A Compliant Visco-Plastic Particle Contact Model Based on Differential Variational Inequalities [☆]

A. Tasora[a], M. Anitescu[b], S. Negrini[c], D. Negrut[d]

[a]*Università degli Studi di Parma, Dipartimento di Ingegneria Industriale, 43100 Parma, Italy, tasora@ied.unipr.it*
[b]*Mathematics and Computer Science Division, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, IL 60439, USA, anitescu@mcs.anl.gov*
[c]*Politecnico di Milano, Dept. of Mechanical Engineering, V.La Masa 1, 20126 Milan, Italy, silvia.negrini@mail.polimi.it*
[d]*University of Wisconsin - Madison, Department of Mechanical Engineering, 1513 University Avenue, Madison, WI 53706, USA, negrut@cae.wisc.edu*

Abstract

This work describes an approach to simulate contacts between three-dimensional shapes with compliance and damping using the framework of the differential variational inequality theory. Within the context of nonsmooth dynamics, we introduce an extension to the classical set-valued model for frictional contacts between rigid bodies, allowing contacts to experience local compliance, viscosity, and plasticization. Different types of yield surfaces can be defined for various types of contact, a versatile approach that contains the classic dry Coulomb friction as a special case. The resulting problem is a differential variational inequality that can be solved, at each integration time step, as a variational inequality over a convex set.

Keywords: Variational inequalities, contacts, plasticity, friction

1. Introduction

Numerical methods for simulating systems that feature unilateral contacts with friction are critical in many fields of applied mechanics. Real-life problems include cases with few contacts, as in keylock mechanisms and

clock escapements, as well as many-body dynamics applications involving millions of contacts encountered when analyzing earth-moving machines operating on granular soils. When a detailed analysis of the contact is required, finite-element methods can be used to model the contact surfaces. However, since this approach is prohibitively compute intensive when handling complex many-body problems, it is most often abandoned for pointwise contact models based on a limited number of parameters. In this category, discrete element method (DEM) approaches [1], drawing on a regularization of the frictional contact forces, require very small integration timesteps. Alternatively, one can adopt a method based on nonsmooth mechanics [2], leveraging a more robust differential variational inequality (DVI) integration scheme [3, 4, 5].

Being able to deal efficiently with multibody problems with large amounts of objects, DVI methods have been used to simulate granular assemblies and bulk materials, for example, in [6, 5, 7]; but in most cases a simple Coulomb contact model is used. Such a model is defined by a single friction coefficient that neglects the effect of rolling friction, cohesion, compliance, and so on. In this paper we extend the DVI contact model beyond the single friction coefficient, yet retaining the efficiency of a formalism based on set-valued functions. In previous work we described a DVI model that targeted only the case of infinitely stiff frictional contacts [8]: that model had certain limitations when the rigid body assumption needed to be relaxed. Recent extensions to the DVI framework, encompassing compliance, are discussed in [9, 10]. In our work here, we extend the model by introducing compliance $C_i \in \mathbb{R}^{3 \times 3}$ and damping $R_i \in \mathbb{R}^{3 \times 3}$ matrices associated with each contact event in the simulation and by defining yield surfaces Υ_i for optional plastic flow in contacts that could also account for phenomena such as cohesion.

The proposed approach is similar to that in [11], where the authors handled mostly mesh-based surfaces; in our model we also added plasticization as an optional feature.

2. Background

In this section we introduce concepts and notations in convex analysis that will be used in the rest of the article. More details can be found in [12].

A set $\mathcal{K} \in \mathbb{R}^n$ is a n-dimensional *cone* if, for all $\mathbf{x} \in \mathcal{K}$, we have that $\beta \mathbf{x} \in \mathcal{K}$ for all $\beta \in \mathbb{R}^+$. As for other sets, cones can be optionally *convex*,

2

closed, or *compact*. A cone is said to be *pointed* or *salient* if it satisfies $\mathcal{K} \cap -\mathcal{K} = \{\emptyset\}$.

A set $\mathcal{K}$ in a generic real vector space equipped with an inner product has a *dual cone* $\mathcal{K}^*$ that is always convex regardless of the fact that $\mathcal{K}$ is a cone too, or convex too. It is expressed as

$$\mathcal{K}^* = \left\{ \mathbf{y} \in \mathbb{R}^n : \langle \mathbf{y}, \mathbf{x} \rangle \geq 0 \quad \forall \mathbf{x} \in \mathcal{K} \right\}. \tag{1}$$

The *polar cone*, strictly related to the dual cone of a set, is defined as

$$\mathcal{K}^\circ = \left\{ \mathbf{y} \in \mathbb{R}^n : \langle \mathbf{y}, \mathbf{x} \rangle \leq 0 \quad \forall \mathbf{x} \in \mathcal{K} \right\} = -\mathcal{K}^*. \tag{2}$$

The *normal cone* and the *tangent cone* to a closed convex set $\mathcal{K}$ at the point $\mathbf{x} \in \mathcal{K}$ are closed and convex and are respectively defined as

$$\mathcal{N}_\mathcal{K}(\mathbf{x}) = \{ \mathbf{y} \in \mathbb{R}^n : \langle \mathbf{y}, \mathbf{x} - \mathbf{z} \rangle \geq 0, \forall \mathbf{z} \in \mathcal{K} \} \tag{3}$$

$$\mathcal{T}_\mathcal{K}(\mathbf{x}) = \mathrm{cl}\{ \beta(\mathbf{y} - \mathbf{x}) : \mathbf{y} \in \mathcal{K}, \beta \in \mathbb{R}^+ \} = \mathcal{N}_\mathcal{K}(\mathbf{x})^\circ. \tag{4}$$

Note that if $\mathbf{x}$ is an interior point of $\mathcal{K}$, it is always $\mathcal{N}_\mathcal{K}(\mathbf{x}) = \{\emptyset\}$.

The *recession cone* or *horizon cone* of $\mathcal{K}$ is defined as

$$\mathcal{K}_v^\infty = \left\{ \lim_{k \to \infty} t_k \mathbf{x}_k : t_k \downarrow 0, \mathbf{x}_k \in \mathcal{K}_v \right\}, \tag{5}$$

and it can be verified that $\mathcal{K}_v^\infty = \{\emptyset\}$ if $\mathcal{K}$ is bounded.

