_{ji}^{\mathrm{T}}\delta\mathbf{u}$ and $\delta\gamma_{ji} = \mathbf{w}_{ji}^{\mathrm{T}}\delta\mathbf{u}$ according to (11), one obtains for the generalized impulsive force $\mathbf{R}_{ji} = \mathbf{w}_{ji}\Lambda_{ji}$ as observed in (10).

The local velocities γ_{ji} in (11) are typically the relative velocities that are directly affected by the Newtonian impact law. We denote by γ_{ji}^+ and γ_{ji}^- the post- and pre-impact relative velocities, and by ε_i the Newtonian coefficient of restitution of the impact element i that relates them together. The magnitude ξ_{ji} in (12) is just an auxiliary variable, used to make the associated impact law more compact, as indicated in (13). The impact laws (13) on impact element component level are not yet further specified, because the individual directions j within the same impact element i may interact with each other as, e.g. for two-dimensional friction. This interaction may more clearly be represented by the impact laws on impact element level (17). Note however, that the impact laws (13) have to be regarded as constitutive laws, as they put kinematic magnitudes and force entities in a certain functional relationship.

Any explicit dependence on time is prohibited for the following terms in (10)–(13): For the mass matrix $\mathbf{M}$, for the generalized force directions $\mathbf{w}$, but also for the relative velocities γ which are usually of the form $\gamma = \mathbf{w}^{\mathrm{T}}\mathbf{u} + \chi$ with an inhomogeneity term χ that reflects external kinematic excitation. The latter may bring energy into the system, as for moving walls which the system may impact against. Such kind of time-dependence will not be allowed in this article to keep things reasonably simple. The *much more* complicated case of external kinematic excitation *without* assuming a specific impact law is treated in full generality for perfect unilateral geometric constraints in [2]; see also [3] for excerpts.

In order to rewrite (10)–(13) on impact element level, we collect for each impact element i its associated components j by setting $\mathbf{W}_i := (\mathbf{w}_{1i}, \ldots, \mathbf{w}_{m(i)i})$, $\boldsymbol{\Lambda}_i := (\Lambda_{1i}, \ldots, \Lambda_{m(i)i})^{\mathrm{T}}$, $\boldsymbol{\gamma}_i := (\gamma_{1i}, \ldots, \gamma_{m(i)i})^{\mathrm{T}}$, $\boldsymbol{\xi}_i := (\xi_{1i}, \ldots, \xi_{m(i)i})^{\mathrm{T}}$. In this way, (10)–(13) becomes

$$\mathbf{M}\left(\mathbf{u}^+ - \mathbf{u}^-\right) = \sum_{i=1}^{n} \mathbf{W}_i \boldsymbol{\Lambda}_i, \tag{14}$$

$$\boldsymbol{\gamma}_i = \mathbf{W}_i^{\mathrm{T}}\mathbf{u}, \tag{15}$$

$$\boldsymbol{\xi}_i = \boldsymbol{\gamma}_i^+ + \varepsilon_i \boldsymbol{\gamma}_i^-, \tag{16}$$

$$\boldsymbol{\xi}_i \in \mathscr{N}_{\mathscr{C}_i}(-\boldsymbol{\Lambda}_i), \tag{17}$$

where we have now introduced in (17) the impact laws of the individual impact elements as normal cone inclusions, which implies that the impulsive forces $-\boldsymbol{\Lambda}_i$ are restricted to the sets $\mathscr{C}_i$, i.e. $-\boldsymbol{\Lambda}_i \in \mathscr{C}_i$. We call an impact event a *standard impact* if it meets the structure laid down in equations (14) and (15). Non-standard impacts are encountered if, for example, sharp bends in the configuration manifold have to be crossed transversally. In this case, the generalized force directions $\mathbf{w}_{ji}$ may not be constant for certain indices j, but may instantaneously change when passing the bend. A more concrete example is a point mass which is confined to a rigid tube with a corner, constituting a bilateral constraint with kink. At the very instant at which the point mass reaches the kink, an impulsive force is required to instantaneously change the direction of the velocity of the mass, thereby keeping it on the constraint. We call (16) and (17) an *impact law of Newton type*, because it imposes a kinematic condition on the impact via the restitution coefficient ε_i, i.e. a condition on the pre- and post-impact relative velocities $\boldsymbol{\gamma}_i^{\pm}$ as in Newton's original law. We further call (16) and (17) an *inequality impact law*, because it handles different cases via the inequalities used to define the normal cone (1). Each of the sets $\mathscr{C}_i$ is required to be convex and to contain the 0-element, such that (2) holds in the form

$$0 \in \mathscr{C}_i \quad \Longrightarrow \quad \boldsymbol{\xi}_i^{\mathrm{T}} \boldsymbol{\Lambda}_i \leq 0. \tag{18}$$

The sets $\mathscr{C}_i$ are the reservoirs of negative impulsive forces $-\boldsymbol{\Lambda}_i$ that the impact elements can provide. Unfortunately, we have to allow that at least the size of some of these sets $\mathscr{C}_i$ may vary with the values $\boldsymbol{\Lambda}_k$ $(i \neq k)$ of other impact elements. This destroys some of the mathematical structure and beauty, but is necessary to take Coulomb type friction elements into account, for which the maximal tangential loads depend on the normal load.

To achieve an even more condensed formulation of (14)–(17), we set $\mathbf{W} := (\mathbf{W}_1, \ldots, \mathbf{W}_n)$, $\boldsymbol{\Lambda} := (\boldsymbol{\Lambda}_1^{\mathrm{T}}, \ldots, \boldsymbol{\Lambda}_n^{\mathrm{T}})^{\mathrm{T}}$, $\boldsymbol{\gamma} := (\boldsymbol{\gamma}_1^{\mathrm{T}}, \ldots, \boldsymbol{\gamma}_n^{\mathrm{T}})^{\mathrm{T}}$, $\boldsymbol{\xi} := (\boldsymbol{\xi}_1^{\mathrm{T}}, \ldots, \boldsymbol{\xi}_n^{\mathrm{T}})^{\mathrm{T}}$ and $\boldsymbol{\varepsilon} := \operatorname{diag}(\varepsilon_i \mathbf{I}_i)$, where $\mathbf{I}_i$ is the $m(i) \times m(i)$ identity matrix. We obtain

$$\mathbf{M}\big(\mathbf{u}^+ - \mathbf{u}^-\big) = \mathbf{W}\boldsymbol{\Lambda}, \tag{19}$$

$$\boldsymbol{\gamma} = \mathbf{W}^{\mathrm{T}}\mathbf{u}, \tag{20}$$

$$\boldsymbol{\xi} = \boldsymbol{\gamma}^+ + \boldsymbol{\varepsilon}\boldsymbol{\gamma}^-, \tag{21}$$

$$\text{impact law } (\boldsymbol{\xi}, \boldsymbol{\Lambda}). \tag{22}$$

Due to the unhealthy property $\mathscr{C}_i = \mathscr{C}_i(\boldsymbol{\Lambda}_k)$, it is *not* possible to state the impact law (22) as one condensed normal cone inclusion $\boldsymbol{\xi} \in \mathscr{N}_{\mathscr{C}}(-\boldsymbol{\Lambda})$ with $\mathscr{C} = \mathscr{C}_1 \times \cdots \times \mathscr{C}_n$, but it has to be left as it is. Nevertheless, it still holds that

$$\boldsymbol{\xi}^{\mathrm{T}} \boldsymbol{\Lambda} \leq 0, \tag{23}$$

which can be concluded from (18) by summing up the n individual terms $\boldsymbol{\xi}_i^{\mathrm{T}} \boldsymbol{\Lambda}_i$.

In order to solve the impact problem (19)–(22), one proceeds in the following way which has become standard during the last couple of years: One first states the difference in the post- and pre-impact relative velocities by (20) as

$$\boldsymbol{\gamma}^+ - \boldsymbol{\gamma}^- = \mathbf{W}^{\mathrm{T}}\big(\mathbf{u}^+ - \mathbf{u}^-\big), \tag{24}$$

and then eliminates with the help of (19) the term $(\mathbf{u}^+ - \mathbf{u}^-)$ to obtain

$$\boldsymbol{\gamma}^+ - \boldsymbol{\gamma}^- = \mathbf{G}\boldsymbol{\Lambda}, \tag{25}$$

where we have introduced the Delassus operator

$$\mathbf{G} := \mathbf{W}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{W}, \tag{26}$$

which is a symmetric and at least positive semi-definite matrix, set up by the sub-matrices $\mathbf{G}_{ik} = \mathbf{W}_i^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{W}_k$ of the n individual impact elements. Note that the sub-matrices on the diagonal, i.e. $\mathbf{G}_{ii} = \mathbf{W}_i^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{W}_i$, are symmetric and positive definite.

At this point, one has to decide whether the impact problem is to be solved forward or—if possible at all—backward in time. In the first case, the pre-impact relative velocities $\boldsymbol{\gamma}^-$ are given and should stay in the equations, whereas the unknown post-impact relative velocities $\boldsymbol{\gamma}^+$ have to be eliminated from them in favor for $\boldsymbol{\xi}$ which is needed for the impact law (22). In the second case, things are just the other way around. For completeness, we will show how both cases are treated. Equation (21) is solved for the variable that is going to be eliminated,

$$\boldsymbol{\gamma}^+ = \boldsymbol{\xi} - \boldsymbol{\varepsilon}\boldsymbol{\gamma}^- \quad \text{or} \quad \boldsymbol{\gamma}^- = \boldsymbol{\varepsilon}^{-1}\left(\boldsymbol{\xi} - \boldsymbol{\gamma}^+\right) \tag{27}$$

and then plugged into (25) which gives

$$\boldsymbol{\xi} = \mathbf{G}\boldsymbol{\Lambda} + (\mathbf{I} + \boldsymbol{\varepsilon})\boldsymbol{\gamma}^- \quad \text{or} \quad \boldsymbol{\xi} = -\boldsymbol{\varepsilon}\mathbf{G}\boldsymbol{\Lambda} + (\mathbf{I} + \boldsymbol{\varepsilon})\boldsymbol{\gamma}^+, \tag{28}$$

where $\mathbf{I}$ denotes the identity matrix. Either one in (28), together with the impact law (22), may then be processed by an inequality solver to get the solution $(\boldsymbol{\xi}, \boldsymbol{\Lambda})$. Once this solution has been obtained, the unknown relative velocities can be calculated from (27), and the unknown generalized velocities from (19). We will need Eqs. (28) again in Sect. 5 for setting up the changes in the kinetic energy at the impact in an appropriate form.

4 Impact elements

In view of a sound mechanical setting, the whole impact problem (14)–(17) has to be checked for kinetic, kinematic, and energetic consistency. Energetic consistency will be discussed in Sect. 5. We call a system kinetically consistent if the impact equations (14) and the force restrictions $-\boldsymbol{\Lambda}_i \in \mathscr{C}_i$ are met, which is obviously the case for any post-impact velocity $\mathbf{u}^+$ determined from (14)–(17). Newtonian impacts according to (14)–(17) are therefore always kinetically consistent.

Kinematic consistency is more subtle for Newtonian impact laws, because the relative velocities of the various impact elements are directly affected by the choice of the restitution coefficients. The only way for ensuring kinematic consistency is to analyze each of the impact elements separately, to get the appropriate restrictions on the restitution coefficient. We therefore discuss in each of the following subsections one particular impact element, also in view of its mechanical interpretation. The impact laws are all of the form $\boldsymbol{\xi}_i \in \mathscr{N}_{\mathscr{C}_i}(-\boldsymbol{\Lambda}_i)$ and differ from each other only in the choice of $\mathscr{C}_i$. For the kinematic consistency conditions, we assume throughout all subsections that the mechanical system has been in proper operation already before the impact, and that it leaves the impact by being in proper operation as well. Artificial inconsistencies resulting from a wrong initialization of the system via incompatible pre-impact initial conditions as, e.g. a closed contact with a pre-impact relative velocity such that the bodies separate, are excluded by this, but may be of importance for numerics.

4.1 Geometric unilateral constraints

The geometric unilateral constraint is the archetype of any collision model. It is represented by the inequality constraint $g_i(\mathbf{q}) \geq 0$ on the gap function g_i, which we assume to be equal to zero for contact, greater than zero for separation, and less than zero for forbidden overlap. If g_i measures the Euclidean distance between two boundary points on a common normal of the contacting bodies, then it can be shown that $\gamma_i = \dot{g}_i$ is the normal contact relative velocity with $\gamma_i = \mathbf{w}_i^{\mathrm{T}}\mathbf{u}$ and $\mathbf{w}_i^{\mathrm{T}} = \partial g_i/\partial \mathbf{q}$. We further recognize that the geometric unilateral constraint is a one-dimensional impact element, thus $m(i) = 1$ in (10).

As already a first part of the impact law we assume that non-vanishing impulsive normal forces Λ_i occur only for a closed contact $g_i = 0$, which is a reasonable assumption for collisions. For kinematic consistency, we have to take into account that the system has to arrive at $g_i = 0$ from an admissible state in the past, which necessarily requires a non-positive pre-impact normal relative velocity, $\gamma_i^- \leq 0$. Values $\gamma_i^- < 0$ correspond to an approaching process, at which the bodies touch under a non-vanishing velocity, whereas $\gamma_i^- = 0$ may correspond to a soft touchdown or to a contact that has been closed somewhere in the past. For the post-impact normal relative velocity, kinematic consistency requires $\gamma_i^+ \geq 0$, where strict inequality describes immediate separation as bouncing off, and equality indicates that the contact stays closed *on velocity level*. Note that the case $\gamma_i^+ = 0$ still leaves open whether there is an adjoined interval of motion for which the bodies keep contact, or for which they immediately separate on acceleration level, i.e. under the influence of non-impulsive forces from the adjoined impact-free motion.

We are now going to write down Newton's impact law for closed gaps $g_i = 0$ in an extended version that allows to differentiate between contacts which participate in the impact by non-vanishing impulsive forces Λ_i, and others which are superfluous for the impact process, although their gaps are closed, also. This distinction is based on a sign restriction on the impulsive normal force, which should act as a compressive magnitude $\Lambda_i \geq 0$ but not pull on the contact. In the case of a non-vanishing impulsive force, we apply Newton's impact law as usual, i.e. we "invert" the normal relative velocity according to the rule

$$\Lambda_i > 0 \quad \Longrightarrow \quad \gamma_i^+ = -\varepsilon_i \gamma_i^- . \tag{29}$$

The magnitude ε_i denotes Newton's coefficient of restitution. Due to kinematic consistency $\gamma_i^- \leq 0$, $\gamma_i^+ \geq 0$, its values have to be restricted to $\varepsilon_i \geq 0$. The case $\varepsilon_i = 0$ corresponds to a completely inelastic impact with vanishing post-impact relative velocity ($\gamma_i^+ = 0$), whereas $\varepsilon_i = 1$ leads to an inversion of the relative velocity ($\gamma_i^+ = -\gamma_i^-$).