3. Differential Variational Inequalities

Classical approaches based on the time integration of ordinary differential equations (ODEs) or differential-algebraic equations (DAEs) enforce friction and unilateral contacts by means of regularization, leading to sharp constraints and friction forces being approximated as stiff force fields. Regularization approaches in ODEs and DAEs place a heavy computational burden on the numerical integrator, which must handle highly nonlinear forces. This in turn leads to extremely short integration timesteps required to maintain numerical stability. Compared with ODEs and DAEs, DVIs have only recently gained widespread acceptance in the computational dynamics community. DVIs can be used to describe mechanical models that are subject to nonsmooth events and set-valued constraints, such as those caused by sticking and sliding friction. Additionally, they can deal with nonsmooth

3

phenomena directly, resulting in greater efficiency, better numerical stability, and enhanced robustness.

Since the seminal works by J. J. Moreau [13, 14] on measure differential inclusions (MDIs), other authors have contributed to the topic of nonsmooth mechanics, developing both theoretical and practical issues of great interest in the field of mechanical simulations; see, for instance, [3, 15, 16, 17, 2, 18].

The theory of DVI provides a generalization of DAE problems with discontinuities and a set-valued right-hand side; in this context, MDIs are special cases of DVIs [19]. Convergence results, existence, and uniqueness of solutions can be found in [16, 19]. The formulation of DVIs exploits and extends the theory of variational inequalities (VI), a fertile research topic in mathematics that was shown to be consequential in application fields such as continuum mechanics, tribology, game theory, and economy [20].

A generic VI is a problem of the type

$$\mathbf{x} \in \mathbb{K} \quad : \quad \langle F(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \geq 0 \quad \forall \mathbf{y} \in \mathbb{K}, \tag{6}$$

with $\mathbb{K}$ closed and convex and continuous $F(\mathbf{x}) : \mathbb{K} \to \mathbb{R}^n$. We call $\mathsf{SOL}(\mathbb{K}, F)$ the solution of problem (6). Then, a DVI is defined as the problem of finding the absolutely continuous function $\mathbf{x}$ on $[0, T]$:

$$\frac{d\mathbf{x}}{dt} = f(t, \mathbf{x}, \mathbf{u}) \tag{7}$$

$$\mathbf{u} \in \mathsf{SOL}\left(\mathbb{K}, F(t, \mathbf{x}(t), \cdot)\right), \tag{8}$$

along with boundary conditions $\Xi(\mathbf{x}(0), \mathbf{x}(T)) = 0$.

Esistence and uniqueness of $\mathbf{u}$ in term (8) are assured, respectively, in the case of coercive and monotone $F(\cdot)$.

A special case is when $\mathbb{K} = \mathbb{R}^m$ and Eq. (8) becomes a constraint function of the type $F(t, \mathbf{x}(t), \mathbf{u}(t)) = 0$: in this case, the DVI becomes a classic DAE, as in mechanical systems with only smooth, bilateral constraints.

For our purposes, in nonsmooth frictional problems, $\mathbf{x}$ is the state of the system: $\mathbf{x} = \{\mathbf{v}^T, \mathbf{q}^T\}^T$. We must allow discontinuities in speeds to accommodate impacts, if any. Moreover, it is known that there are problems such as the Painlevé paradoxes that cannot be solved without this assumption [21]. This extension leads to *measure* differential inclusions, where we assume $\mathbf{v}$ to be a (discontinuous) function of bounded variation; that is, $\bigvee_a^b \mathbf{v}$ is finite in $[a, b] \subset [0, T]$, with vector Borel measure $\nu(dt) = d\mathbf{v}(t)$ inducing the Riemann-Stieltjes integrals $\int \phi(t)\nu(dt)$ for any continuous $\phi(t)$. This

4

allows impulsive events, given the singular decomposition of measures $\nu = \mu_s + j\lambda_0$, where μ_s is a signed vector Borel singular measure with Lebesgue-null support, λ_0 is a Lebesgue measure on $\mathbb{R}$, and j is L^1.

In general, the MDI for the mechanical problem is expressed as

$$\frac{d\mathbf{x}(t)}{dt} = \left\{\begin{matrix} d\mathbf{v}(t)/dt \\ d\mathbf{q}(t)/dt \end{matrix}\right\} \in \left\{\begin{matrix} \mathcal{K}_v(t, \mathbf{v}, \mathbf{q}) \\ \mathbf{v}(t) \end{matrix}\right\} = \mathcal{K}(\mathbf{x}, t), \tag{9}$$

where the nontrivial part is the MDI

$$\frac{d\mathbf{v}(t)}{dt} \in \mathcal{K}_v(t, \mathbf{v}, \mathbf{q}), \tag{10}$$

whose weak solution, according to the Stewart generalization of Moreau's definition, is defined as

$$\frac{\int \phi(t)\nu(dt)}{\int \phi(t)dt} \in \overline{co} \bigcup_{\tau:\phi\tau\neq 0} \mathcal{K}_v(\tau), \tag{11}$$

when $\mathcal{K}_v(t, \mathbf{v}, \mathbf{q})$ is a set-valued map with closed convex values and with closed graph. Under the assumption of the pointedness of the recession cone $\mathcal{K}_v^\infty$, that is when $\mathcal{K}_v^\infty \cap -\mathcal{K}_v^\infty = \emptyset$, it has been demonstrated [16] that the weak solution is equivalent to the strong solution of (10).

Under the above assumptions, most timestepping numerical methods for the solution of (9) generate approximating sequences of discrete solutions $\mathbf{v}_k$ and $\mathbf{q}_k$ with $k \in \mathbb{N}$ and timesteps h, exhibiting convergence $\mathbf{q}_k(\cdot) \to \mathbf{q}(\cdot)$ with $h \downarrow 0$ uniformly, $\mathbf{v}_k(\cdot) \to \mathbf{v}(\cdot)$ with $h \downarrow 0$ pointwise, and $\nu_k(\cdot) \overset{*}{\rightharpoonup} \nu(\cdot)$ in weak* sense. Weak* convergence for sequences of measures means that $\int \phi(t)\nu_k(dt) \to \int \phi(t)\nu(dt)$, that is, $\int \phi(t)\nu(dt) = \lim_{k\to\infty} \int \phi(t)\nu_k(dt)$ for any continuous bounded $\phi(t)$ [22]. From a practical point of view this means that, in terms of accelerations and speeds, the outputs of DVI methods are not smooth functions but, rather, generalized functions, or *distributions*. In particular, the computed reactions γ in constraints and contacts appear as sequences of impulses rather than continuous force functions; yet it is guaranteed that $\gamma_k \overset{*}{\rightharpoonup} \gamma$ in weak* sense.[1]

---

[1]For plotting reasons one can transform the impulse distribution into "smoothed" force graphs by dividing measures by timesteps. Note that in the absence of discontinuities this would give the same plot of reaction forces as in a classical ODE.

4. The DVI Model

We assume that the system state is defined by a vector of generalized co-ordinates $\mathbf{q} \in \mathbb{R}^{m_q}$ and a vector of generalized speeds $\mathbf{v} \in \mathbb{R}^{m_v}$. In multibody dynamics it often happens that $m_q > m_v$; for instance, we represent the rotations of rigid bodies in three-dimensional space with unimodular quaternions $\epsilon \in \mathbb{H}_1$ in order to avoid singularities in the parametrization of $\mathsf{SO}(\mathbb{R}, 3)$. Regardless, it is straightforward to define a linear map $\dot{\mathbf{q}} = \Gamma(\mathbf{q})\mathbf{v}$ if $\mathbf{q}$ is needed.

We also introduce generalized force fields $\mathbf{f}_e(\mathbf{q}, \mathbf{v}, t)$ and gyroscopic forces $\mathbf{f}_c(\mathbf{q}, \mathbf{v})$, giving a total force field $\mathbf{f}_t(\mathbf{q}, \mathbf{v}, t) \in \mathbb{R}^{m_v}$.

The inertial properties of the system are represented by the mass matrix $M(\mathbf{q}) \in \mathbb{R}^{m_v \times m_v}$, assumed positive definite, usually block-diagonal for the rigid body case only.

Bilateral constraints, if any, are introduced by using a set $\mathcal{G}_B$ of n_b scalar constraint equations, assumed differentiable everywhere:

$$\Psi^i(\mathbf{q}, t) = 0, \quad i \in \mathcal{G}_\mathcal{B}. \tag{12}$$

We introduce $\nabla_q \Psi^i = [\partial \Psi^i / \partial \mathbf{q}]^T$ and $\nabla \Psi^{iT} = \nabla_q \Psi^{iT} \Gamma(\mathbf{q})$ to express the constraint (12) at the velocity level after differentiation:

$$\frac{d\Psi^i(\mathbf{q}, t)}{dt} = \nabla \Psi^{iT} \mathbf{v} + \frac{\partial \Psi^i}{\partial t} = 0, \quad i \in \mathcal{G}_\mathcal{B}. \tag{13}$$

The set of unilateral contacts is denoted by $\mathcal{G}_A$. We assume that for each unilateral contact there is a contact distance function $\Phi(\mathbf{q})$ that is continuously differentiable; this leads to the assumption that a contact normal can be defined for each contact point, as well as a couple of contact points aligned along this normal. For each contact $i \in \mathcal{G}_A$, we introduce the orthogonal space generator $\mathbf{D}^i_{\gamma_n} = \nabla_q \Phi^{iT} \Gamma(\mathbf{q})$ and the two tangent space generators $\mathbf{D}^i_{\gamma_u}, \mathbf{D}^i_{\gamma_w}$, with $\mathbf{D}^i_{\gamma_n}, \mathbf{D}^i_{\gamma_u}, \mathbf{D}^i_{\gamma_w} \in \mathbb{R}^{m_v}$; for details about their formulation see, for instance, [23].

For each contact point i, we consider a normal force $\widehat{\gamma}^i_n$ and two tangential forces $\widehat{\gamma}^i_u$ and $\widehat{\gamma}^i_w$. Also, we introduce the normal contact displacement y^i_n and the two tangential displacements y^i_u and y^i_w. For brevity, we will write

$$\widehat{\gamma}^i_\mathcal{A} = \{\widehat{\gamma}^i_n, \widehat{\gamma}^i_u, \widehat{\gamma}^i_w\}^T \tag{14}$$

$$\mathbf{y}^i = \{y^i_n, y^i_u, y^i_w\}^T \tag{15}$$

$$D^i = \left[\mathbf{D}^i_{\gamma_n} | \mathbf{D}^i_{\gamma_u} | \mathbf{D}^i_{\gamma_w}\right]. \tag{16}$$

In the case of only bilateral constraints, the Lagrange multipliers $\widehat{\gamma}_{\mathcal{B}}^{i}$ are associated with the reaction forces in constraints, and the equations of motion lead to an index 3 set of DAEs:

$$\dot{\mathbf{q}} = \Gamma(\mathbf{q})\mathbf{v} \tag{17}$$

$$M(\mathbf{q})\frac{d\mathbf{v}}{dt} = \sum_{i \in \mathcal{G}_{\mathcal{B}}} \widehat{\gamma}_{\mathcal{B}}^{i} \nabla \Psi^{i} + \mathbf{f}_{t}(t, \mathbf{q}, \mathbf{v}) \tag{18}$$

$$\Psi^{i}(\mathbf{q}, t) = 0, \quad i \in \mathcal{G}_{\mathcal{B}}. \tag{19}$$

Introducing also unilateral contacts where contact forces are subject to VI, one gets the following DVI:

$$\dot{\mathbf{q}} = \Gamma(\mathbf{q})\mathbf{v} \tag{20}$$

$$M(\mathbf{q})\frac{d\mathbf{v}}{dt} = \sum_{i \in \mathcal{G}_{\mathcal{B}}} \widehat{\gamma}_{\mathcal{B}}^{i} \nabla \Psi^{i} + \sum_{i \in \mathcal{G}_{\mathcal{A}}} \widehat{\gamma}_{\mathcal{A}}^{i} D^{i} + \mathbf{f}_{t}(t, \mathbf{q}, \mathbf{v}) \tag{21}$$

$$\Psi^{i}(\mathbf{q}, t) \in \emptyset, \quad i \in \mathcal{G}_{\mathcal{B}} \tag{22}$$

$$\widehat{\gamma}_{\mathcal{A}}^{i} \in \mathsf{SOL}\left(\Upsilon^{i}, F(t, \mathbf{q}(t), \mathbf{v}(t), \cdot)\right), \quad i \in \mathcal{G}_{\mathcal{A}}. \tag{23}$$

This represents a DVI problem with $\widehat{\gamma}_{\mathcal{A}}^{i}$ and $\widehat{\gamma}_{\mathcal{B}}^{i}$ representing $\mathbf{u}$ of Eqs. (7) and (8). Likewise, given that Eq. (22) is a special case of a VI over an unbounded set $\mathbb{R}^{n_{b}}$, one has that

$$\widehat{\gamma}_{\mathcal{B}}^{i} \in \mathsf{SOL}\left(\mathbb{R}^{n_{b}}, F(t, \mathbf{q}(t), \mathbf{v}(t), \cdot)\right), \quad i \in \mathcal{G}_{\mathcal{B}}.$$

By defining proper sets Υ^{i} and functions $F(t, \mathbf{q}(t), \mathbf{v}(t), \cdot)$ in the VIs of the formulation above, one can develop different constitutive laws for contact points, ranging from rigid to compliant and plastic behavior.