Suppose now that, for any reason, the contact does *not* participate in the impact, i.e. that the value of the impulsive force is zero, although the contact is closed. This may happen in multi-contact configurations, such as in the example in Fig. 5(a). A rigid bar is resting on two unilateral supports and is hit by a rigid ball. We expect an impulsive reaction between the ball and the bar, but also between the bar and the left support, such that (29) holds. As a consequence, the ball and the left tip of the bar bounce back from each other, such that the bar is set in an instantaneous counter-clockwise rotation about the point at which it contacts the left support. This contact stays closed after the impact. Due to the post-impact angular velocity of the bar, a strictly positive relative velocity is induced by rigid body kinematics in the contact between the bar and the right support, without the need on having an impulsive force there. Such cases can be taken into account in the impact law by allowing post-impact relative velocities *higher* than what is prescribed by Newton's impact law, in combination with a vanishing impulsive force, i.e.

$$\Lambda_i = 0 \quad \Longrightarrow \quad \gamma_i^+ \geq -\varepsilon_i \gamma_i^- . \tag{30}$$

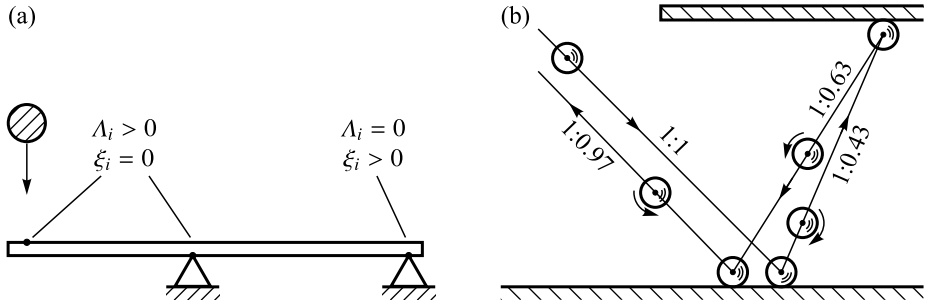

Fig. 5 (**a**) A configuration for which Newton's law is required in inequality form to allow the right contact to open at the impact. (**b**) Tangential restitution as needed for a ball to follow the displayed trajectory

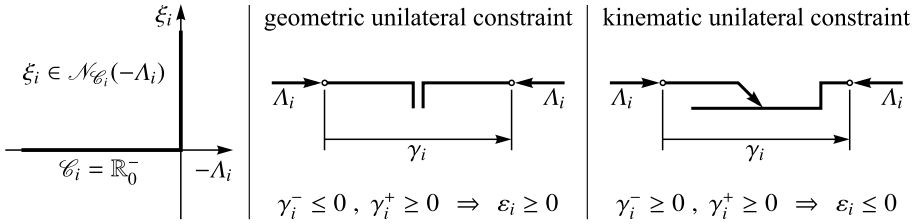

Fig. 6 Unilateral constraints: Graph of the normal cone inclusion, kinematic consistency conditions, and consequences on the impact coefficients

Closed contacts that are physically superfluous to support the impact process can be modeled in this way. Or, in other words, if the impact process would have been re-calculated by omitting the contacts for which (30) applies, the result would have been the same.

In order to write (29) and (30) in a more condensed form, we introduce according to (16) the abbreviation

$$\xi_i = \gamma_i^+ + \varepsilon_i \gamma_i^- . \tag{31}$$

With it, the full impact law (29), (30) is represented by one of the equivalent statements

$$\xi_i \ge 0, \quad \Lambda_i \ge 0, \quad \xi_i \Lambda_i = 0 \quad \text{or} \quad -\Lambda_i \in \operatorname{Upr}(\xi_i) \quad \text{or} \quad \xi_i \in \mathcal{N}_{\mathbb{R}_0^-}(-\Lambda_i), \tag{32}$$

i.e. by the inequality-complementarity conditions (8), by the Upr-inclusion (7), or by the associated normal cone inclusion (1) with $\mathscr{C}_i = \mathbb{R}_0^-$. The graph of the latter is displayed in Fig. 6, together with the kinematic consistency conditions and the symbol used to represent geometric unilateral constraints.

4.2 Kinematic unilateral constraints

The kinematic unilateral constraint is a linear inequality constraint on velocity level of the form $\gamma_i = \mathbf{w}_i^{\mathrm{T}} \mathbf{u} \ge 0$. It permits motion without any resistance in one direction but blocks in the other direction. The technical realization of a kinematic unilateral constraint is a sprag clutch. Kinematic unilateral constraints cannot impact by themselves, but they can react on impacts with unbounded forces in the blocked direction, $\Lambda_i \ge 0$. The dimension of this

impact element is $m(i) = 1$. Kinematic consistency requires for both the pre- and post-impact relative velocity $\gamma_i^- \geq 0$ and $\gamma_i^+ \geq 0$.

The only possibility to set up an impact law for kinematic unilateral constraints within the concept of Newton-type inequality impact laws (14)–(17) is to adopt the same rules (29)–(31) as for the geometric unilateral constraints. The physical interpretation of (29) for kinematic unilateral constraints is the following: If an impact, induced by, e.g. a collision somewhere else in the system, tries to move the sprag clutch in the blocked direction, there will be a reaction $\Lambda_i > 0$. The relative velocity in the sprag clutch will then be adjusted according to $\gamma_i^+ = -\varepsilon_i \gamma_i^-$, where the Newtonian impact coefficient has to fulfill $\varepsilon_i \leq 0$ to ensure kinematic consistency. Note that the relative velocity at a kinematic unilateral constraint does not change sign at the impact. The physical interpretation of (30) is as for geometric unilateral constraints, i.e. that the sprag clutch could be removed from the system without changing the impact process whenever $\Lambda_i = 0$.

The resulting impact law can again be stated by either of the three conditions (32), leading to the same graph as displayed in Fig. 6. The symbol of a kinematic unilateral constraint, also depicted in Fig. 6 together with the kinematic consistency conditions and the resulting restriction on the impact coefficient, has to be interpreted in the following way: The two ends of the sprag clutch can be moved without any resistance apart from each other. By trying to push the ends together, the sprag clutch instantaneously blocks and prevents any motion in this direction.

Hard stops and sprag clutches are important elements in engineering. However, if their impact behavior is modeled by Newton-type inequality impact laws according to (14)–(17) and Fig. 6, serious problems with respect to energetic consistency may arise. We will show an example in Sect. 6 in which the combination of just one geometric and one kinematic unilateral constraint may already lead to an increase in the overall kinetic energy when working with different impact coefficients. This situation is fortunately relieved if both impact coefficients are chosen to be equal to zero, for which everything works in perfect manner. Another reason to study kinematic unilateral constraints is to gain a better understanding of (Coulomb-type) friction elements and the problems with energy increase related to them. It will be shown in Sect. 9 that the one-dimensional friction element is nothing else than a certain arrangement of sprag clutches and constant loads.

4.3 Geometric bilateral constraints

Geometric bilateral constraints are not visible in the system's equations if minimal coordinates are used. For certain applications, it might however be more convenient to take them into account as superimposed constraints $g_i(\mathbf{q}) = 0$ by a Lagrangian multiplier technique. The Lagrangian multipliers for impact-free motion constitute the forces that are necessary to keep the system on the constraint. These forces are assumed to be of any size and, if necessary, even of impulsive nature Λ_i. As sprag clutches, geometric bilateral constraints cannot impact by themselves, but they have to react on impacts such that the constraint $g_i(\mathbf{q}) = 0$ will not be left afterward because of the generalized velocities that have been changed by the impact to $\mathbf{u}^+$. Geometric bilateral constraints have therefore to be processed in (14)–(17) together with all the other impact elements, to ensure that they leave the impact with compatible velocities $\mathbf{u}^+$ such that $g_i(\mathbf{q}) = 0$ can be maintained.

In order to include them in (14)–(17), an index reduction step has to be performed to bring them from the position to the velocity level. This is normally accomplished by choosing $\mathbf{q}_0 = \mathbf{q}(t_0)$ at a specific time instance t_0 such that $g_i(\mathbf{q}_0) = 0$ holds, and by ensuring that this value of $g_i(\mathbf{q})$ is maintained for all times. The latter is achieved by setting the time derivative

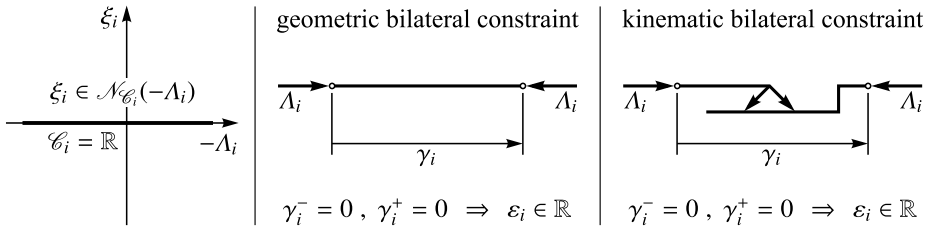

Fig. 7 Bilateral constraints: Graph of the normal cone inclusion, kinematic consistency conditions, and consequences on the impact coefficients

of $g_i(\mathbf{q})$ equal to zero, i.e. by demanding that $\dot{g}_i = \gamma_i = \mathbf{w}_i^{\mathrm{T}}\mathbf{u} = 0$ with $\mathbf{w}_i^{\mathrm{T}} = \partial g_i/\partial \mathbf{q}$. This condition has to apply in particular for the pre- and post-impact relative velocities, which reveals at once the kinematic consistency conditions for geometric bilateral constraints as $\gamma_i^- = 0$ and $\gamma_i^+ = 0$. The dimension of the associated impact element is again $m(i) = 1$.

The impact law for the geometric bilateral constraint can easily be brought into the form (16), (17). We take into account that the impulsive force Λ_i is not restricted in any way, and we formally apply Newton's impact law,

$$\Lambda_i \in \mathbb{R} \quad \Longrightarrow \quad \gamma_i^+ = -\varepsilon_i \gamma_i^- . \tag{33}$$

With $\gamma_i^- = 0$ due to pre-impact kinematic consistency, the result will always be $\gamma_i^+ = 0$ and ensure post-impact kinematic consistency, independent of the choice of the impact coefficient ε_i. For geometric bilateral constraints, there is no kinematic restriction on the impact coefficient. We now use again the abbreviation

$$\xi_i = \gamma_i^+ + \varepsilon_i \gamma_i^- \tag{34}$$

as in (16), and state (33) by either of the two conditions

$$\xi_i = 0, \quad \Lambda_i \in \mathbb{R} \quad \text{or} \quad \xi_i \in \mathscr{N}_{\mathbb{R}}(-\Lambda_i). \tag{35}$$

Note that $-\Lambda_i$ is always in the interior of the set $\mathscr{C}_i = \mathbb{R}$, which by (1) causes the normal cone to consist of the 0-element only, thus $\xi_i = 0$. In other words, the normal cone inclusion is single-valued for this case, and is identified as the trivial function $\xi_i(-\Lambda_i) = 0$. The associated graph together with the symbol for the geometric bilateral constraint and the kinematic consistency conditions are depicted in Fig. 7. Finally, note that a geometric bilateral constraint $g_i(\mathbf{q}) = 0$ can always be generated by a parallel connection of two geometric unilateral constraints $g_i(\mathbf{q}) \geq 0$, $g_k(\mathbf{q}) \geq 0$ with $g_k = -g_i$. For an exhaustive geometric analysis of this interplay, we refer to [2].

4.4 Kinematic bilateral constraints

Kinematic bilateral constraints are not visible in the system's equation if minimal velocities are used. This, however, would destroy our assumption $\dot{\mathbf{q}} = \mathbf{u}$ which would have to be replaced by $\dot{\mathbf{q}} = \mathbf{H}(\mathbf{q})\mathbf{u}$ and would slightly affect the structure supposed in (14)–(17). Although not a big hurdle, we prefer to stick to the aforementioned structure for clarity. We therefore only allow kinematic bilateral constraints to be superimposed on the system with $\dot{\mathbf{q}} = \mathbf{u}$. As known from classical mechanics, they always are of the form $\gamma_i = \mathbf{w}_i^{\mathrm{T}}\mathbf{u} = 0$, which leads to the same kinematic consistency conditions $\gamma_i^- = 0$, $\gamma_i^+ = 0$ as for geometric

bilateral constraints. Moreover, everything said about the geometric bilateral constraint in Sect. 4.3 stays true for the kinematic bilateral constraint, except from the index reduction process that is not needed here. In particular, the impact law and its interpretation (33)–(35) may be adopted without any changes, which leads to the results displayed in Fig. 7. Also, the last statement of the previous subsection remains true in a slightly modified form, i.e. that a kinematic bilateral constraint $\gamma_i = \mathbf{w}_i^{\mathrm{T}}\mathbf{u} = 0$ can always be generated by a parallel connection of two kinematic unilateral constraints $\gamma_i = \mathbf{w}_i^{\mathrm{T}}\mathbf{u} \geq 0$, $\gamma_k = \mathbf{w}_k^{\mathrm{T}}\mathbf{u} \geq 0$ with $\gamma_k = -\gamma_i$. This is even indicated by the symbol that we have chosen to represent kinematic bilateral constraints.

4.5 Kinematic step constraints of Coulomb type

The kinematic step constraint of Coulomb type is actually the impact element that is used to model frictional effects in collisions for the planar case. It therefore constitutes a one-dimensional ($m(i) = 1$) Coulomb-type friction element that acts on the tangential relative velocity $\gamma_i = \mathbf{w}_i^{\mathrm{T}}\mathbf{u}$ by the tangential impulsive force Λ_i. We call this impact element kinematic, because it is formulated by nature on velocity but not on displacement level. We call it a step constraint, because it is represented by a set-valued sign function as we will see, which forms an upward step of a priory unknown size by limiting the realizable tangential impulsive forces Λ_i. We call it of Coulomb type, because the reservoir $\mathscr{C}_i(\Lambda_k) = \mu_i \Lambda_k[-1, 1]$ of the tangential impulsive forces $-\Lambda_i$ is chosen such that it depends by the coefficient of friction μ_i on the impulsive force $\Lambda_k \geq 0$ of another impact element, which in most cases is the geometric unilateral constraint that models the collision in the normal direction of the contact.

We are now going to review the Coulomb-type friction element as originally presented in [20], and to state it as a normal cone inclusion according to (16), (17). As already mentioned, the value of the tangential impulsive force is restricted to $|\Lambda_i| \leq \mu_i \Lambda_k$, where μ_i denotes the coefficient of friction. For the case $|\Lambda_i| < \mu_i \Lambda_k$, pure Newtonian collision behavior is proposed,

$$|\Lambda_i| < \mu_i \Lambda_k \quad \Longrightarrow \quad \gamma_i^+ = -\varepsilon_i \gamma_i^-, \tag{36}$$

which yields a reversion of the tangential relative velocity in the same fashion as for the geometric unilateral constraint (29). Tangential restitution is taken into account by the coefficient ε_i. Although misleading, we call (36) the *stick state* of the impact element, according to the terminology of impact-free motion. Note that the post-impact relative velocity γ_i^+ does *not* vanish, i.e. is *unequal* to zero for $\gamma_i^- \neq 0$ and $\varepsilon_i \neq 0$.

Similar to (30), one completes now the tangential impact law for the case $|\Lambda_i| = \mu_i \Lambda_k$ by setting

$$\begin{aligned} \Lambda_i = -\mu_i \Lambda_k \quad &\Longrightarrow \quad \gamma_i^+ \geq -\varepsilon_i \gamma_i^-, \\ \Lambda_i = +\mu_i \Lambda_k \quad &\Longrightarrow \quad \gamma_i^+ \leq -\varepsilon_i \gamma_i^-, \end{aligned} \tag{37}$$

which is called the *slip state*. One easily recognizes that Coulomb type friction is present, but also that the slip state still carries something from the Newtonian collision behavior.

There is no kinematical restriction on the pre- or post-impact relative velocities γ_i^- or γ_i^+, hence the impact coefficient ε_i can freely be chosen according to the physical needs. By setting

$$\xi_i = \gamma_i^+ + \varepsilon_i \gamma_i^-, \tag{38}$$

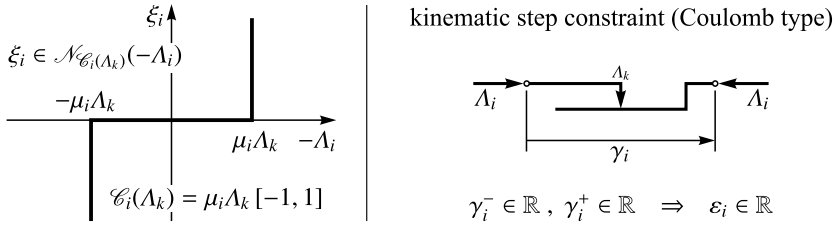

Fig. 8 One-dimensional Coulomb type friction: Graph of the normal cone inclusion, kinematic consistency conditions, and consequences on the impact coefficients

both cases (36) and (37) can be expressed by the set-valued sign function (6) or likewise the normal cone inclusion (5),

$$-\Lambda_i \in \mu_i \Lambda_k \,\mathrm{Sgn}(\xi_i) \quad \text{or} \quad \xi_i \in \mathcal{N}_{\mu_i \Lambda_k [-1,1]}(-\Lambda_i). \tag{39}$$

The graph of the latter, together with the symbol for the friction impact element and the kinematic consistency conditions are depicted in Fig. 8. Tangential impact coefficients ε_i unequal to zero are indeed needed to compute certain collision behavior, such as the trajectory displayed in Fig. 5(b) of a highly elastic rubber ball that is thrown under a table. After the second collision, it reverts the direction of its horizontal velocity component and bounces back, see [8] for the full calculation. If ε_i would have been set equal to zero in this example, the ball would stay under the table as observed for ping-pong balls. The associated trajectory may be found in [4] or [22].