5. On Multibody Contacts

In the context of multibody mechanics, interactions are usually modeled by sets of contact rules acting at the point of contact. The model discussed in [4, 8], although efficient, is based on rigid frictional contacts; this assumption is too restrictive in case we need to simulate, for instance, granular materials that exploit cohesive phenomena and local elastoplastic deformations of the particles such as when simulating earth-moving machines operating on soft soils. On the other hand, ODE methods have difficulty representing rigid or almost-rigid contacts.
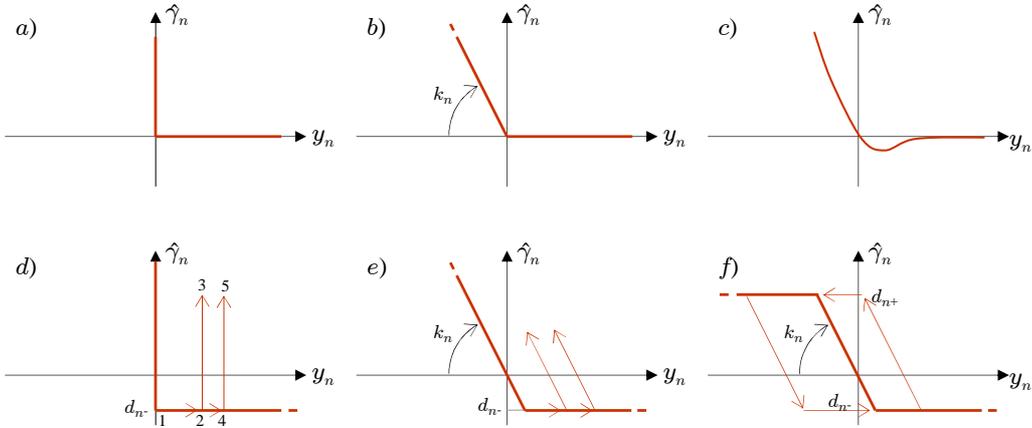
Figure 1: Basic constitutive relations for normal reaction.

This motivates our effort to identify a versatile contact model that can accommodate effects such as cohesion, friction, compliance and plasticization and that can be expressed as a set-valued function that fits well into our DVI setting. By changing parameters, such a model must be able to encompass the special case of dry friction between rigid bodies.

5.1. Compliance

The theories of Hertz, Mindlin, Cattaneo and Deresiewicz are often used as the basis for many contact models for locally compliant bodies [24]. In general, these models embed elements such as springs, dashpots, and sliders. The slider accounts for friction. Moreover, the stiffness of the spring may be hysteretic and/or nonlinear. A practical and often used model is the so-called Hertz-Mindlin linear damped spring model, which uses two Kelvin-Voigt spring dashpot systems in normal and tangential directions.

As an example, restricted to the normal component $\widehat{\gamma}_n$ of the contact force and the contact distance y_n, Figure 1 shows various types of constitutive laws $y_n \mapsto \widehat{\gamma}_n$.

Figure 1-a shows that the DVI approach previously used in [8] is a set-valued function where, for $y_n = 0$, $\widehat{\gamma}_n \in \mathbb{R}^+$, and no penetration is allowed; this is equivalent to the complementarity constraint of the Signorini contact condition: $y_n \geq 0, \widehat{\gamma}_n \geq 0, \langle y_n, \widehat{\gamma}_n \rangle = 0$.

Conversely, Figure 1-b refers to a typical DEM contact and shows that penetration is allowed depending on a finite stiffness. Introducing nonlinear stiffness such as in Figure 1-c does not require much complication in DEM

simulation codes. However, the higher the tangent stiffness, the more difficult the DEM integration is. Hence, the case of $k \uparrow \infty$ to approach the DVI rigid contact of Figure 1-a) is out of the question in DEM.

In the following sections, however, we show how to extend the original DVI formulation in order to take into account contact compliance as in Figure 1-b yet still allowing the original case of infinite stiffness in Figure 1-a. Hence, the proposed DVI method has the advantages of both DVI and DEM methods.

5.2. Cohesion

Cohesion between surfaces is caused by different physical mechanisms, such as liquid bridges between particles and chemical bonds. Such complex physical phenomena can significantly affect the macroscopic behavior of granular assemblies of particles.

At atomic scales, interactions belong to the classes of bonds, van der Waals and electrostatic forces between molecules, that can be modeled, for example, by using the method of the attractive-repulsive potential of Lennard-Jones. We remark that these intermolecular interactions can be predominant when studying very small particles such as in the mechanics of powders and cannot be neglected.

Since one cannot always simulate the smallest details of molecular dynamics, various models have been proposed in order to capture the cumulative effect of molecular forces over an entire contact point, for example, the Johnson-Kendall-Roberts (JKR) or the Derjaguin-Muller-Toporov (DMT) models. In general, these short-range interactions are modeled as nonlinear forces acting between two particles such as in the example of Figure 1-c, with a negative part in the fourth quadrant.

Difficulties in handling cohesive forces can happen either when they show high tangent stiffness or when they exhibit a dissipative nature, as in the case of the irreversible breaking of bonds. Difficulties of the first kind lead to stiff yet smooth problems that can be solved with ODEs and implicit integration; however, the second kind of problem, given the nonsmooth nature of those phenomena, can fit well in the broader DVI context. As such, these dissipative phenomena can be considered a special case of plasticity, a case discussed in the following section.

5.3. Plasticity

When dissipative phenomena take place, the contact force might depend on the load history, which is a well-covered topic in the theory of plasticity. Here we use the term plasticity in a broad sense, embracing either the classical mechanism of material flow in continua, or other micro- and mesoscopic phenomena that cannot be described by linear/nonlinear elasticity alone, such as void collapse, crushing, or changes in chemical properties.

Figure 1-*d* shows how the original DVI contact model can be modified to handle plasticization. For example, by exercising a pulling action, the contact can withstand a cohesion up to 1, where the limit d_{n-} is met; then a displacement (a plastic *flow*) with constant cohesion will follow; when one applies again a pushing action, the force will go from 2 to 3, and so on, following further cycles as in points 4 and 5. This simple plasticization can be more versatile if nonzero compliance is used, as in Figure 1-*e*. Optionally, a compression plasticization limit can also be used, as in Figure 1-*f*.

Described here is a DVI contact model that can handle the case of Figure 1-*f*, while retaining the ability of handling all other cases. Hardening and softening plasticity [25] can also be characterized by introducing a few other parameters. The introduction of tangential effects, such as friction, can be accommodated in a general DVI setting described next.

6. Elasto-Plastic DVI Contact Model

In this context, we assume that the contact force $\widehat{\gamma}_{\mathcal{A}}^i = \{\widehat{\gamma}_n^i, \widehat{\gamma}_u^i, \widehat{\gamma}_w^i\}^T$ must satisfy a convex-set inclusion:

$$\widehat{\gamma}_{\mathcal{A}}^i \in \widehat{\Upsilon}^i, \tag{24}$$

where $\widehat{\Upsilon}^i$ is the *yield surface* of the contact point. Yield surfaces are used also in continuum mechanics to define possible values of the strain tensor, for example, the Mohr-Coulomb and the Drucker-Prager yield surfaces. A single yield surface cannot exactly describe what happens if two particles with different plastic yields come into contact. However, here we assume a unique yield surface as a sufficient way to characterize microscale failure phenomena in the contact as a whole, with a reasonable number of parameters that can fit into simulations with millions of particles. The parameters for the shape of the yield surface can be fitted by using experimental data, as typically done when using a DEM approach [26].

The plastic component $\mathbf{y}_P^i$ of the displacement (plastic flow) and the elastic component $\mathbf{y}_E^i$ are related as

$$\mathbf{y}^i = \mathbf{y}_E^i + \mathbf{y}_P^i . \tag{25}$$

Given that the displacement of the contact is $\mathbf{y}^i$, we introduce the following assumptions.

Assumption 1. *Associated plastic flow. The increment to the plastic displacement is normal to the yield surface:*

$$\dot{\mathbf{y}}_P^i \in -\mathcal{N}_{\widehat{\Upsilon}^i}(\widehat{\gamma}_{\mathcal{A}}^i). \tag{26}$$

This imposes the condition that the derivative of the plastic component of the displacement is orthogonal to the yield surface when $\widehat{\gamma}_{\mathcal{A}}^i$ is on the boundary of the yield surface, a hypothesis that in the context of continuum mechanics is called associated plasticity. Although this assumption can be limiting in some cases discussed later, it can make the problem more tractable from a numerical point of view. We remark that, following the definition of normal cones, the plastic flow is null when $\widehat{\gamma}_{\mathcal{A}}^i$ is in the interior of Υ^i.

Along with the inclusion of Eqs. (24) and (26), given $\mathbf{y}_E^i = \mathbf{y}^i - \mathbf{y}_P^i$ for Eq.(25), we introduce the complete elastoplastic contact model:

$$\widehat{\gamma}_{\mathcal{A}}^i = -K^i \left(\mathbf{y}^i - \mathbf{y}_P^i \right) \tag{27}$$

$$\dot{\mathbf{y}}_P^i \in -\mathcal{N}_{\widehat{\Upsilon}^i}(\widehat{\gamma}_{\mathcal{A}}^i) \quad ; \quad \widehat{\gamma}_{\mathcal{A}}^i \in \widehat{\Upsilon}^i. \tag{28}$$

In Eq. (27) we use a stiffness matrix $K^i \in \mathbb{R}^{3 \times 3}$. A common case is when K^i is a diagonal matrix having the normal and the two tangential stiffness values on the diagonal. The minus sign reflects the convention that positive reactions $\widehat{\gamma}_{\mathcal{A}}$ are associated with pushing contact.

By differentiation, assuming K^i constant in time, Eq. (27) becomes

$$\dot{\widehat{\gamma}}_{\mathcal{A}}^i = -K^i \left(\dot{\mathbf{y}}^i - \dot{\mathbf{y}}_P^i \right) .$$

Given that $\dot{\mathbf{y}}^i = D^{i^T} \mathbf{v}$, one can discretize in time with timestep $h = t^{l+1} - t^l$, hence getting

$$\frac{\widehat{\gamma}_{\mathcal{A}}^{i,l+1} - \widehat{\gamma}_{\mathcal{A}}^{i,l}}{h} = -K^i \left(D^{i^T} \mathbf{v}^{l+1} - \dot{\mathbf{y}}_P^i \right) .$$

After algebraic manipulations, one obtains

$$\dot{\mathbf{y}}_P^i = D^{i^T}\mathbf{v}^{l+1} + \left(hK^i\right)^{-1}\widehat{\gamma}_{\mathcal{A}}^{i,l+1} - \left(hK^i\right)^{-1}\widehat{\gamma}_{\mathcal{A}}^{i,l} \quad \in -\mathcal{N}_{\widehat{\Upsilon}^i}(\widehat{\gamma}_{\mathcal{A}}^i).$$

The MDI approach requires working with impulses rather than with forces. Therefore, we use the notation $h\widehat{\gamma} = \gamma$ and obtain

$$\dot{\mathbf{y}}_P^i = D^{i^T}\mathbf{v}^{l+1} + \left(h^2K^i\right)^{-1}\gamma_{\mathcal{A}}^{i,l+1} - \left(h^2K^i\right)^{-1}\gamma_{\mathcal{A}}^{i,l} \quad \in -\mathcal{N}_{\Upsilon^i}(\gamma_{\mathcal{A}}^i). \qquad (29)$$

In the MDI representation we used the cone $\Upsilon = h\widehat{\Upsilon}$. Incidentally, it is easy to show that

$$\mathcal{N}_{\widehat{\Upsilon}^i}(\widehat{\gamma}_{\mathcal{A}}^i) = \mathcal{N}_{\Upsilon^i}(\gamma_{\mathcal{A}}^i).$$

Equation (29) demonstrates that a time integration method must be aware of the impulse at the previous step $\gamma_{\mathcal{A}}^{i,l}$. However, from Eq. (27), it follows that $\widehat{\gamma}_{\mathcal{A}}^{i,l} = -K^i\left(\mathbf{y}^{i,l} - \mathbf{y}_P^{i,l}\right)$, so $(hK^i)^{-1}\gamma_{\mathcal{A}}^{i,l} = -(\mathbf{y}^{i,l} - \mathbf{y}_P^{i,l})$. Using this in Eq. (29) yields

$$\dot{\mathbf{y}}_P^i = D^{i^T}\mathbf{v}^{l+1} + \left(h^2K^i\right)^{-1}\gamma_{\mathcal{A}}^{i,l+1} - \frac{1}{h}\left(\mathbf{y}^{i,l} - \mathbf{y}_P^{i,l}\right) \quad \in -\mathcal{N}_{\Upsilon^i}(\gamma_{\mathcal{A}}^i), \qquad (30)$$

which requires knowledge of $\mathbf{y}_P^{i,l}$, from the previous timestep. In a timestepping scheme, this means that one must update

$$\mathbf{y}_P^{i,l+1} = \mathbf{y}_P^{i,l} + h\dot{\mathbf{y}}_P^i .$$

Introducing the notation

$$E^i = -\left(h^2K^i\right)^{-1} \qquad (31)$$

$$\mathbf{c}^i = -\frac{1}{h}\left(\mathbf{y}^{i,l} - \mathbf{y}_P^{i,l}\right) \qquad (32)$$

leads to

$$\dot{\mathbf{y}}_P^i = D^{i^T}\mathbf{v}^{l+1} - E^i\widehat{\gamma}_{\mathcal{A}}^{i,l+1} - \mathbf{c}^i \quad \in -\mathcal{N}_{\Upsilon^i}(\widehat{\gamma}_{\mathcal{A}}^i) . \qquad (33)$$