Although the above example shows the necessity for taking into account tangential restitution for certain cases, it normally does not play an important role in applications and is therefore not considered in many papers on the subject. We will keep tangential restitution in our formulation, because precisely the impact coefficients ε_i are needed to derive the energetic consistency conditions in Sect. 5, which forces us not to throw them away too early. Nevertheless, some words should be spent on the case $\varepsilon_i = 0$, i.e. pure Coulomb type friction without restitution, due to its practical importance. The most explicit form of the impact law for this special case is directly obtained from (36) and (37) and reads as

$$\begin{aligned} |\Lambda_i| < \mu_i \Lambda_k \quad &\Longrightarrow \quad \gamma_i^+ = 0, \\ \Lambda_i = -\mu_i \Lambda_k \quad &\Longrightarrow \quad \gamma_i^+ \geq 0, \\ \Lambda_i = +\mu_i \Lambda_k \quad &\Longrightarrow \quad \gamma_i^+ \leq 0. \end{aligned} \tag{40}$$

One observes that the stick case from the first line in (40) now indeed yields post-impact tangential relative velocities $\gamma_i^+ = 0$, fulfilling the kinematic non-slip condition.

In order to discuss the four possible state combinations at a frictional impact, we rewrite (40) in inverted form,

$$\begin{aligned} \gamma_i^+ = 0 \quad &\Longrightarrow \quad |\Lambda_i| \leq \mu_i \Lambda_k, \\ \gamma_i^+ \neq 0 \quad &\Longrightarrow \quad \Lambda_i = -\mu_i \Lambda_k \,\mathrm{sgn}\big(\gamma_i^+\big). \end{aligned} \tag{41}$$

For a transition of pre-impact slip ($\gamma_i^- \neq 0$) to post-impact stick ($\gamma_i^+ = 0$), the impact law behaves in accordance with Coulomb's law by limiting the tangential impulsive force to $|\Lambda_i| \leq \mu_i \Lambda_k$, as seen from the first line in (41). For slip without reversal, i.e. $\gamma_i^\pm \neq 0$ with

$\operatorname{sgn}(\gamma_i^+) = \operatorname{sgn}(\gamma_i^-)$, the impulsive force is $\Lambda_i = -\mu_i \Lambda_k \operatorname{sgn}(\gamma_i^+)$ by the second line in (41) and again in accordance with Coulomb's law.

The remaining two cases, i.e. slip reversal with $\gamma_i^\pm \neq 0$ and $\operatorname{sgn}(\gamma_i^+) = -\operatorname{sgn}(\gamma_i^-)$, as well the stick-slip transition with $\gamma_i^- = 0$ and $\gamma_i^+ \neq 0$, are again driven by the second line in (41) and yield a tangential impulsive force equal to $\Lambda_i = -\mu_i \Lambda_k \operatorname{sgn}(\gamma_i^+)$. Compared to Coulomb's law, this value is too big, because successive integration over a phase of sliding and reversed sliding would lower the absolute value of the resulting impulsive force. The same is true for any phase of sticking, for which the absolute value of the impulsive tangential force is strictly less than $\mu_i \Lambda_k$.

This problem has already been mentioned in [18] and thoroughly discussed in [4] and [22]. It is a consequence on the impossibility to evaluate the integral of the tangential forces without additional assumptions. The impact parameter ε_i helps sometimes to downplay this problem, but does not remove its source. For more information on how this tangential impact law behaves, we refer in particular to the paper [21], in which the frictional impact of a rigid bar against an inelastic half-space has been analyzed in full detail.

Due to the said discrepancies of the original Coulomb law for impact-free motion, we prefer to call this impact element of Coulomb *type*. It still shows the main properties of Coulomb friction, which are the dependence on the normal load, and the proper realization of a stick phase. We further want to anticipate that the above discrepancy is related to but *not* responsible for the energy gain in Kane's example.

4.6 Isotropic Coulomb-type friction

Isotropic Coulomb-type friction is the simplest two-dimensional version ($m(i) = 2$) of the Coulomb-type kinematic step constraint from the previous section, containing all the nice and less nice properties of the former. Together with the geometric unilateral constraint, it is used to model spatial contact situations. In order to set up the associated impact law as a normal cone inclusion, we choose two directions in the common tangent plane of the contacting bodies, orthogonal to each other in the Euclidean metric, in which the two components $\gamma_{1i} = \mathbf{w}_{1i}^{\mathrm{T}}\mathbf{u}$, $\gamma_{2i} = \mathbf{w}_{2i}^{\mathrm{T}}\mathbf{u}$ of the relative velocity $\boldsymbol{\gamma}_i = (\gamma_{1i}, \gamma_{2i})^{\mathrm{T}}$ and the two components Λ_{1i}, Λ_{2i} of the tangential impulsive force $\boldsymbol{\Lambda}_i = (\Lambda_{1i}, \Lambda_{2i})^{\mathrm{T}}$ are measured. The friction coefficient is again denoted by μ_i, and the associated impulsive normal force by Λ_k; see Fig. 9.

We now generalize (36) and (37) straight forwardly to the two-dimensional situation, in the same way as commonly done for impact-free motion: For the stick state $\|\boldsymbol{\Lambda}_i\| < \mu_i \Lambda_k$, we claim pure Newtonian collision behavior as in (36),

$$\|\boldsymbol{\Lambda}_i\| < \mu_i \Lambda_k \quad \Longrightarrow \quad \boldsymbol{\gamma}_i^+ = -\varepsilon_i \boldsymbol{\gamma}_i^- . \tag{42}$$

Note that the direction of $\boldsymbol{\gamma}_i^-$ is reversed by the impact, as if the contact would bounce against a wall perpendicular to $\boldsymbol{\gamma}_i^-$ or $\boldsymbol{\gamma}_i^+$. In order to treat the slip state, we first define the slip direction $\mathbf{e}_i$ as opposite to the impulsive force $\boldsymbol{\Lambda}_i$,

$$\mathbf{e}_i(\boldsymbol{\Lambda}_i) := -\frac{\boldsymbol{\Lambda}_i}{\|\boldsymbol{\Lambda}_i\|} . \tag{43}$$

If slip is valid, we allow as in (37) the post-impact velocities $\boldsymbol{\gamma}_i^+$ to exceed the value of $-\varepsilon_i \boldsymbol{\gamma}_i^-$, *but only in the slip direction* $\mathbf{e}_i$,

$$\|\boldsymbol{\Lambda}_i\| = \mu_i \Lambda_k \quad \Longrightarrow \quad \boldsymbol{\gamma}_i^+ = -\varepsilon_i \boldsymbol{\gamma}_i^- + \kappa_i \mathbf{e}_i(\boldsymbol{\Lambda}_i) \quad \text{with } \kappa_i \geq 0. \tag{44}$$

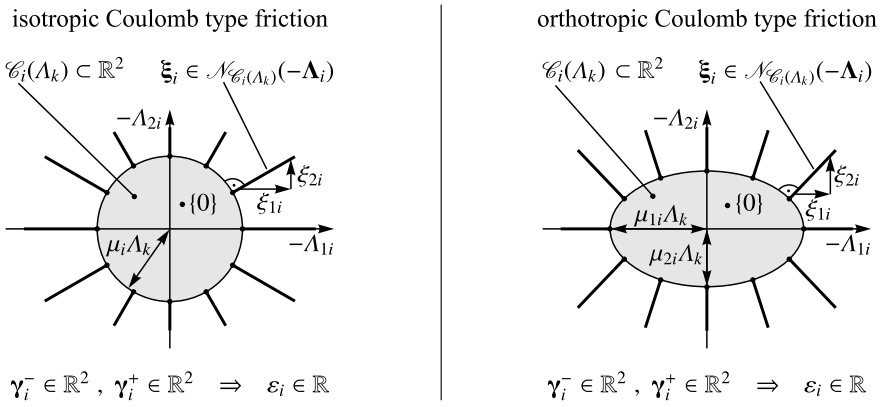

Fig. 9 Two-dimensional Coulomb type friction: Sets $\mathscr{C}_i$ of negative admissible impulsive forces $\boldsymbol{\Lambda}_i$, associated normal cones, kinematic consistency conditions, and consequences on the impact coefficients

This equation seems to be very artificial, but is needed in this form to finally reach the structure of a normal cone inclusion (16), (17). Comments on the justification of the proposed approach are made in Sect. 4.8.

There is no kinematic restriction on the pre- or post-impact relative velocities $\boldsymbol{\gamma}_i^-$ or $\boldsymbol{\gamma}_i^+$, and hence ε_i can freely be chosen. According to (16), we introduce

$$\boldsymbol{\xi}_i = \boldsymbol{\gamma}_i^+ + \varepsilon_i \boldsymbol{\gamma}_i^- \tag{45}$$

and write the two cases (42) and (44) together as

$$\begin{aligned} \|\boldsymbol{\Lambda}_i\| < \mu_i \Lambda_k \quad &\Longrightarrow \quad \boldsymbol{\xi}_i = 0, \\ \|\boldsymbol{\Lambda}_i\| = \mu_i \Lambda_k \quad &\Longrightarrow \quad \boldsymbol{\xi}_i = \kappa_i \mathbf{e}_i(\boldsymbol{\Lambda}_i), \quad \kappa_i \geq 0. \end{aligned} \tag{46}$$

From the restrictions on $\boldsymbol{\Lambda}_i$ in (46) one identifies the reservoir $\mathscr{C}_i(\Lambda_k)$ of the tangential impulsive forces as

$$\mathscr{C}_i(\Lambda_k) = \left\{ \boldsymbol{\Gamma}_i \in \mathbb{R}^2 \,|\, \|\boldsymbol{\Gamma}_i\| \leq \mu_i \Lambda_k \right\} \quad \text{with } -\boldsymbol{\Lambda}_i \in \mathscr{C}_i(\Lambda_k), \tag{47}$$

which is a circular disk with radius $\mu_i \Lambda_k$ as displayed in the left diagram of Fig. 9. The impact law (46) can now again be formulated as a normal cone inclusion $\boldsymbol{\xi}_i \in \mathscr{N}_{\mathscr{C}_i(\Lambda_k)}(-\boldsymbol{\Lambda}_i)$ as explicitly shown in the next section. As seen from the second line in (46), the normal cone becomes the half-line $\bigcup_{\kappa_i \geq 0} \kappa_i \mathbf{e}_i$ generated by the slip direction $\mathbf{e}_i$ if $-\boldsymbol{\Lambda}_i$ is at the boundary of $\mathscr{C}_i(\Lambda_k)$. We finally want to point out that $\boldsymbol{\Lambda}_i, \mathbf{e}_i, \boldsymbol{\xi}_i$ are aligned for slip, but *not* necessarily $\boldsymbol{\xi}_i, \boldsymbol{\gamma}_i^-, \boldsymbol{\gamma}_i^+$. However, if $\varepsilon_i = 0$, then $\boldsymbol{\xi}_i$ and $\boldsymbol{\gamma}_i^+$ have the same direction which is also the direction of $\mathbf{e}_i$ and $-\boldsymbol{\Lambda}_i$.

4.7 Orthotropic Coulomb-type friction

In this section, we propose an impact model for orthotropic Coulomb type friction ($m(i) = 2$) that is strictly derived from a normal cone inclusion. For the reservoir of tangential impulsive forces $-\boldsymbol{\Lambda}_i$, we assume an elliptical set $\mathscr{C}_i(\Lambda_k)$ with its semi-axes propor-

tional to the two friction coefficients μ_{1i}, μ_{2i} as depicted in the right part of Fig. 9,

$$\mathscr{C}_i(\Lambda_k) = \left\{ \boldsymbol{\Gamma}_i \in \mathbb{R}^2 \Big| \frac{\Gamma_{1i}^2}{(\mu_{1i}\Lambda_k)^2} + \frac{\Gamma_{2i}^2}{(\mu_{2i}\Lambda_k)^2} \leq 1 \right\}. \tag{48}$$

In order to obtain the normal cone via the conditions in (4), this set needs to be represented in the form

$$\mathscr{C}_i(\Lambda_k) = \big\{ \boldsymbol{\Gamma}_i \,|\, f(\boldsymbol{\Gamma}_i) \leq 0 \big\}, \tag{49}$$

by an appropriate differentiable function f which we choose as

$$f(\boldsymbol{\Gamma}_i) = \frac{\Gamma_{1i}^2}{(\mu_{1i}\Lambda_k)^2} + \frac{\Gamma_{2i}^2}{(\mu_{2i}\Lambda_k)^2} - 1. \tag{50}$$

Further, the differential of f

$$\frac{\partial f}{\partial \boldsymbol{\Gamma}_i} = \frac{2}{\Lambda_k^2}\left(\frac{\Gamma_{1i}}{\mu_{1i}^2}, \frac{\Gamma_{2i}}{\mu_{2i}^2} \right) \tag{51}$$

is needed to state the desired normal cone inclusion (17) by the conditions in (4),

$$\begin{aligned} &\boldsymbol{\xi}_i \in \mathscr{N}_{\mathscr{C}_i(\Lambda_k)}(-\boldsymbol{\Lambda}_i) \\ &\iff \quad \boldsymbol{\xi}_i^{\mathrm{T}} = \alpha_i \frac{\partial f}{\partial \boldsymbol{\Gamma}_i}(-\boldsymbol{\Lambda}_i), \quad f(-\boldsymbol{\Lambda}_i) \leq 0, \quad \alpha_i \geq 0, \quad \alpha_i f(-\boldsymbol{\Lambda}_i) = 0. \end{aligned} \tag{52}$$

By further evaluating the inequality-complementarity conditions in (52), one obtains

$$\begin{aligned} &\boldsymbol{\xi}_i \in \mathscr{N}_{\mathscr{C}_i(\Lambda_k)}(-\boldsymbol{\Lambda}_i) \\ &\iff \quad \begin{cases} f(-\boldsymbol{\Lambda}_i) < 0 \implies \boldsymbol{\xi}_i^{\mathrm{T}} = 0 \\ f(-\boldsymbol{\Lambda}_i) = 0 \implies \boldsymbol{\xi}_i^{\mathrm{T}} = \alpha_i \frac{\partial f}{\partial \boldsymbol{\Gamma}_i}(-\boldsymbol{\Lambda}_i), \quad \alpha_i \geq 0. \end{cases} \end{aligned} \tag{53}$$

If a representation of (53) in terms of a normalized slip direction $\mathbf{e}_i(\boldsymbol{\Lambda}_i)$ is desired, one may finally set

$$\kappa_i := \alpha_i \left\| \frac{\partial f}{\partial \boldsymbol{\Gamma}_i} \right\| (-\boldsymbol{\Lambda}_i) \quad \text{and} \quad \mathbf{e}_i^{\mathrm{T}}(\boldsymbol{\Lambda}_i) := \left(\frac{\partial f}{\partial \boldsymbol{\Gamma}_i} \Big/ \left\| \frac{\partial f}{\partial \boldsymbol{\Gamma}_i} \right\| \right)(-\boldsymbol{\Lambda}_i) \tag{54}$$

and rewrite the conditions on the right of (53) as

$$\begin{aligned} f(-\boldsymbol{\Lambda}_i) < 0 \quad &\implies \quad \boldsymbol{\xi}_i = 0, \\ f(-\boldsymbol{\Lambda}_i) = 0 \quad &\implies \quad \boldsymbol{\xi}_i = \kappa_i \mathbf{e}_i(\boldsymbol{\Lambda}_i), \quad \kappa_i \geq 0. \end{aligned} \tag{55}$$

Equation (55) has the same structure as the impact law (46) for the isotropic case. The first line in (55) represents the stick state, for which the impulsive forces $\boldsymbol{\Lambda}_i$ are in the interior of the set $\mathscr{C}_i$ and the relative velocities are inverted by $\boldsymbol{\xi}_i = 0$, i.e. $\boldsymbol{\gamma}_i^+ = -\varepsilon_i \boldsymbol{\gamma}_i^-$. The second line describes slip with the impulsive forces $\boldsymbol{\Lambda}_i$ on the boundary of $\mathscr{C}_i$ and the slip direction $\mathbf{e}_i(\boldsymbol{\Lambda}_i)$ essentially determined by the differential (51) of f. Note that the normal cone for the case of slip is still a half-line, generated by $\mathbf{e}_i(\boldsymbol{\Lambda}_i)$ and orthogonal to the boundary of $\mathscr{C}_i$, but is in general no longer in radial direction as for the isotropic case. Finally, note that isotropic friction is a special case of orthotropic friction, which is obtained when taking equal friction coefficients $\mu_i := \mu_{1i} = \mu_{2i}$ in (48), which makes (55) to become identical with (46).

4.8 Comments on impact modeling

The procedure shown in Sect. 4.7 is general and may be applied to any convex set $\mathscr{C}_i$ to formulate the associated Newton-type impact law in form of a normal cone inclusion. One further example is Coulomb–Contensou friction which takes into account the interaction between planar sliding friction and pivoting friction [9, 14], and thus requires a three-dimensional impact law ($m(i) = 3$).

Impact laws are constitutive laws. They are never general, and they cannot be derived from any principle in dynamics. They have to be identified by measurements, which is a very hard task, in particular for spatial collisions. In addition, the impact laws have to match the associated constitutive laws for impact-free motion. This is guaranteed in (14)–(17) by taking sets $\mathscr{C}_i$ with the very same shape as those used for impact-free motion, on which the normal cones are then built. In this way, hard contacts and dry friction are represented for impacts and impact-free motion by the same mathematical objects, and allow for a combined numerical treatment when time integration is performed.

Normal cone inclusions are by no means artificial mathematical entities, but they are required to generalize d'Alembert/Lagrange's principle in a natural way to situations in which constraint forces are restricted to convex sets. As d'Alembert/Lagrange's principle, i.e. the definition of a perfect constraint, normal cone inclusions describe force interactions in an idealized way that may not apply satisfactorily for all practical applications. In such cases, it seems to be advisable to start with a normal cone inclusion nonetheless, and to carefully adjust it afterward according to the physical needs. This, however, affects immediately the energetic consistency conditions derived in the next section, and one has to make sure by oneself that nothing goes wrong.

5 Energetic consistency conditions

In this section, we derive several sufficient conditions under which the impact according to (14)–(17) does not lead to an increase in energy. Since the positions $\mathbf{q}$ are assumed to be constant during the impact, only the difference in kinetic energy counts. We consider systems for which the kinetic energy is purely quadratic,

$$T = \frac{1}{2}\mathbf{u}^{\mathrm{T}}\mathbf{M}\mathbf{u}, \tag{56}$$

with $\mathbf{M}(\mathbf{q})$ the symmetric and positive definite mass matrix and $\mathbf{u}$ the generalized velocities. This applies, e.g. for the case that the particle coordinates do not explicitly depend on time. The difference of post-impact and pre-impact kinetic energy is

$$T^+ - T^- = \frac{1}{2}\mathbf{u}^{+\mathrm{T}}\mathbf{M}\mathbf{u}^+ - \frac{1}{2}\mathbf{u}^{-\mathrm{T}}\mathbf{M}\mathbf{u}^- = \frac{1}{2}\big(\mathbf{u}^+ + \mathbf{u}^-\big)^{\mathrm{T}}\mathbf{M}\big(\mathbf{u}^+ - \mathbf{u}^-\big). \tag{57}$$

Elimination of $\mathbf{M}(\mathbf{u}^+ - \mathbf{u}^-)$ with the help of (19) yields

$$T^+ - T^- = \frac{1}{2}\big(\mathbf{u}^+ + \mathbf{u}^-\big)^{\mathrm{T}}\mathbf{W}\boldsymbol{\Lambda} = \frac{1}{2}\boldsymbol{\Lambda}^{\mathrm{T}}\mathbf{W}^{\mathrm{T}}\big(\mathbf{u}^+ + \mathbf{u}^-\big), \tag{58}$$

which can be rewritten by using (24) as

$$T^+ - T^- = \frac{1}{2}\boldsymbol{\Lambda}^{\mathrm{T}}\big(\boldsymbol{\gamma}^+ + \boldsymbol{\gamma}^-\big) =: W. \tag{59}$$

In this form, the energy difference may also be interpreted as the overall work W done by the n individual impact elements,

$$W = \sum_{i=1}^{n} \frac{1}{2} \boldsymbol{\Lambda}_i^{\mathrm{T}} \left(\boldsymbol{\gamma}_i^{+} + \boldsymbol{\gamma}_i^{-} \right) =: \sum_{i=1}^{n} W_i \tag{60}$$

for which W_i is the contribution of a single impact element. We define an impact problem to be *energetically consistent* iff the overall impact work W is non-positive,

$$W \leq 0. \tag{61}$$

Note that we do *not* require the individual contributions W_i to be non-positive as done by so-called energetic impact laws, which would again lead to (61) as a consequence, but would in our opinion restrict the energy transfer within the system in a too harsh way.

In order to prepare (59) for first conclusions, we state the impact work W as a function of only the impact law variables $\boldsymbol{\xi}$ and $\boldsymbol{\Lambda}$ and the impact coefficients $\boldsymbol{\varepsilon}$. To do so, we solve (28) for the relative velocities $\boldsymbol{\gamma}^{-}$ and $\boldsymbol{\gamma}^{+}$,

$$\boldsymbol{\gamma}^{-} = (\mathbf{I} + \boldsymbol{\varepsilon})^{-1} (\boldsymbol{\xi} - \mathbf{G}\boldsymbol{\Lambda}) \quad \text{and} \quad \boldsymbol{\gamma}^{+} = (\mathbf{I} + \boldsymbol{\varepsilon})^{-1} (\boldsymbol{\xi} + \boldsymbol{\varepsilon}\mathbf{G}\boldsymbol{\Lambda}) \tag{62}$$

and put the result back in (59),

$$W = \frac{1}{2} \boldsymbol{\Lambda}^{\mathrm{T}} (\mathbf{I} + \boldsymbol{\varepsilon})^{-1} \left(2\boldsymbol{\xi} - (\mathbf{I} - \boldsymbol{\varepsilon}) \mathbf{G} \boldsymbol{\Lambda} \right). \tag{63}$$

In the following subsections, conditions on the impact parameters $\boldsymbol{\varepsilon}$ are derived to ensure energetic consistency.

5.1 Isolated impact elements

In a first step, we investigate the conditions under which the individual impact elements guarantee energetic consistency when isolated from each other. This includes the special case of just and only one arbitrary impact element in the mechanical system. Mutual interference of the impact elements can formally be excluded by *not* allowing for far distance effects as in Newton's cradle [2], *and* by assuming a block-diagonal Delassus operator $\mathbf{G}$, in which only the symmetric and positive definite diagonal entries $\mathbf{G}_{ii} = \mathbf{W}_i^{\mathrm{T}} \mathbf{M}^{-1} \mathbf{W}_i$ are present, and all the off-diagonal terms $\mathbf{G}_{ik} = \mathbf{W}_i^{\mathrm{T}} \mathbf{M}^{-1} \mathbf{W}_k$ $(i \neq k)$ are equal to zero. Under this hypothesis, the contact work (63) becomes

$$W = \sum_{i=1}^{n} W_i = \sum_{i=1}^{n} \frac{1}{2} \boldsymbol{\Lambda}_i^{\mathrm{T}} (\mathbf{I}_i + \varepsilon_i \mathbf{I}_i)^{-1} \left(2\boldsymbol{\xi}_i - (\mathbf{I}_i - \varepsilon_i \mathbf{I}_i) \mathbf{G}_{ii} \boldsymbol{\Lambda}_i \right), \tag{64}$$

where $\mathbf{I}_i$ again denote the $m(i) \times m(i)$ identity matrices. We derive now the bounds on the impact coefficients that ensure energetic consistency $W = \sum_{i=1}^{n} W_i \leq 0$ under any and all possible choices of ε_i for the case of isolated impact elements. Note that the following calculation is *only* based on the work expression (64), meaning that the kinematic restrictions on the impact coefficients as derived in Sect. 4 are *not* taken into account here.

Under the above restrictions, the impact work W_i of an individual i can directly be identified from (64). We claim that

$$W_i = \frac{1}{1 + \varepsilon_i} \left(\boldsymbol{\Lambda}_i^{\mathrm{T}} \boldsymbol{\xi}_i - \frac{1}{2} (1 - \varepsilon_i) \boldsymbol{\Lambda}_i^{\mathrm{T}} \mathbf{G}_{ii} \boldsymbol{\Lambda}_i \right) \overset{!}{\leq} 0 \quad \text{for each individual } i \tag{65}$$

are n necessary conditions for $W = \sum_{i=1}^{n} W_i \leq 0$ to sufficiently hold. To see this, suppose that $W_i > 0$ for one particular i. If we are able to enforce all other impact works to be $W_k = 0$ by admissible choices of ε_k, then we have already shown contradiction. We first evaluate the inequality in (65) to get the desired restrictions on ε_i, and show afterward that $W_k = 0$ can be realized within these restrictions.

5.1.1 Case 1: $\varepsilon_i > -1$

From (65), we obtain

$$\boldsymbol{\Lambda}_i^{\mathrm{T}} \boldsymbol{\xi}_i \leq \frac{1}{2}(1-\varepsilon_i)\boldsymbol{\Lambda}_i^{\mathrm{T}} \mathbf{G}_{ii} \boldsymbol{\Lambda}_i . \tag{66}$$

It holds that $\boldsymbol{\Lambda}_i^{\mathrm{T}} \boldsymbol{\xi}_i \leq 0$ due to (18), with the most restrictive case for ε_i in (66) when $\boldsymbol{\Lambda}_i^{\mathrm{T}} \boldsymbol{\xi}_i = 0$. We therefore get

$$0 \leq \frac{1}{2}(1-\varepsilon_i)\boldsymbol{\Lambda}_i^{\mathrm{T}} \mathbf{G}_{ii} \boldsymbol{\Lambda}_i . \tag{67}$$

Further, $\mathbf{G}_{ii}$ is positive definite, i.e. $\boldsymbol{\Lambda}_i^{\mathrm{T}} \mathbf{G}_{ii} \boldsymbol{\Lambda}_i > 0, \forall \boldsymbol{\Lambda}_i \neq 0$, which yields $\varepsilon_i \leq 1$. Altogether, we obtain for Case 1

$$-1 < \varepsilon_i \leq 1 \tag{68}$$

as the restriction on the impact coefficient ε_i.

5.1.2 Case 2: $\varepsilon_i < -1$

From (65) and (18), we obtain

$$\frac{1}{2}(1-\varepsilon_i)\boldsymbol{\Lambda}_i^{\mathrm{T}} \mathbf{G}_{ii} \boldsymbol{\Lambda}_i \leq \boldsymbol{\Lambda}_i^{\mathrm{T}} \boldsymbol{\xi}_i \leq 0 \tag{69}$$

and thus $\varepsilon_i \geq 1$, which contradicts $\varepsilon_i < -1$ and therefore excludes this case.

5.1.3 Case 3: $\varepsilon_i = -1$

Because of the division by $(1+\varepsilon_i)$, Eq. (65) cannot directly be evaluated. One has to go back to (28), from which one gets

$$\boldsymbol{\xi}_i = \mathbf{G}_{ii} \boldsymbol{\Lambda}_i . \tag{70}$$

Note that $(1+\varepsilon_i)\boldsymbol{\gamma}_i^-$, i.e. the term that drives the impact, is fully eliminated now. With (70), the impact work (65) becomes

$$W_i = \frac{1}{1+\varepsilon_i}\left(\boldsymbol{\Lambda}_i^{\mathrm{T}} \mathbf{G}_{ii} \boldsymbol{\Lambda}_i - \frac{1}{2}(1-\varepsilon_i)\boldsymbol{\Lambda}_i^{\mathrm{T}} \mathbf{G}_{ii} \boldsymbol{\Lambda}_i \right) = \frac{1}{2}\boldsymbol{\Lambda}_i^{\mathrm{T}} \mathbf{G}_{ii} \boldsymbol{\Lambda}_i , \tag{71}$$

where the critical term $1/(1+\varepsilon_i)$ has now canceled out. To get the value of W_i, the impact law (17) has to be evaluated by the definition of the normal cone (1),

$$\boldsymbol{\xi}_i \in \mathscr{N}_{\mathscr{C}_i}(-\boldsymbol{\Lambda}_i) \quad \Longleftrightarrow \quad \boldsymbol{\xi}_i^{\mathrm{T}}\left(\boldsymbol{\Lambda}_i - \boldsymbol{\Lambda}_i^*\right) \leq 0 \quad \text{for } -\boldsymbol{\Lambda}_i \in \mathscr{C}_i, \quad \forall -\boldsymbol{\Lambda}_i^* \in \mathscr{C}_i . \tag{72}$$

Insertion of (70) in the right-hand side of (72) yields

$$\boldsymbol{\Lambda}_i^{\mathrm{T}} \mathbf{G}_{ii} \boldsymbol{\Lambda}_i \leq \boldsymbol{\Lambda}_i^{\mathrm{T}} \mathbf{G}_{ii} \boldsymbol{\Lambda}_i^* \quad \text{for } -\boldsymbol{\Lambda}_i \in \mathscr{C}_i, \quad \forall -\boldsymbol{\Lambda}_i^* \in \mathscr{C}_i , \tag{73}$$

Table 1 Restrictions on the impact coefficients from kinematic and isolated energetic consistency

Geometric unilateral constraint	Kinematic unilateral constraint	Bilateral constraints	Coulomb type friction
$0 \leq \varepsilon_i \leq 1$	$-1 < \varepsilon_i \leq 0$	$-1 < \varepsilon_i \leq 1$	$-1 < \varepsilon_i \leq 1$

which has to hold for *all* $-\boldsymbol{\Lambda}_i^* \in \mathscr{C}_i$. Since $0 \in \mathscr{C}_i$, we choose $\boldsymbol{\Lambda}_i^* = 0$ to obtain

$$\boldsymbol{\Lambda}_i^{\mathrm{T}} \mathbf{G}_{ii} \boldsymbol{\Lambda}_i \leq 0. \tag{74}$$

Due to the definiteness of $\mathbf{G}_{ii}$, the only possible value for $\boldsymbol{\Lambda}_i$ is $\boldsymbol{\Lambda}_i = 0$. Together with (70), (16), and (71), one finally gets

$$\boldsymbol{\Lambda}_i = 0 \quad \Longrightarrow \quad \boldsymbol{\xi}_i = 0, \quad \boldsymbol{\gamma}_i^+ = \boldsymbol{\gamma}_i^-, \quad W_i = 0. \tag{75}$$

Although not violating the inequality $W_i \leq 0$, we do not further consider $\varepsilon_i = -1$ as a meaningful impact parameter, and leave the strict inequality in (68) as it is. The value $\varepsilon_i = -1$ totally eliminates the impact element from the system and hence from any interference with other elements by not providing any impulsive force and letting the relative velocities unchanged.

5.1.4 Enforcing $W_i = 0$

As already announced, it finally has to be shown that $W_i = 0$ can be enforced for all impact elements by a proper choice of the impact parameter ε_i from the range of admissible values (68). By taking $\varepsilon_i = 1$, the impact work (65) becomes $W_i = \frac{1}{2} \boldsymbol{\Lambda}_i^{\mathrm{T}} \boldsymbol{\xi}_i$. The product $\boldsymbol{\Lambda}_i^{\mathrm{T}} \boldsymbol{\xi}_i$ is equal to zero for the geometric and kinematic unilateral constraint by (32), for the geometric and kinematic bilateral constraint by (35), and can be enforced to be equal to zero for all friction constraints by setting the friction coefficients large enough in (36), (46), (55).

We have shown that the restrictions (68) on the various impact parameters provide a sufficient condition to ensure energetic consistency if the impact elements are isolated from each other by zero-entries in the off-diagonal terms of the Delassus operator. Together with the kinematic restrictions from Figs. 6–9 of Sect. 4, one obtains the values summarized in Table 1 which are henceforth used.