Given that

$$M\frac{d\mathbf{v}}{dt} = \sum_{i \in \mathcal{G}_{\mathcal{A}}} D^i\widehat{\gamma}_{\mathcal{A}}^i + \mathbf{f}(\mathbf{q}, \mathbf{v}, t) ,$$

one can discretize in timesteps and write in terms of (vector signed) measures:

$$M\mathbf{v}^{l+1} = M\mathbf{v}^l + \sum_{i \in \mathcal{G}_{\mathcal{A}}} D^i\gamma_{\mathcal{A}}^{i,l+1} + h\mathbf{f}(\mathbf{q}, \mathbf{v}, t) .$$

12

That is,

$$\mathbf{v}^{l+1} = \mathbf{v}^l + M^{-1} \sum_{i \in \mathcal{G}_\mathcal{A}} D^i \gamma_\mathcal{A}^{i,l+1} + hM^{-1}\mathbf{f}(\mathbf{q}, \mathbf{v}, t) . \tag{34}$$

Next, we define

$$\gamma_\mathcal{E} = \left\{ \gamma_\mathcal{A}^{1^T}, \gamma_\mathcal{A}^{2^T}, \ldots \right\}^T \tag{35}$$

$$D_\mathcal{E} = \left[D^1 | D^2 | \ldots \right] \tag{36}$$

$$\mathbf{c} = \left\{ \mathbf{c}^{1^T}, \mathbf{c}^{2^T}, \ldots \right\}^T . \tag{37}$$

Moreover, we define the set Υ as the Cartesian product of all yield surfaces:

$$\Upsilon = \underset{i \in \mathcal{G}_\mathcal{A}}{\times} \Upsilon^i . \tag{38}$$

Substituting Eq. (34) into Eq. (30) yields the following inclusion:

$$\dot{\mathbf{y}}_P = D_\mathcal{E}^T M D_\mathcal{E} \gamma_\mathcal{E}^{l+1} + D_\mathcal{E}^T \left(\mathbf{v}^l + hM^{-1}\mathbf{f}(\mathbf{q}, \mathbf{v}, t) \right) - E_\mathcal{E} \gamma_\mathcal{E}^{l+1} - \mathbf{c} \quad \in -\mathcal{N}_\Upsilon(\widehat{\gamma}_\mathcal{E}) .$$

That is, by rearranging terms, we get

$$\dot{\mathbf{y}}_P = \left[D_\mathcal{E}^T M D_\mathcal{E} - E_\mathcal{E} \right] \gamma_\mathcal{E}^{l+1} + D_\mathcal{E}^T \left(\mathbf{v}^l + hM^{-1}\mathbf{f}(\mathbf{q}, \mathbf{v}, t) \right) - \mathbf{c} \quad \in -\mathcal{N}_\Upsilon(\widehat{\gamma}_\mathcal{E}) .$$

Then, by introducing the vector $\mathbf{r}$ and the square matrix N,

$$N \equiv \left[D_\mathcal{E}^T M D_\mathcal{E} - E_\mathcal{E} \right] \tag{39}$$

$$\mathbf{r} \equiv D_\mathcal{E}^T \left(\mathbf{v}^l + hM^{-1}\mathbf{f}(\mathbf{q}, \mathbf{v}, t) \right) - \mathbf{c} , \tag{40}$$

we can write the entire problem as a VI that must be solved at each integration time step:

$$N\gamma_\mathcal{E}^{l+1} + \mathbf{r} \in -\mathcal{N}_\Upsilon(\gamma_\mathcal{E}) ; \quad \gamma_\mathcal{E}^{l+1} \in \Upsilon . \tag{41}$$

Another way to write Eq.(41) is as a classical $\mathrm{VI}(F(\gamma_\mathcal{E}^{l+1}), \Upsilon)$ with affine $F(\gamma_\mathcal{E}^{l+1}) = N\gamma_\mathcal{E}^{l+1} + \mathbf{r}$:

$$\boxed{\gamma_\mathcal{E}^{l+1} \in \Upsilon : \quad \left\langle N\gamma_\mathcal{E}^{l+1} + \mathbf{r}, \mathbf{z} - \gamma_\mathcal{E}^{l+1} \right\rangle \geq 0 \quad \forall \mathbf{z} \in \Upsilon} \tag{42}$$

13

This type of VI can be seen also as a QP with convex conic constraints. In fact, given a functional $f(\gamma_{\mathcal{E}})$, $f \in C^1(\Upsilon)$, it is known that

$$\min_{\gamma_{\mathcal{E}} \in \Upsilon} f(\gamma_{\mathcal{E}}) \quad \Leftrightarrow \quad \gamma_{\mathcal{E}} \in \Upsilon, \ \mathrm{grad} f(\gamma_{\mathcal{E}}) \in -\mathcal{N}_{\Upsilon}(\gamma_{\mathcal{E}}) \ . \tag{43}$$

Thus, Eq. (41) represents the first order optimality condition for the following QP:

$$\min_{\gamma_{\mathcal{E}} \in \Upsilon} \frac{1}{2} \gamma_{\mathcal{E}}^T N \gamma_{\mathcal{E}} + \gamma_{\mathcal{E}}^T \mathbf{r} \ . \tag{44}$$

Also, when Υ itself is a cone, Eq. (41) is equivalent to a cone complementarity problem (CCP).

7. Visco-Elasto-Plastic DVI Contact Model

We can extend the previous contact model to the case where damping is also taken into account. Again we introduce a yield surface $\widehat{\Upsilon}^i$ for $\widehat{\gamma}_{\mathcal{A}}^i \in \widehat{\Upsilon}^i$, and we make the assumption of associate flow $\dot{\mathbf{y}}_P^i \in -\mathcal{N}_{\widehat{\Upsilon}^i}(\widehat{\gamma}_{\mathcal{A}}^i)$.