Note again that we have *not* used the kinematic restrictions by deriving the bounds on ε_i from the work inequality (65). Since kinematic consistency, i.e. additional conditions on $\boldsymbol{\gamma}_i^-$ and $\boldsymbol{\gamma}_i^+$ directly affect the work expression (60), some values for ε_i in (68) are too restrictive. This can be observed for the bilateral constraints, which cannot impact by themselves as soon as $\gamma_i^+ = \gamma_i^- = 0$ is taken into account, leading to $\frac{1}{2} \Lambda_i (\gamma_i^+ + \gamma_i^-) = 0$ in (60) independent of the choice of ε_i. On the other hand, these additional values for ε_i do not bring any new physics into the system, and we therefore stick to the intervals given in Table 1.

5.2 Fully coupled systems

Energetic consistency of the individual impact elements as studied in the last section may get lost when the impact elements are coupled among each other by the off-diagonal terms in the Delassus matrix. In order to address this case, we keep the restrictions $-1 < \varepsilon_i \leq 1$

from (68) and derive additional conditions on ε_i to ensure energetic consistency also for fully coupled systems. By expanding the overall impact work W, we obtain from (63)

$$W = \boldsymbol{\Lambda}^{\mathrm{T}}(\mathbf{I}+\boldsymbol{\varepsilon})^{-1}\boldsymbol{\xi} - \frac{1}{2}\boldsymbol{\Lambda}^{\mathrm{T}}(\mathbf{I}+\boldsymbol{\varepsilon})^{-1}(\mathbf{I}-\boldsymbol{\varepsilon})\mathbf{G}\boldsymbol{\Lambda}. \tag{76}$$

By keeping the restrictions $-1 < \varepsilon_i \leq 1$ in (76), we set now

$$\boldsymbol{\delta} := (\mathbf{I}+\boldsymbol{\varepsilon})^{-1}(\mathbf{I}-\boldsymbol{\varepsilon}), \quad \boldsymbol{\delta} = \mathrm{diag}(\delta_i \mathbf{I}_i) \quad \text{with } \delta_i = \frac{1-\varepsilon_i}{1+\varepsilon_i} \geq 0 \tag{77}$$

and state energetic consistency as

$$W = \boldsymbol{\Lambda}^{\mathrm{T}}(\mathbf{I}+\boldsymbol{\varepsilon})^{-1}\boldsymbol{\xi} - \frac{1}{2}\boldsymbol{\Lambda}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{G}\boldsymbol{\Lambda} \overset{!}{\leq} 0. \tag{78}$$

With the help of (18), we obtain for the first summand

$$\boldsymbol{\Lambda}^{\mathrm{T}}(\mathbf{I}+\boldsymbol{\varepsilon})^{-1}\boldsymbol{\xi} = \sum_{i=1}^{n} \frac{1}{1+\varepsilon_i}\boldsymbol{\Lambda}_i^{\mathrm{T}}\boldsymbol{\xi}_i \leq 0. \tag{79}$$

In order to ensure (78) as a whole, it remains to find conditions under which the matrix $(\boldsymbol{\delta}\mathbf{G})$ is positive semi-definite, i.e. under which $\frac{1}{2}\boldsymbol{\Lambda}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{G}\boldsymbol{\Lambda} \geq 0$ for arbitrary $\boldsymbol{\Lambda}$. Such conditions have been derived in [16] for the cases of similar and small impact coefficients ε_i as shown in the following:

5.2.1 Similar impact coefficients

We denote by $\varepsilon_{\min}$ the smallest and by $\varepsilon_{\max}$ the largest of the impact coefficients $\varepsilon_i \in (-1, 1]$ and set

$$\delta_{\max} := \frac{1-\varepsilon_{\min}}{1+\varepsilon_{\min}}, \qquad \delta_{\min} := \frac{1-\varepsilon_{\max}}{1+\varepsilon_{\max}}. \tag{80}$$

With the abbreviation

$$\bar{\delta} := \frac{1}{2}(\delta_{\max} + \delta_{\min}), \tag{81}$$

the double of the second summand in (78) is now rewritten as

$$\boldsymbol{\Lambda}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{G}\boldsymbol{\Lambda} = \bar{\delta}\boldsymbol{\Lambda}^{\mathrm{T}}\mathbf{G}\boldsymbol{\Lambda} - \boldsymbol{\Lambda}^{\mathrm{T}}(\bar{\delta}\mathbf{I} - \boldsymbol{\delta})\mathbf{G}\boldsymbol{\Lambda}. \tag{82}$$

Since $\mathbf{G}$ is symmetric and real, it follows for the first term on the right-hand side of (82) that

$$\bar{\delta}\boldsymbol{\Lambda}^{\mathrm{T}}\mathbf{G}\boldsymbol{\Lambda} \geq \bar{\delta}\lambda_{\min}\|\boldsymbol{\Lambda}\|^2, \tag{83}$$

where $\lambda_{\min}$ denotes the smallest (real) eigenvalue of $\mathbf{G}$. For the second term on the right-hand side of (82), it holds that

$$\boldsymbol{\Lambda}^{\mathrm{T}}(\bar{\delta}\mathbf{I} - \boldsymbol{\delta})\mathbf{G}\boldsymbol{\Lambda} \leq \|\bar{\delta}\mathbf{I} - \boldsymbol{\delta}\|\|\mathbf{G}\|\|\boldsymbol{\Lambda}\|^2 = \frac{1}{2}(\delta_{\max} - \delta_{\min})\lambda_{\max}\|\boldsymbol{\Lambda}\|^2, \tag{84}$$

where $\lambda_{\max}$ is the largest (real) eigenvalue of $\mathbf{G}$, and $\|\mathbf{A}\|$ denotes the matrix norm induced by the Euclidean vector norm, i.e. $\|\mathbf{A}\| := \max_{\mathbf{x}}\{\|\mathbf{A}\mathbf{x}\| \mid \|\mathbf{x}\| = 1\}$. By the estimates (83) and (84), we conclude that the left-hand side of (82) is non-negative if

$$\bar{\delta}\lambda_{\min}\|\boldsymbol{\Lambda}\|^2 \geq \frac{1}{2}(\delta_{\max} - \delta_{\min})\lambda_{\max}\|\boldsymbol{\Lambda}\|^2 \quad \Longrightarrow \quad \boldsymbol{\Lambda}^{\mathrm{T}}(\boldsymbol{\delta}\mathbf{G})\boldsymbol{\Lambda} \geq 0 \tag{85}$$

or, after canceling $\boldsymbol{\Lambda}$ and back-substituting (81), that the inequality

$$(\delta_{\max} + \delta_{\min})\lambda_{\min} \geq (\delta_{\max} - \delta_{\min})\lambda_{\max} \quad \Longrightarrow \quad (\boldsymbol{\delta}\mathbf{G}) \text{ is PSD} \tag{86}$$

is sufficient for $(\boldsymbol{\delta}\mathbf{G})$ to be positive semi-definite. Together with (80), we conclude that similar impact coefficients ε_i in the sense that

$$\frac{\varepsilon_{\max} - \varepsilon_{\min}}{1 - \varepsilon_{\max}\varepsilon_{\min}} \equiv \frac{\delta_{\max} - \delta_{\min}}{\delta_{\max} + \delta_{\min}} \leq \frac{\lambda_{\min}}{\lambda_{\max}} \tag{87}$$

provide a sufficient condition for the impact to be energetically consistent, i.e. to fulfill the impact work inequality (78).

5.2.2 Small impact coefficients

Another sufficient condition for energetic consistency, *independent* of the one in (87), is obtained when writing the second summand in (78) as

$$\boldsymbol{\Lambda}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{G}\boldsymbol{\Lambda} = \boldsymbol{\Lambda}^{\mathrm{T}}\mathbf{G}\boldsymbol{\Lambda} - \boldsymbol{\Lambda}^{\mathrm{T}}(\mathbf{I} - \boldsymbol{\delta})\mathbf{G}\boldsymbol{\Lambda}. \tag{88}$$

As in (83), we have for the first term on the right-hand side of (88)

$$\boldsymbol{\Lambda}^{\mathrm{T}}\mathbf{G}\boldsymbol{\Lambda} \geq \lambda_{\min}\|\boldsymbol{\Lambda}\|^2. \tag{89}$$

For the second term on the right-hand side of (88), it holds that

$$\boldsymbol{\Lambda}^{\mathrm{T}}(\mathbf{I} - \boldsymbol{\delta})\mathbf{G}\boldsymbol{\Lambda} \leq \|\mathbf{I} - \boldsymbol{\delta}\|\|\mathbf{G}\|\|\boldsymbol{\Lambda}\|^2 = (1 - \delta_{\min})\lambda_{\max}\|\boldsymbol{\Lambda}\|^2. \tag{90}$$

By the estimates (89) and (90), we conclude that the left-hand side of (88) is non-negative if

$$\lambda_{\min}\|\boldsymbol{\Lambda}\|^2 \geq (1 - \delta_{\min})\lambda_{\max}\|\boldsymbol{\Lambda}\|^2 \quad \Longrightarrow \quad \boldsymbol{\Lambda}^{\mathrm{T}}(\boldsymbol{\delta}\mathbf{G})\boldsymbol{\Lambda} \geq 0 \tag{91}$$

or, after canceling $\boldsymbol{\Lambda}$ and back-substituting (80), that the inequality

$$\frac{2\varepsilon_{\max}}{1 + \varepsilon_{\max}} \equiv 1 - \delta_{\min} \leq \frac{\lambda_{\min}}{\lambda_{\max}} \tag{92}$$

is sufficient for $(\boldsymbol{\delta}\mathbf{G})$ to be positive semidefinite. In other words, impact coefficients that are small in the sense of (92) provide another sufficient condition for the impact to be energetically consistent, i.e. to fulfill the impact work inequality (78).

5.2.3 Equal impact coefficients

A special case of *similar impact coefficients* is met when they are all equal to each other with value ε,

$$\varepsilon := \varepsilon_1 = \cdots = \varepsilon_n. \tag{93}$$

In this case, we have $\varepsilon_{\max} = \varepsilon_{\min} = \varepsilon$, and (87) reduces to $0 \leq \lambda_{\min}/\lambda_{\max}$ which is always satisfied, because the eigenvalues of the positive semi-definite Delassus matrix $\mathbf{G}$ are all non-negative. Alternatively, energetic consistency may directly be verified on the second summand in (78). With

$$\delta := \delta_1 = \cdots = \delta_n \geq 0 \tag{94}$$

and $\boldsymbol{\delta} = \delta \mathbf{I}$ as a consequence, one obtains

$$\boldsymbol{\Lambda}^{\mathrm{T}} \boldsymbol{\delta} \mathbf{G} \boldsymbol{\Lambda} = \delta \boldsymbol{\Lambda}^{\mathrm{T}} \mathbf{G} \boldsymbol{\Lambda} \geq 0 \quad \forall \boldsymbol{\Lambda}. \tag{95}$$

Hence, the impact work inequality (78) is fulfilled by (79) and (95).

5.2.4 Completely elastic impact

We call an impact *completely elastic* iff all impact coefficients are equal to one, i.e. if $\varepsilon = 1$ in (93). A completely elastic impact is therefore a special case of *equal impact coefficients* which by itself is a special case of *similar impact coefficients*, and is always energetically consistent. Note, however, that completely elastic impacts exclude kinematic unilateral constraints, because their impact coefficients can NOT be set equal to one by Table 1 without violating kinematic consistency. Completely elastic impacts may be energy preserving. With $\varepsilon = 1$, one obtains $\delta = 0$ by (77), and the impact work (78) reduces with the help of (79) to

$$W = \sum_{i=1}^{n} \frac{1}{2} \boldsymbol{\Lambda}_i^{\mathrm{T}} \boldsymbol{\xi}_i \leq 0. \tag{96}$$

As already elaborated in Sect. 5.1, the product $\boldsymbol{\Lambda}_i^{\mathrm{T}} \boldsymbol{\xi}_i$ is equal to zero for the geometric and kinematic unilateral constraint by (32), and for the geometric and kinematic bilateral constraint by (35). It can be enforced to be equal to zero for all friction constraints when the friction coefficients are chosen large enough in (36), (46), (55), i.e. when the impact is processed in the stick states. Under these conditions, we obtain $W = 0$ which is conservation of kinetic energy.

5.2.5 Completely inelastic impact

We call an impact *completely inelastic* iff all impact coefficients are equal to zero, i.e. if $\varepsilon = 0$ in (93). A completely inelastic impact is therefore a special case of *equal impact coefficients* which by itself is a special case of *similar impact coefficients*, and is always energetically consistent. Moreover, a completely inelastic impact is a special case of *small impact coefficients* with $\varepsilon_{\max} = 0$, for which (92) always holds true. The completely inelastic impact is the *only* case for which energetic consistency can be guaranteed for any and all combinations of impact elements in a system, since $\varepsilon_i = 0$ is the only common value of the various impact elements listed in Table 1.

5.3 Special impact elements

In this section, we consider systems with geometric unilateral, geometric bilateral, and kinematic bilateral constraints only. In other words, kinematic unilateral constraints as well as any Coulomb type frictional constraints are excluded. For such systems, energetic consistency is assured as shown in the following: The impact work (59) can be stated as

$$W = \frac{1}{2}\boldsymbol{\Lambda}^{\mathrm{T}}\boldsymbol{\xi} + \frac{1}{2}\boldsymbol{\Lambda}^{\mathrm{T}}(\mathbf{I} - \boldsymbol{\varepsilon})\boldsymbol{\gamma}^{-}, \tag{97}$$

where we have used the left equation in (27) to eliminate the post-impact relative velocities $\boldsymbol{\gamma}^{+}$. Written as a sum over the n individual one-dimensional impact elements, Eq. (97) becomes

$$W = \sum_{i=1}^{n} \frac{1}{2}\Lambda_i \xi_i + \frac{1}{2}(1 - \varepsilon_i)\Lambda_i \gamma_i^{-}. \tag{98}$$

With $\Lambda_i \xi_i \leq 0$ by (18) and $\varepsilon_i \leq 1$ by (68), it just has to be shown that $\Lambda_i \gamma_i^{-} \leq 0$ for the contact work to be non-positive. For geometric unilateral constraints, we have $\Lambda_i \geq 0$ by (32) and $\gamma_i^{-} \leq 0$ by kinematic consistency (Fig. 6), hence $\Lambda_i \gamma_i^{-} \leq 0$. For all bilateral constraints, we have $\gamma_i^{-} = 0$ by kinematic consistency (Fig. 7), hence $\Lambda_i \gamma_i^{-} = 0$, which completes the proof. Note that this proof is *not* as immediate as the ones in Sect. 5.2, because it requires *in addition* the kinematic consistency conditions of the impact elements.

6 A slide-push mechanism as an example for violation of energetic consistency

The purpose of this section is to construct the most basic example that contradicts energetic consistency. At least two impact elements, coupled by the Delassus operator, are necessary. Otherwise, the case of isolated impact elements would apply, which has been proven in Sect. 5.1 to be energetically consistent. One of the two impact elements must be a geometric unilateral constraint to initiate the impact, because all the other impact elements can only react on impacts but not impact by themselves. The second impact element must not be another geometric unilateral constraint or any bilateral constraint, because energetic consistency has been proven for this case in Sect. 5.3. We therefore choose from the remaining impact elements the most basic one, which is the kinematic unilateral constraint. The impact parameters of the two impact elements must not be equal to each other, i.e. both equal to zero, because energetic consistency would then apply due to Sect. 5.2.

The resulting system is displayed in Fig. 10. A similar, slightly modified version of this system was originally designed in [17], sharing all the properties of the one considered here, and was introduced to the community in [10]. The system in Fig. 10 consists of two bodies with mass m, a unilateral geometric constraint (hard contact) between body 1 and the environment (wall), and a unilateral kinematic constraint (sprag clutch) between body 1

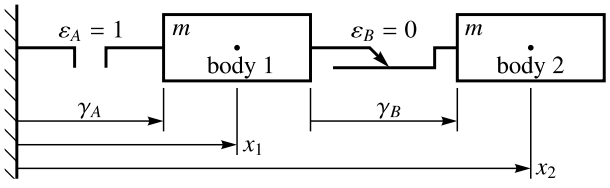

Fig. 10 The slide-push mechanism as the most basic example that contradicts energetic consistency for Newtonian impacts

and body 2. The sprag clutch is arranged such that the two bodies can move away from each other, but not approach each other. The restitution coefficient of the hard contact is $\varepsilon_A = 1$, realizing a completely elastic Newtonian impact, whereas the impact parameter of the sprag clutch is set to $\varepsilon_B = 0$. The inertial positions of the masses are denoted by x_1 and x_2, and the relative velocities in the impact elements by γ_A and γ_B.

For the pre-impact configuration, we assume that body 1 is moving with velocity v towards the wall, whereas body 2 is at rest, causing the sprag clutch to work in its unconstrained direction. The pre-impact kinetic energy is therefore $T^- = \frac{1}{2}mv^2$. At the impact time-instant, the velocity of body 1 is inverted, such that it moves afterward away from the wall with velocity v. The sprag clutch prevents body 1 from approaching body 2. The latter must therefore move with velocity v to the right, also. All together, both masses are now moving with the same velocity v to the right, the kinetic energy $T^+ = \frac{1}{2}mv^2 + \frac{1}{2}mv^2 = mv^2$ has doubled at the impact, resulting in the impact work

$$W = T^+ - T^- = \frac{1}{2}mv^2 > 0. \tag{99}$$

If the impact coefficient ε_B would have been chosen strictly less than zero but within the range given in Table 1, the result would even be worse, because body 2 would then move with an even higher post-impact velocity to the right.

In order to derive the result (99) more formally, we write down the set of Eqs. (14)–(19) componentwise with $\varepsilon_A = 1$, $\varepsilon_B = 0$ as

$$\begin{aligned}
&m\left(u_1^+ - u_1^-\right) = \Lambda_A - \Lambda_B, && m\left(u_2^+ - u_2^-\right) = \Lambda_B,\\
&\gamma_A = u_1, && \gamma_B = u_2 - u_1,\\
&\xi_A = \gamma_A^+ + \gamma_A^-, && \xi_B = \gamma_B^+,\\
&\xi_A \geq 0, \quad \Lambda_A \geq 0, \quad \xi_A \Lambda_A = 0, && \xi_B \geq 0, \quad \Lambda_B \geq 0, \quad \xi_B \Lambda_B = 0,
\end{aligned} \tag{100}$$

where we have set $u_1 := \dot{x}_1$ and $u_2 := \dot{x}_2$. The impact laws for the hard contact and the sprag clutch in the last line have been formulated as inequality-complementarity conditions as displayed in (32). According to (56) and (59), the kinetic energy T and the impact work W for this example read as

$$T = \frac{1}{2}mu_1^2 + \frac{1}{2}mu_2^2, \qquad W = \frac{1}{2}\left(\gamma_A^+ + \gamma_A^-\right)\Lambda_A + \frac{1}{2}\left(\gamma_B^+ + \gamma_B^-\right)\Lambda_B. \tag{101}$$

The pre-impact generalized velocities are given by

$$u_1^- = -v \quad \text{and} \quad u_2^- = 0, \tag{102}$$

from which we get by the second line in (100) the pre-impact relative velocities

$$\gamma_A^- = -v \quad \text{and} \quad \gamma_B^- = v. \tag{103}$$

By physical intuition, we anticipate that the impulsive forces Λ_A and Λ_B are both strictly greater than zero. This assumption, which has to be proven afterward, enforces ξ_A and ξ_B to be equal to zero by the complementarity conditions from the last line in (100),

$$\xi_A = \gamma_A^+ + \gamma_A^- = 0 \quad \text{and} \quad \xi_B = \gamma_B^+ = 0. \tag{104}$$

From (104) and (103), we obtain the post-impact relative velocities as

$$\gamma_A^+ = v \quad \text{and} \quad \gamma_B^+ = 0, \tag{105}$$

which leads to the post-impact generalized velocities

$$u_1^+ = v \quad \text{and} \quad u_2^+ = v. \tag{106}$$

By plugging (102) and (106) into the first line of (100), one finally determines the values of the impulsive forces as

$$\Lambda_A = 3mv > 0 \quad \text{and} \quad \Lambda_B = mv > 0. \tag{107}$$

They both are strictly greater than zero, which justifies our assumption. With (103), (105), (107), the impact work (101) becomes

$$W = \frac{1}{2} m v^2 \tag{108}$$

which is the same as already stated in (99). This result may also be obtained when computing the pre- and post-impact kinetic energies (101) by using (102) and (106), respectively,

$$T^- = \frac{1}{2} m v^2 \quad \text{and} \quad T^+ = m v^2, \tag{109}$$

and taking their difference, $W = T^+ - T^-$. Note that only *kinematic* relations have been used to determine in (102)–(106) the post-impact state of the system, and that no *kinetic* quantities have been involved. This is typical for Newtonian impacts. Of course may such an approach easily lead to a violation of energetic principles.

In order to finally evaluate the consistency conditions (87) and (92), we write the first two lines of (100) in matrix notation,

$$\begin{pmatrix} m & 0 \\ 0 & m \end{pmatrix} \begin{pmatrix} u_1^+ - u_1^- \\ u_2^+ - u_2^- \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Lambda_A \\ \Lambda_B \end{pmatrix}, \qquad \begin{pmatrix} \gamma_A \\ \gamma_B \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \tag{110}$$

and identify from them the matrices $\mathbf{M}$ and $\mathbf{W}$ according to (19) and (20). Accordingly, the Delassus operator $\mathbf{G}$ in (26) is calculated as

$$\mathbf{M} = \begin{pmatrix} m & 0 \\ 0 & m \end{pmatrix}, \qquad \mathbf{W} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \quad \Longrightarrow \quad \mathbf{G} = \mathbf{W}^{\mathrm{T}} \mathbf{M}^{-1} \mathbf{W} = \frac{1}{m} \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}. \tag{111}$$

Note that the off-diagonal entries of $\mathbf{G}$ are unequal to zero. The two eigenvalues of $\mathbf{G}$ are

$$\lambda_{\min} = \frac{1}{2m}(3 - \sqrt{5}), \qquad \lambda_{\max} = \frac{1}{2m}(3 + \sqrt{5}), \tag{112}$$

and their ratio used in the consistency conditions (87) and (92) is

$$\frac{\lambda_{\min}}{\lambda_{\max}} = \frac{1}{2}(7 - 3\sqrt{5}) \approx 0.146. \tag{113}$$

On the other hand, we have chosen in this examples the two impact coefficients as

$$\varepsilon_{\min} = 0, \qquad \varepsilon_{\max} = 1, \tag{114}$$

which gives for the desired expressions in (87) and (92)

$$\frac{\varepsilon_{\max} - \varepsilon_{\min}}{1 - \varepsilon_{\max}\varepsilon_{\min}} = 1, \qquad \frac{2\varepsilon_{\max}}{1 + \varepsilon_{\max}} = 1. \tag{115}$$

As expected, the consistency conditions for both, similar and small impact coefficients fail.

7 Kane's example: Frictional impact at a double pendulum

In [12], Kane reported on the paradoxical situation of an increase in kinetic energy for a double pendulum under a Newtonian impact with Coulomb friction. The same example is later found in the book [13], together with a table that shows some parameters for which such an increase can be observed. The purpose of this section is to put Kane's example into our framework and to verify that the impact elements introduced so far are able to reproduce the original results. The approach is semi-analytical, and all numerical values are rounded to four digits.

The double pendulum (Fig. 11) consists of two identical uniform rods with mass m and length l, and is hinged at the origin O of the inertial frame $(\mathbf{e}_x^I, \mathbf{e}_y^I, \mathbf{e}_z^I)$. The system is described by the generalized coordinates $\mathbf{q} = (\varphi_1 \varphi_2)^{\mathrm{T}}$ which measure the absolute angles of the two rods. The end C of the lower pendulum strikes a horizontal surface by a frictional impact. The contact angles and the pre-impact generalized velocities are as specified in Fig. 11. All data are kept in accordance with those of the original example in [13].

7.1 Pre-impact state and Delassus operator

In this section, we allocate all operators and data that are later needed to process the impact by the left equation in (28) together with the impact laws. The generalized coordinates and their derivatives are denoted by

$$\mathbf{q} = \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}, \qquad \dot{\mathbf{q}} =: \mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \tag{116}$$

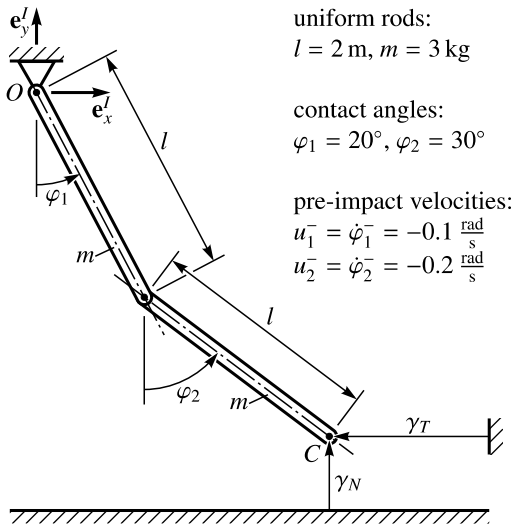

Fig. 11 Double pendulum striking the environment by a frictional impact

For the contact terms, we use the notation

$$\boldsymbol{\gamma} = \begin{pmatrix} \gamma_N \\ \gamma_T \end{pmatrix}, \qquad \boldsymbol{\xi} = \begin{pmatrix} \xi_N \\ \xi_T \end{pmatrix}, \qquad \boldsymbol{\Lambda} = \begin{pmatrix} \Lambda_N \\ \Lambda_T \end{pmatrix}, \qquad \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_N & 0 \\ 0 & \varepsilon_T \end{pmatrix}. \tag{117}$$

The relative velocities γ_T and γ_N at the contact point C are identified by Fig. 11 as the negative first and the second component of the absolute velocity ${}_I\mathbf{v}_C$ of point C, displayed in the inertial frame. The latter is obtained by direct differentiation of the associated position vector ${}_I\mathbf{r}_{OC}$ which gives

$$\begin{aligned} {}_I\mathbf{r}_{OC} &= \begin{pmatrix} l\sin\varphi_1 + l\sin\varphi_2 \\ -l\cos\varphi_1 - l\cos\varphi_2 \\ 0 \end{pmatrix}, \\ {}_I\mathbf{v}_C &= \begin{pmatrix} lu_1\cos\varphi_1 + lu_2\cos\varphi_2 \\ lu_1\sin\varphi_1 + lu_2\sin\varphi_2 \\ 0 \end{pmatrix} =: \begin{pmatrix} -\gamma_T \\ \gamma_N \\ 0 \end{pmatrix}. \end{aligned} \tag{118}$$

With $\gamma_N = \mathbf{w}_N^{\mathrm{T}}\mathbf{u}$ and $\gamma_T = \mathbf{w}_T^{\mathrm{T}}\mathbf{u}$ according to (11) for $m(i) = 1$, the generalized force directions $\mathbf{w}_N$ and $\mathbf{w}_T$ are identified from (118) as

$$\mathbf{w}_N = \begin{pmatrix} l\sin\varphi_1 \\ l\sin\varphi_2 \end{pmatrix} = \begin{pmatrix} 0.6840 \\ 1.0 \end{pmatrix} \mathrm{m}, \qquad \mathbf{w}_T = \begin{pmatrix} -l\cos\varphi_1 \\ -l\cos\varphi_2 \end{pmatrix} = \begin{pmatrix} -1.8794 \\ -1.7321 \end{pmatrix} \mathrm{m}. \tag{119}$$

The kinetic energy (56) of the double pendulum reads as

$$T = \frac{1}{2}\mathbf{u}^{\mathrm{T}}\mathbf{M}\mathbf{u} = \frac{1}{2}ml^2\left(\frac{4}{3}u_1^2 + \frac{1}{3}u_2^2 + u_1u_2\cos(\varphi_2 - \varphi_1)\right), \tag{120}$$

from which one conveniently extracts the mass matrix $\mathbf{M}$ to be

$$\mathbf{M} = ml^2\begin{pmatrix} \frac{4}{3} & \frac{1}{2}\cos(\varphi_2 - \varphi_1) \\ \frac{1}{2}\cos(\varphi_2 - \varphi_1) & \frac{1}{3} \end{pmatrix} = \begin{pmatrix} 16.0 & 5.9088 \\ 5.9088 & 4.0 \end{pmatrix} \mathrm{kg\,m^2}. \tag{121}$$

With $\mathbf{W} := (\mathbf{w}_N\,\mathbf{w}_T)$ from (119) and $\mathbf{M}$ from (121), the Delassus operator $\mathbf{G} = \mathbf{W}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{W}$ from (26) becomes

$$\mathbf{G} = \begin{pmatrix} 0.3365 & -0.5071 \\ -0.5071 & 0.8134 \end{pmatrix} \mathrm{kg}^{-1}. \tag{122}$$

The two eigenvalues of $\mathbf{G}$ and their ratio are

$$\lambda_{\min} = 0.0146\ \mathrm{kg}^{-1}, \qquad \lambda_{\max} = 1.1354\ \mathrm{kg}^{-1}, \qquad \frac{\lambda_{\min}}{\lambda_{\max}} = 0.0129, \tag{123}$$

revealing already that the frictional impact problem is ill-conditioned to a certain amount. The pre-impact velocities and pre-impact kinetic energy are according to Fig. 11 and Eqs. (20), (120)

$$\mathbf{u}^- = \begin{pmatrix} -0.1 \\ -0.2 \end{pmatrix} \frac{\mathrm{rad}}{\mathrm{s}}, \qquad \boldsymbol{\gamma}^- = \begin{pmatrix} -0.2684 \\ 0.5343 \end{pmatrix} \frac{\mathrm{m}}{\mathrm{s}}, \qquad T^- = 0.2782\ \mathrm{J}. \tag{124}$$

The impact problem is now prepared insofar as the left equation in (28), i.e.

$$\boldsymbol{\xi} = \mathbf{G}\boldsymbol{\Lambda} + (\mathbf{I} + \boldsymbol{\varepsilon})\boldsymbol{\gamma}^- \tag{125}$$

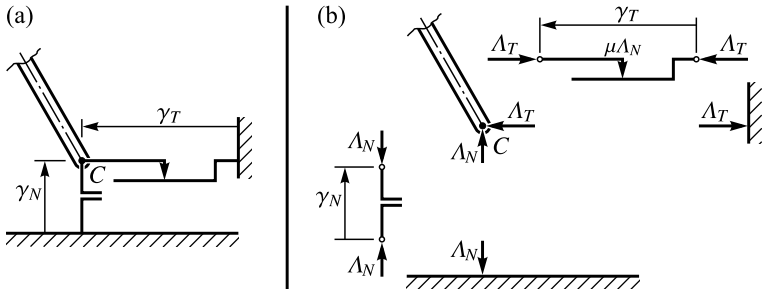

Fig. 12 Impact elements at the contact point and free body diagram

together with the impact laws can be solved for the unknowns $\boldsymbol{\xi}$ and $\boldsymbol{\Lambda}$. According to the frictional contact in Fig. 11, we choose as impact elements a geometric unilateral constraint (Fig. 6) and a one-dimensional Coulomb type friction element (Fig. 8) with the impact laws as specified in Eqs. (32) and (39), respectively,

$$-\Lambda_N \in \mathrm{Upr}(\xi_N), \qquad -\Lambda_T \in \mu\Lambda_N \,\mathrm{Sgn}(\xi_T). \tag{126}$$

These impact elements are shown in a close-up of the contact point in Fig. 12(a). Figure 12(b) displays the free body diagram associated with the impact elements, which is later used to properly determine the *signs* of the impulsive contact forces. Once the values for $\boldsymbol{\Lambda}$ and $\boldsymbol{\xi}$ have been computed from (125) and (126), the post-impact relative velocities $\boldsymbol{\gamma}^+$ can be calculated from (21), and the generalized post-impact velocities $\mathbf{u}^+$ from (19).