When we add the damping $R^i \in \mathbb{R}^{3 \times 3}$ for each contact point, the DVI constitutive model for the visco-elasto-plastic contact becomes

$$\widehat{\gamma}_{\mathcal{A}}^i = -K^i \left(\mathbf{y}^i - \mathbf{y}_P^i \right) - R^i \left(\dot{\mathbf{y}}^i - \dot{\mathbf{y}}_P^i \right) \tag{45}$$

$$\dot{\mathbf{y}}_P^i \in -\mathcal{N}_{\widehat{\Upsilon}^i}(\widehat{\gamma}_{\mathcal{A}}^i) \quad ; \quad \widehat{\gamma}_{\mathcal{A}}^i \in \widehat{\Upsilon}^i \ . \tag{46}$$

By differentiating Eq. (45) and by using finite differences with timestep h, one gets

$$\frac{\widehat{\gamma}_{\mathcal{A}}^{i,l+1} - \widehat{\gamma}_{\mathcal{A}}^{i,l}}{h} = -K^i \dot{\mathbf{y}}^i + K^i \dot{\mathbf{y}}_P^i - R^i \left(\frac{\dot{\mathbf{y}}^{l+1} - \dot{\mathbf{y}}^l}{h} \right) + R^i \left(\frac{\dot{\mathbf{y}}_P^{l+1} - \dot{\mathbf{y}}_P^l}{h} \right) \ .$$

After a few algebraic manipulations, and recalling that $\dot{\mathbf{y}}^{l+1} = D^{i^T} \mathbf{v}^{l+1}$, we have that

$$\dot{\mathbf{y}}_P^i = D^{i^T} \mathbf{v}^{l+1} + \left(h^2 K^i + h R^i \right)^{-1} \gamma_{\mathcal{A}}^{i,l+1} - \left(h^2 K^i + h R^i \right)^{-1} \left(\gamma_{\mathcal{A}}^{i,l} + h R^i (\dot{\mathbf{y}}^l - \dot{\mathbf{y}}_P^l) \right) \ . \tag{47}$$

For the entire system, by repeating the same steps that we did for the previous section, one again obtains a VI of the type $N \gamma_{\mathcal{E}}^{l+1} + \mathbf{r} \in -\mathcal{N}_{\Upsilon}(\gamma_{\mathcal{E}})$; $\gamma_{\mathcal{E}}^{l+1} \in$

Υ; but in this case the formulation of N and $\mathbf{r}$ includes the effect of damping because

$$E^i = -\left(h^2 K^i + hR^i\right)^{-1} \tag{48}$$

$$\mathbf{c}^i = \left(h^2 K^i + hR^i\right)^{-1}\left(\gamma_{\mathcal{A}}^{i,l} + hR^i(\dot{\mathbf{y}}^l - \dot{\mathbf{y}}_P^l)\right) . \tag{49}$$

To simplify the computation of $\left(h^2 K^i + hR^i\right)^{-1}$, we define R^i as a Rayleigh damping $R^i = \alpha_M^i M^i + \alpha_K^i K^i$ with M^i being the mass matrix in the coordinates of the contact. This has also the effect of reducing the number of parameters of the constitutive model: only the two values α_M^i and α_K^i are required for defining damping, instead of the nine components of R^i. Even simpler, one can set

$$R^i = \alpha_K^i K^i.$$

Moreover, the term of Eq. (49) can be simplified by multiplying Eq. (45) by h and rearranging, so $hR^i\left(\dot{\mathbf{y}}^i - \dot{\mathbf{y}}_P^i\right) = -\gamma_{\mathcal{A}}^i - hK^i\left(\mathbf{y}^i - \mathbf{y}_P^i\right)$, thus giving the following handy expressions of Eqs. (48) and (49):

$$E^i = -\frac{1}{h(h + \alpha_K^i)} K^{i^{-1}} \tag{50}$$

$$\mathbf{c}^i = -\frac{1}{h + \alpha_K^i}\left(\mathbf{y}^{i,l} - \mathbf{y}_P^{i,l}\right) . \tag{51}$$

8. Yield Surfaces

Different shapes of yield surface, along with expressions for the K^i stiffness and R^i damping, can be considered to approximate the physical properties of the contact. However, here we are interested only in simple shapes that can depend on a small number of parameters. In this section we discuss few special cases of yield surfaces that can be useful for approximating various physical phenomena in interparticle contacts.

The first case is depicted in Figure 2, where Υ^i is a Lorentz second-order cone. This forces the tangential component of the contact force to be contained in a circle whose radius is the normal force multiplied by a coefficient μ, that is $\sqrt{(\gamma_u^i)^2 + (\gamma_w^i)^2} \leq \mu\gamma_n^i$. The angle of the cone is $\phi = \tan^{-1}(\mu)$. The model corresponds to the Amontons-Coulomb friction law for dry contact, except that the sliding motion $\dot{\mathbf{y}}_P$ is on the surface of the polar cone $\Upsilon^{i\circ}$ rather than being horizontal, causing a dilatancy effect. Nevertheless, in [8]

it has been shown that the stabilization in the $\mathbf{c}^i$ terms makes the dilatancy have a small finite limit that assumes the value zero for small sliding speeds, namely for $h \downarrow 0$ or for $\mu \downarrow 0$. Moreover, by leaving $\dot{\mathbf{y}}_P^{i,l} = 0$ through all timesteps, one gets $\mathbf{c}^i = \frac{1}{h}(\mathbf{y}^{i,l})$, which is the same stabilization term of [8].

The second case is shown in Figure 3. This is the simplest way to define a cohesive set-valued contact law by introducing a single parameter, shown as c in the figure. The result is a still a cone but translated along the vertical axis. A practical consequence of this yield surface is that the contact can withstand a normal pulling impulse up to c; beyond this limit, the cohesive contact will drift apart, without breaking.[2]

Figure 4 shows an alternative where the yield surface is a "rounded" version of the cone of Figure 3. That is, Υ^i is the Minkowski sum of an Amontons-Coulomb friction cone and a ball $\mathcal{B} \in \mathbb{R}^3$. Again, a simple parameter d is required to define the cohesion. A more configurable surface could be defined by introducing an ellipsoid instead of a sphere in the Minkowski sum, hence allowing anisothropy, but at the cost of three parameters instead of one. Again, here we recall that each additional parameter not only makes the model calibration more difficult but also impacts the memory footprint of very large simulations. Hence, as long as possible, one should try to use simple yield surfaces. To this end, those in Figures 2, 3, and 4 are good candidates.

Figure 5 illustrates how, at the cost of a single additional parameter r. a simple modification of the approach captured in Figure 4 can lead to yield *capping*. The yield surface is now bound also in the pushing direction, and the contact can withstand plasticization also in the compression phase.

9. P-matrix Cone Complementarity Problems for Nonassociative Plasticity

An important modeling issue is nonassociative plasticity, whereby displacement at the plasticity limit is not orthogonal to the yield surface.