7.2 Post-impact stick

In [13], 2nd line of Table P14.6(b) on p. 348, an energy increase of 0.16 J with stick at the end is reported when the impact is processed with a restitution coefficient of 0.5, a static friction coefficient of 0.5, and a kinetic friction coefficient of 0.4. A restitution coefficient for the friction element is not mentioned as it is implicitly understood to be equal to zero. Since we do not use different friction coefficients for slip and stick, we take 0.5 for our friction coefficient to reproduce the above results, as stick is expected at the end of the impact. With the parameters

$$\varepsilon_N = 0.5, \qquad \varepsilon_T = 0, \qquad \mu = 0.5, \tag{127}$$

the solution of (125) and (126) is

$$\boldsymbol{\xi} = 0, \qquad \boldsymbol{\Lambda} = \begin{pmatrix} 3.4079 \\ 1.4676 \end{pmatrix} \mathrm{N\,s} \tag{128}$$

with stick at the end as expected. Note also that the stick condition $|\Lambda_T| < \mu\Lambda_N$ is fulfilled, and that the impulsive tangential force Λ_T counter-intuitively acts as depicted in Fig. 12, trying to push the lower end C of the pendulum to the *left* to bring it to stick. The post-impact velocities and kinetic energy are

$$\boldsymbol{\gamma}^+ = \begin{pmatrix} 0.1342 \\ 0 \end{pmatrix} \frac{\mathrm{m}}{\mathrm{s}}, \qquad \mathbf{u}^+ = \begin{pmatrix} -0.3346 \\ 0.3631 \end{pmatrix} \frac{\mathrm{rad}}{\mathrm{s}}, \qquad T^+ = 0.4416\,\mathrm{J}, \tag{129}$$

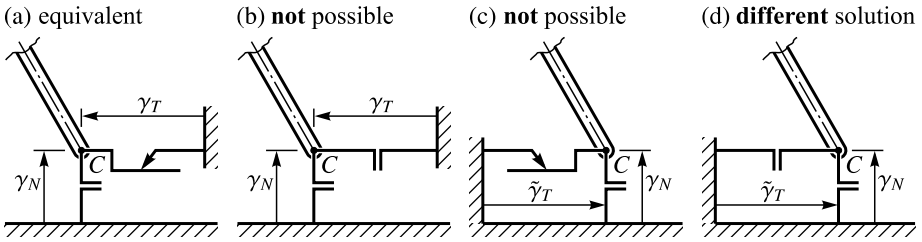

Fig. 13 Various arrangements of unilateral constraints in the tangential direction

which leads together with (124) to an energy increase of

$$W = T^+ - T^- = 0.1634\ \mathrm{J}, \tag{130}$$

as reported in [13]. So far, we were able to confirm this result, which also means that the impact laws as used in classical literature are special cases of our inequality formulation. Further, note that the impact coefficients ε_N and ε_T in (127) are by (123) neither similar nor small in the sense of (87) and (92), such that energetic consistency cannot be guaranteed.

The impact terminates with stick in the tangential direction, i.e. with a vanishing tangential relative velocity $\gamma_T^+ = 0$. One might ask whether the same post-impact state could alternatively be reached by an impact element which is more basic than the friction element, i.e. by a unilateral constraint realizing a completely inelastic impact $\varepsilon_T = 0$. All possible arrangements of geometric and kinematic unilateral constraints are depicted in Fig. 13, but only the sprag clutch in (a) leads to the desired result: The pre-impact relative velocity by (124) is kinematically consistent, $\gamma_T^- > 0$, and the impulsive force Λ_T is directed as needed, see Fig. 12(b). An evaluation of the impact law (32) for this case yields $\xi_T = 0$ and $\Lambda_T > 0$ with the very same value as in (128).

In contrast, the other three elements depicted in (b)–(d) cannot be used to reproduce the result (128): The impact contact in (b) provides an impulsive force in the right direction, but is initiated with a kinematically inconsistent pre-impact velocity $\gamma_T^- \not\leq 0$. The sprag clutch in (c) cannot provide an impulsive force in the direction which is needed, and the pre-impact velocity $\tilde{\gamma}_T^- = -\gamma_T^- \not\geq 0$ is kinematically inconsistent, also. The impact contact in (d) is kinematically admissible $\tilde{\gamma}_T^- = -\gamma_T^- \leq 0$, but its impulsive force cannot act in the appropriate direction. Processing the impact with the impact element (d) would lead to a solution with $\Lambda_T = 0$, which is different from (128). Another reason why the impact contacts (b) and (d) cannot generate the solution (128) is the proof of energetic consistency in Sect. 5.3 which would contradict the energy gain (130).

In summary, we have seen that the only duty of the friction element in Fig. 12 for the given parameter set is to realize the stick condition $\gamma_T^+ = 0$ at the end of the impact. We have also seen that the same can be accomplished by the sprag clutch in Fig. 13(a), leading together with the normal impact contact to the *same* post-impact velocities $\boldsymbol{\gamma}^+$ as in (129). For the given parameter set, the original and the modified double pendulum have therefore to be regarded as *equivalent* for the impact. Furthermore, we have obtained by the modified pendulum a mechanical system with two degrees of freedom, one impact contact, and one sprag clutch, just as in the slide-push mechanism from Sect. 6. In other words, all the ingredients needed to produce the energetic inconsistency are present, and one might ask whether the mechanism leading to the energy increase is again the one observed in Sect. 6. This question will positively be answered in Sect. 8.

7.3 Post-impact slip

A second case which leads to an energetic inconsistency is reported in the fourth line of Table P14.6(b) in [13]. In contrast to the second line, the impact terminates with slip when the restitution coefficient is chosen as 0.7, the static friction coefficient as 0.51, and the kinetic friction coefficient as 0.5. In order to verify this result, we process here the impact with the parameters

$$\varepsilon_N = 0.7, \qquad \varepsilon_T = 0, \qquad \mu = 0.5, \tag{131}$$

which yields by (125) and (126)

$$\boldsymbol{\xi} = \begin{pmatrix} 0 \\ -0.0177 \end{pmatrix} \frac{\mathrm{m}}{\mathrm{s}}, \qquad \boldsymbol{\Lambda} = \begin{pmatrix} 5.4995 \\ 2.7498 \end{pmatrix} \mathrm{N\,s}. \tag{132}$$

Note that slip is valid by $\xi_T \neq 0$ for the post-impact state, and that the impulsive forces fulfill the slip condition $\Lambda_T = \mu \Lambda_N$. Again, as in Sect. 7.2, the value of the tangential impulsive force Λ_T is greater than zero, trying to push the lower end C of the pendulum to the *left*. In contrast to the previous section, Λ_T is now *not* big enough to realize the stick condition at the end, so that point C is expected to slide to the *right*. This is confirmed by the value of γ_T^+ in the post-impact state

$$\boldsymbol{\gamma}^+ = \begin{pmatrix} 0.1879 \\ -0.0177 \end{pmatrix} \frac{\mathrm{m}}{\mathrm{s}}, \qquad \mathbf{u}^+ = \begin{pmatrix} -0.4430 \\ 0.4909 \end{pmatrix} \frac{\mathrm{rad}}{\mathrm{s}}, \qquad T^+ = 0.7671\ \mathrm{J}, \tag{133}$$

which is indeed less than zero, indicating slip reversal. Furthermore, the energy increase as computed from (124) and (133) is

$$W = T^+ - T^- = 0.4889\ \mathrm{J} \tag{134}$$

and coincides with the value given in [13]. Energetic consistency can again not be guaranteed by (87) or (92), as the impact coefficients ε_N and ε_T are neither similar nor small by (123).

One could again try to reproduce the results (132), (133) by one of the unilateral constraints in Fig. 13; this time with a tangential impact coefficient $\varepsilon_T \neq 0$ to achieve the non-vanishing post-impact slip value for γ_T^+. This, however, is *not* possible: Configurations (b)–(d) fail because of the very same reasons as before, but also configuration (a) cannot be used, as $\gamma_T^+ < 0$ according to (133) is kinematically inconsistent for the sprag clutch and would require by $\gamma_T^+ = -\varepsilon_T \gamma_T^-$ an impact coefficient $\varepsilon_T = 0.0331$ *greater* than zero, violating the admissible range for kinematic unilateral constraints given in Table 1.

As all the unilateral constraints in Fig. 13 fail, the most basic representation of the tangential impact is by an impressed impulsive force $\Lambda_T = \mu \Lambda_N$, which is depicted in Fig. 14(a). This configuration is the simplest but not the only one representing the impact. In Figs. 14(b)–(e), alternative configurations with increasing complexity are shown. They all lead to the same post-impact state (132) when processed with the data (131). For the following interpretation of Fig. 14, one has to keep in mind that the tangential impact has been computed with $\varepsilon_T = 0$, meaning by (12) that $\xi_T \equiv \gamma_T^+$ and causing the sign function (126) to relate the tangential impulsive force Λ_T and the post-impact relative velocity γ_T^+ to each other.

As shown in Fig. 14(a), the tangential impulsive force $\mu \Lambda_N$ acts to the left while point C is moving after the impact to the right. This motion is not hindered by the additional sprag clutch in Fig. 14(b), as its right end is not connected to the environment. Furthermore,

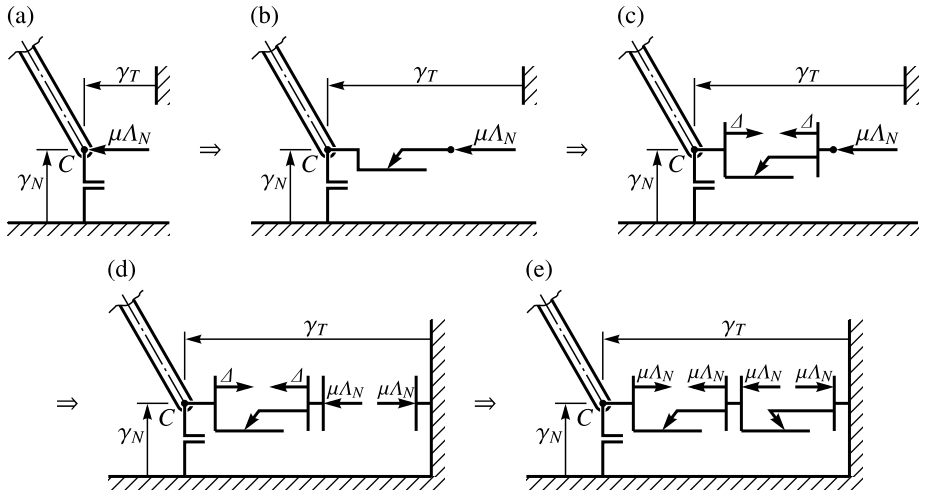

Fig. 14 From an impressed impulsive tangential force $\Lambda_T = \mu\Lambda_N$ for slip (**a**) to the full one-dimensional friction element (**e**)

the sprag clutch is arranged such that it acts in its blocked direction. It therefore allows to transmit $\mu\Lambda_N$ from its right end to the point C of the pendulum. Since the sprag clutch is not moving in its unconstrained direction after the impact, it can internally be pre-loaded in its constrained direction by a pair of impulsive forces Δ of any size as shown in Fig. 14(c). Figure 14(d) is the same as Fig. 14(c) with the only difference that it has been completed by the reaction of $\mu\Lambda_N$ which acts against the environment. As a couple, the two impulsive forces $\mu\Lambda_N$ can be regarded as another impact element, arranged in series relative to the Δ-sprag clutch element on its left. After the impact, point C is moving to the right, causing the $\mu\Lambda_N$-couple to approach each other. Another sprag clutch is added in Fig. 14(e), which allows for this motion by again not changing the impact event. In a last step, the values for Δ are chosen as $\mu\Lambda_N$.

The resulting element in Fig. 14(e) does not only lead to the post-impact state (132), (133) when processed with the data set (131), but is a full representation of the one-dimensional Coulomb type friction element as it will be shown in Sect. 9. When used on the data (127) for post-impact stick, it again provides the solution (128), (129) by the following mechanism: The pre-impact velocity $\gamma_T^- > 0$ is kinematically consistent due the left sprag clutch. The impulsive force $\Lambda_T > 0$ according to (128) causes the left sprag clutch to block anyway. Since $|\Lambda_T| < \mu\Lambda_N$ by (128), Λ_T cannot overcome the pre-load $\mu\Lambda_N$ in the right sprag clutch, which is consequently not moving after the impact as well, and the whole element terminates with stick, i.e. $\gamma_T^+ = 0$.

8 Analysis of Kane's example by an impact equivalent system

In this section, it will be shown that the mechanism leading to the energy gain in Kane's example for both, the stick and the slip case, is indeed the same as the one already observed in Sect. 6. In order to draw this conclusion, we try to find a system which is equivalent to Kane's example but shows the topology of the slide-push mechanism depicted in Fig. 10. We call two systems *impact equivalent* if they are driven by the *same* impact equation (28),

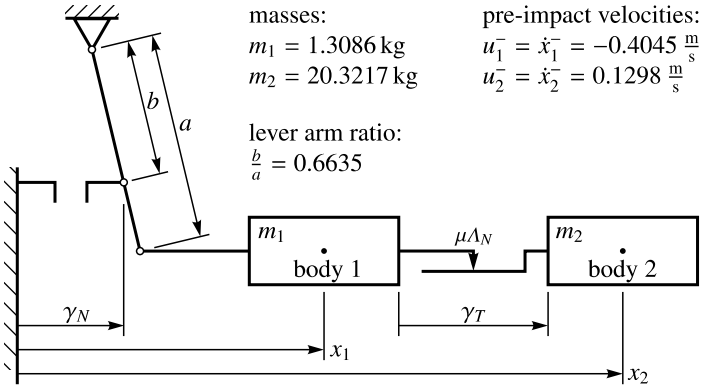

Fig. 15 Impact equivalent system

i.e. if they agree in their Delassus matrices $\mathbf{G}$, and if they contain the *same* impact elements in the *same* order with respect to the entries in $\boldsymbol{\xi}$ or $\boldsymbol{\Lambda}$.

Figure 15 shows the chosen configuration. In contrast to the slide-push mechanism from Fig. 10, the masses of the two bodies are kept different, and a massless lever arm has been added to adjust the velocity γ_N in the impact contact, leading to the kinematic relations

$$\gamma_N = \frac{b}{a} u_1, \quad \gamma_T = u_2 - u_1. \tag{135}$$

With the help of the mass matrix and the generalized force directions, the Delassus operator $\mathbf{G}$ is determined as

$$\begin{aligned} \mathbf{M} &= \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix}, \\ \mathbf{W} &= \begin{pmatrix} \frac{b}{a} & -1 \\ 0 & 1 \end{pmatrix} \quad \Longrightarrow \quad \mathbf{G} = \mathbf{W}^{\mathrm{T}} \mathbf{M}^{-1} \mathbf{W} = \begin{pmatrix} \frac{b^2}{m_1 a^2} & -\frac{b}{m_1 a} \\ -\frac{b}{m_1 a} & \frac{m_1+m_2}{m_1 m_2} \end{pmatrix}. \end{aligned} \tag{136}$$

The values for the two masses and the lever arm ratio are now chosen such that $\mathbf{G}$ in (136) becomes identical to the Delassus operator of the double pendulum (122). This yields

$$m_1 = 1.3086\ \mathrm{kg}, \qquad m_2 = 20.3217\ \mathrm{kg}, \qquad \frac{b}{a} = 0.6635. \tag{137}$$

Note that m_2 is much bigger than m_1, which is a consequence of the poor condition of $\mathbf{G}$. Slight changes in the velocity of m_2 will heavily affect the kinetic energy of the system. If the system is now initiated with the pre-impact relative velocities $\boldsymbol{\gamma}^-$ from (124) and processed with the data (127) and (131), it will terminate the impact with the values $\boldsymbol{\xi}$, $\boldsymbol{\Lambda}$, $\boldsymbol{\gamma}^+$ according to (128), (129), and (132), (133), respectively. For completeness, one can calculate from (135) the associated values for the generalized velocities $\mathbf{u}$ which gives

$$\mathbf{u}^- = \begin{pmatrix} -0.4045 \\ 0.1298 \end{pmatrix} \frac{\mathrm{m}}{\mathrm{s}}, \qquad \mathbf{u}^+_{\mathrm{stick}} = \begin{pmatrix} 0.2023 \\ 0.2023 \end{pmatrix} \frac{\mathrm{m}}{\mathrm{s}}, \qquad \mathbf{u}^+_{\mathrm{slip}} = \begin{pmatrix} 0.2832 \\ 0.2655 \end{pmatrix} \frac{\mathrm{m}}{\mathrm{s}}. \tag{138}$$

The pre-impact velocity $\mathbf{u}^-$ shows that body 1 is moving towards the wall as in the slide-push mechanism of Sect. 6, whereas body 2 is not at rest but is moving slowly to the right.

For post-impact stick, we derived that the friction element in Kane's double pendulum can be replaced by a sprag clutch; see Fig. 13(a). Equivalently, we may replace the friction element in the impact equivalent system (Fig. 15) by a sprag clutch as well, leading to the same arrangement as in the slide-push mechanism of Fig. 10. At the collision, the pre-impact velocity u_1^- of body 1 is reversed and diminished according to (127) by a factor $\varepsilon_N = 0.5$, leading to the post-impact value $u_{1,\text{stick}}^+$ in (138). This value is *larger* than the pre-impact velocity u_2^- of body 2, i.e. body 1 would be approaching body 2 after the impact. This, however, is not possible because of the sprag clutch. Body 2 is therefore pushed to the right by increasing its velocity from u_2^- to $u_{2,\text{stick}}^+ = u_{1,\text{stick}}^+$, which causes the energy gain (130) by the same mechanism as already observed in Sect. 6. Note that the lever arm is only needed to adjust the impulsive force Λ_N to the one in (128).

For post-impact slip, a single sprag clutch cannot substitute the friction element, as the impulsive force Λ_T is too big to withstand the internal pre-loads. However, the mechanism leading to the energy gain (134) is very similar to the one in the stick case: At the collision, the pre-impact velocity u_1^- of body 1 is now reversed and diminished by a factor $\varepsilon_N = 0.7$ according to (131), leading to the post-impact value $u_{1,\text{slip}}^+$ in (138). This value is *again* bigger than the pre-impact velocity u_2^- of body 2, and body 1 would again approach body 2 after the impact against the friction element. In this case, however, the impulsive force Λ_T needed to bring body 2 to the same post-impact velocity as body 1 would be bigger than $\mu\Lambda_N$, which means that the stick condition cannot be realized. Nevertheless, body 2 is again pushed to the right, this time by $\Lambda_T = \mu\Lambda_N$, which increases its velocity from u_2^- to $u_{2,\text{slip}}^+$ and causes the energy gain.

9 Decomposition of the one-dimensional friction element

In this final section, we want to show that sprag clutches are indeed parts of the one-dimensional friction element, leading to the configuration already depicted in Fig. 14(e). The logic behind this approach is that new force elements in mechanics are usually designed by combining existing ones together. There are only two ways to perform this combination, either by arranging the existing elements in parallel to each other, or in series. One-dimensional friction is a quite complicated set-valued force element on velocity level. The most basic set-valued force element on velocity level is the kinematic unilateral constraint. It should therefore be possible to somehow break down the friction element into kinematic unilateral constraints.

It can easily be demonstrated [5] that a parallel connection of sprag clutches does not lead to the desired friction element. Instead, a serial connection is needed, which is more complicated from the mechanical point of view, because inner variables have to be introduced. The resulting element is depicted in the upper right part of Fig. 16. For impact-free motion, this arrangement has been introduced for a first time in [5], but without any further explanation. It can be found in this form also in [7] and [6]. The entire element consists of a serial connection of two sub-elements. The relative velocity of the entire element is denoted by γ_T, the relative velocities of the right and left sub-element by γ_R and γ_L, where $\gamma_T = \gamma_R - \gamma_L$. Each sub-element consists of a sprag clutch and a pair of constant forces $\mu\Lambda_N$ which are connected in parallel to each other within the sub-element. Note that the two sprag clutches work in opposite directions and that the pre-loads $\mu\Lambda_N$ act in the directions in which the individual sprag clutches block.

In order to understand the mechanics of this device, consider Λ not as impulsive forces but as ordinary, finite forces as they occur for impact-free motion. With this in mind, now

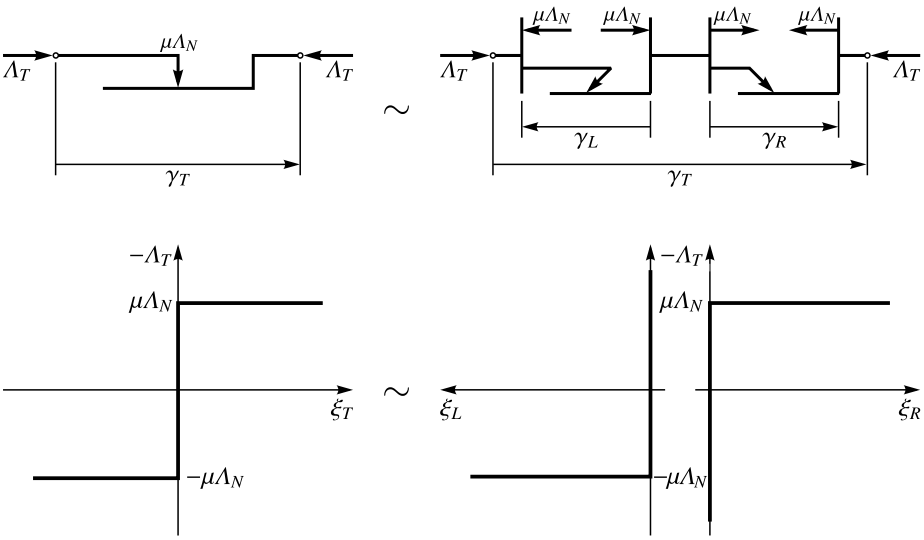

Fig. 16 Friction element and equivalent arrangement of kinematic unilateral constraints

slightly push the two endpoints of the entire force element together. The following will happen: The right sprag clutch will not move, because it is actuated in the blocked direction. The left sprag clutch will not move, too, as long as the applied force is less than the pre-load $\mu\Lambda_N$. This is the stick case. To make the left sprag clutch moving, an applied force of magnitude $\mu\Lambda_N$ is needed to compensate the pre-load. This is the slip case. One can even see the dissipation mechanism from this force element: Realize an oscillatory motion of the two endpoints relative to each other. Every time the endpoints are pushed together, one has to work against the left pair of forces $\mu\Lambda_N$, while the end points of the left sprag clutch are moving towards each other. Every time that the endpoints are pulled apart from each other, one has to work against the right pair of forces $\mu\Lambda_N$, while the endpoints of the right sprag clutch are moving away from each other. In this way, the entire work needed to perform this motion is finally stored in the two sprag clutches and cannot be regained, unless the sprag clutches are "mechanically destroyed".

There is a one-to-one correspondence between the components in the above device and the decomposition of the set-valued sign function into unilateral primitives as presented in Fig. 4 and Eq. (9). The impact law for the friction element (39) is expressed by a set-valued sign function,

$$-\Lambda_T \in \mu\Lambda_N \operatorname{Sgn}(\xi_T), \tag{139}$$

for which the decomposition into unilateral primitives (9) yields

$$-\Lambda_T \in \operatorname{Upr}(\xi_R) + \mu\Lambda_N, \quad -\Lambda_T \in -\operatorname{Upr}(\xi_L) - \mu\Lambda_N, \quad \xi_T = \xi_R - \xi_L. \tag{140}$$

The graphs of the inclusions (139) and (140) are depicted in Fig. 16 below their associated impact elements. In order to relate the various relative velocities γ_A with $A \in \{T, R, L\}$ in the impact elements to the kinematic variables ξ_A in the graphs and in (140), keep in mind that ξ_A is by (16) nothing else than always the same combination $\xi_A = \gamma_A^+ + \varepsilon_A \gamma_A^-$ of the associated relative velocities. In particular for impact free motion, one has by $\gamma_A^+ =$

$\gamma_A^- =: \gamma_A$ the simple relationship $\xi_A = (1 + \varepsilon_A)\gamma_A \sim \gamma_A$. The sprag clutch under the pre-load $\mu\Lambda_N$ in the right sub-element of the decomposed friction element is mathematically represented by an Upr-inclusion which is shifted upward by the value of the pre-load. This yields together with $\xi_R \sim \gamma_R$ the graph depicted below of it, which is represented by the first inclusion in (140). The left sub-element consists again of a sprag clutch under the same pre-load, but this time arranged in the opposite direction. The associated Upr-graph has therefore to be mirrored on the ξ_L-axis and has to be shifted downward, leading to the second inclusion in (140). Note that the ordinates of both sub-graphs carry the *same* entity $-\Lambda_T$, taking into account that the force in a serial connection of sub-elements has to be the same. The last equation in (140) introduces one more kinematic entity when passing from the original friction element to the decomposed one, which is precisely the additional internal variable needed for serial arrangements. As a final remark, the decomposition in (140) is the one to be used when set-valued sign functions are to be treated as linear complementarity problems in standard form [11].

10 Conclusion

A common mathematical structure has been established for various impact elements by equipping them with kinematic restitution coefficients and stating them as normal cone inclusions. This formulation contains as special cases the classical Newtonian impact law and Coulomb's law as usually applied for frictional impacts. Additional features among others are tangential restitution, automated impact state evaluation, and multi-impact capabilities. Hard contacts, sprag clutches, bilateral constraints on displacement and velocity level, one- and two-dimensional Coulomb friction for the isotropic and anisotropic case, and even Coulomb–Contensou friction elements can be combined by this approach in any order within a finite-dimensional Lagrangian system.

The mathematical structure set up so far has to be considered as groundwork for the actual purpose of this contribution, which is to analyze the conditions under which energetic consistency can be guaranteed, and to identify the mechanisms leading to a violation of it. The main results can be summarized as follows: The classical bounds on the restitution coefficients have been justified by requiring energetic consistency for the individual impact elements when they are isolated from each other. For coupled systems, energetic consistency has been related to the condition number of the Delassus operator and can be guaranteed if the impact coefficients are small or similar in a certain sense. For systems with only geometric unilateral, geometric bilateral, and kinematic bilateral constraints, it has been shown that Newtonian impacts are always energetically consistent.

From the consistency proofs, it has been possible to identify a combination of impact elements that is critical for energy increase. As the most basic configuration, one geometric unilateral constraint and one kinematic unilateral constraint, equipped with impact coefficients different from each other, have been arranged in a system with two degrees of freedom. The resulting slide-push mechanism has been shown to produce kinetic energy by a stunningly simple mode of operation, which reveals the true weakness of the Newton's kinematic collision law when combined with sprag clutches. The same mechanism has been observed to be responsible for the energy gain in Kane's example. This conclusion has been drawn by designing an impact equivalent system, again showing the topology of the slide-push mechanism, and the fact that the Coulomb friction element contains hidden sprag clutches. Most interestingly, slip reversal as commonly claimed to be the mechanism responsible for the energy gain, is not necessarily needed, because the phenomenon already occurs for post-impact stick.

Despite of the many deficiencies that Newton type impact laws are known to have, one must not forget about the big number of application problems that have successfully been treated. In the article at hand, Newton type impact laws have been formulated within the concepts of non-smooth dynamics. Normal cone inclusions as used to formulate the impact laws provide a sound mathematical structure, from which a variety of mechanical statements can easily be derived. Newton type impact laws behave in particular critical when the various degrees of freedom of a mechanical system are strongly coupled among each other, and when the impact coefficients have to be of rather different size. For completely inelastic impacts, however, Newtonian impact laws seem to nearly work perfect. One strategy to take advance of this situation is to keep the completely inelastic Newtonian impact as the initial phase of an impact, but to treat the restitution phase by another concept. This will naturally lead to Poisson impacts that are treated in the second part of this contribution.

References

1. Acary, V., Brogliato, B.: Numerical Methods for Nonsmooth Dynamical Systems. Applications in Mechanics and Electronics. Lecture Notes in Applied and Computational Mechanics, vol. 35. Springer, Berlin (2008)
2. Aeberhard, U.: Geometrische Behandlung idealer Stösse. Diss. ETH No. 17695, Zurich (2008)
3. Aeberhard, U., Glocker, Ch.: Energy considerations for excited perfect collisions. In: van Campen, D.H., Lazurko, M.D., van den Oever, W.P.J.M. (eds.) Proc. 5th EUROMECH Nonlinear Oscillations Conference, pp. 422–431. Eindhoven Univ. of Technology, Eindhoven (2005)
4. Glocker, Ch.: Dynamik von Starrkörpersystemen mit Reibung und Stößen. VDI-Fortschrittberichte Mechanik/Bruchmechanik, Reihe 18, Nr. 182. VDI-Verlag, Düsseldorf (1995)
5. Glocker, Ch.: Decomposition of scalar force interactions. In: Pfeiffer, F., Glocker, Ch. (eds.) IUTAM Symposium on Unilateral Multibody Contacts, pp. 15–24. Kluwer, Dordrecht (1999)
6. Glocker, Ch.: Scalar force potentials in rigid multibody systems. In: Pfeiffer, F., Glocker, Ch. (eds.) Multibody Dynamics with Unilateral Contacts. CISM Courses and Lectures, vol. 421, pp. 69–146. Springer, Wien (2000)
7. Glocker, Ch.: Set-Valued Force Laws: Dynamics of Non-Smooth Systems. Lecture Notes in Applied Mechanics, vol. 1. Springer, Berlin (2001)
8. Glocker, Ch.: An introduction to impacts. In: Haslinger, J., Stavroulakis, G. (eds.) Nonsmooth Mechanics of Solids. CISM Courses and Lectures, vol. 485, pp. 45–102. Springer, Vienna (2006)
9. Glocker, Ch.: Reduction techniques for distributed set-valued force laws. In: Baniotopoulos, C.C. (ed.) Nonsmooth/Nonconvex Mechanics with Applications in Engineering, pp. 173–180. Editions Ziti, Thessaloniki (2006)
10. Glocker, Ch.: A geometric interpretation of energetic inconsistency in frictional impact laws. In: Del Piero, G., Maceri, F. (eds.) Abstracts of the Sixth Meeting on Unilateral Problems in Structural Analysis, Castello Maniace, Isola di Ortigia, Syracuse (2007)
11. Glocker, Ch., Studer, C.: Formulation and preparation for numerical evaluation of linear complementarity systems in dynamics. Multibody Syst. Dyn. **13**, 447–463 (2005)
12. Kane, T.R.: A dynamics puzzle. Stanford Mechanics Alumni Club Newsletter **6** (1983)
13. Kane, T.R., Levinson, D.A.: Dynamics: Theory and Applications. McGraw Hill, New York (1985)
14. Leine, R.I., Glocker, Ch.: A set-valued force law for spatial Coulomb–Contensou friction. Eur. J. Mech. A, Solids **22**, 193–216 (2003)
15. Leine, R.I., Nijmeijer, H.: Dynamics and Bifurcations of Non-Smooth Mechanical Systems. Lecture Notes in Applied and Computational Mechanics, vol. 18. Springer, Berlin (2004)
16. Leine, R.I., van de Wouw, N.: Stability and Convergence of Mechanical Systems with Unilateral Constraints. Lecture Notes in Applied and Computational Mechanics, vol. 36. Springer, Berlin (2008)
17. Möller, M.: Consistent integrators for non-smooth dynamical systems. Diss. ETH No. 19715, Zurich (2011)
18. Moreau, J.J.: Unilateral contact and dry friction in finite freedom dynamics. In: Moreau, J.J., Panagiotopoulos, P.D. (eds.) Non-Smooth Mechanics and Applications. CISM Courses and Lectures, vol. 302, pp. 1–82. Springer, Wien (1988)
19. Moreau, J.J.: Bounded variation in time. In: Moreau, J.J., Panagiotopoulos, P.D., Strang, G. (eds.) Topics in Nonsmooth Mechanics, pp. 1–74. Birkhäuser, Basel (1988)

20. Moreau, J.J.: Some numerical methods in multibody dynamics: application to granular materials. Eur. J. Mech. A, Solids **13**(4), 93–114 (1994)
21. Payr, M., Glocker, Ch.: Oblique frictional impact of a bar: analysis and comparison of different impact laws. Nonlinear Dyn. **41**(4), 361–383 (2005)
22. Pfeiffer, F., Glocker, Ch.: Multibody Dynamics with Unilateral Contacts. Wiley, New York (1996)
23. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1972)
24. Studer, C.: Numerics of Unilateral Contacts and Friction. Modeling and Numerical Time Integration in Non-Smooth Dynamics. Lecture Notes in Applied and Computational Mechanics, vol. 47. Springer, Berlin (2009)