---

[2]To simulate cohesive contacts that collapse and break after they reach a limit plasticization, one can make the shapes of the yield surfaces depending on $\mathbf{y}_P$ rather than constant, similar to what happens for hardening/softening plasticity in continua. For example, if we force c to become zero after a given normal limit displacement; hence, switching from Figure 3 to Figure 2, we can simulate breakable adhesion such as in glued contacts.
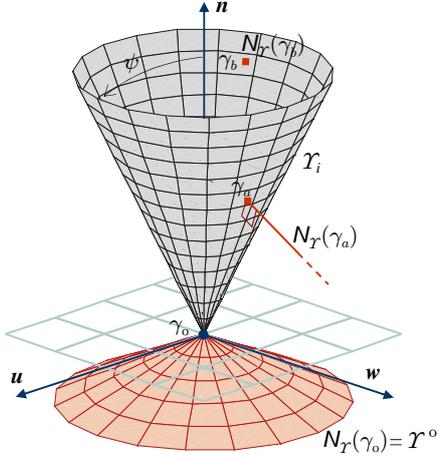
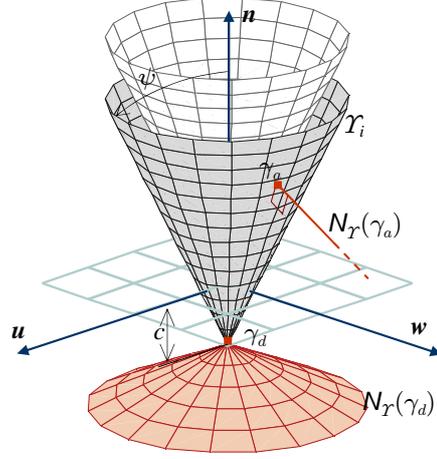Figure 2: Yield surface for Coulomb friction with associative flow.

Figure 3: Yield surface for friction and cohesion, inspired by Mohr-Coulomb.

This raises important well-posedness challenges that we now address. To this end, we investigate the modification of the normal cone when the subject cone is modified by a linear transformation.

Lemma 1. *Let $\mathcal{K}$ be a cone in $\mathbb{R}^n$ and $H \in \mathbb{R}^{n \times n}$ be an invertible matrix. Then, for all $\gamma \in \mathcal{K}$,*

$$\mathcal{N}_{H\mathcal{K}}(H\gamma) = H^{-*}\mathcal{N}_{\mathcal{K}}(\gamma) \ .$$

Proof From the invertibility of H we have that $u = H\gamma \in H\mathcal{K} \Leftrightarrow \gamma \in \mathcal{K}$.

In turn, this implies that

$$v \in \mathcal{N}_{H\mathcal{K}}(u) \Leftrightarrow \langle \widehat{u} - u, v \rangle \leq 0, \forall \widehat{u} \in H\mathcal{K}.$$

Since H is invertible, we have that $\widehat{u} \in H\mathcal{K} \Leftrightarrow H\widehat{\gamma} = \widehat{u}$ for some $\widehat{\gamma} \in \mathcal{K}$. Since $u = H\gamma$, we obtain that

$$v \in \mathcal{N}_{H\mathcal{K}}(u) \Leftrightarrow \langle H(\widehat{\gamma} - \gamma), v \rangle = \langle \widehat{\gamma} - \gamma, H^* v \rangle \leq 0, \forall \widehat{\gamma} \in \mathcal{K}.$$

But the latter expression is precisely equivalent to $H^* v \in \mathcal{N}_{\mathcal{K}}(\gamma)$, which proves the claim.

$\square$

17
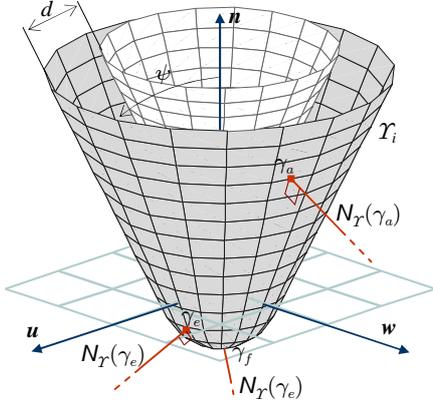
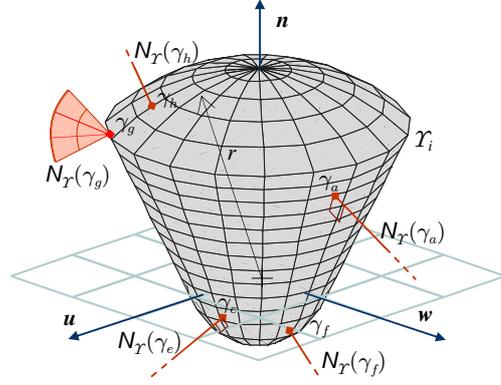Figure 4: Yield surface as a Minkowski sum of a cone and a ball.

Figure 5: Yield with capping for plasticization in compression (crunching).

We now study extensions of (43) that allow the extension of our framework to nonassociative plasticity laws. We thus study the following modified cone complementarity problems,

$$\gamma_{\mathcal{E}} \in \Upsilon, \ N\gamma_{\mathcal{E}} + \mathbf{r} \in -H^{-1}\mathcal{N}_{\Upsilon}(\gamma_{\mathcal{E}}), \tag{52}$$

where H is a symmetric positive definite matrix. In the case of multiple contacts, the cone Υ is a direct sum of three-dimensional cones. One possible modeling approach is for the matrix H to be made of three-dimensional blocks, one per contact. The structure of a generic block H^j is, with respect to the variable ordering per contact $\gamma_n, \gamma_u, \gamma_v$:

$$H^j = \begin{bmatrix} h_{11}^j & 0 & 0 \\ 0 & h_{22}^j & h_{23}^j \\ 0 & h_{23}^j & h_{33}^j \end{bmatrix}.$$

Roughly speaking, this model of plasticity allows the plastic flow to not be normal to the yield surface as long as it makes an angle with the normal of less than 90 degrees.

The question is now, of course, whether the model (52) now has a solution at all. We write it in the equivalent form:

$$\gamma_{\mathcal{E}} \in \Upsilon, \ H\left(N\gamma_{\mathcal{E}} + \mathbf{r}\right) \in -\mathcal{N}_{\Upsilon}(\gamma_{\mathcal{E}}). \tag{53}$$

We begin with the following lemma.

18

Lemma 2. *Consider the eigenvalue decomposition of H, $H = Q^T DQ$, where Q is orthogonal and D is a matrix with a positive diagonal. Problem (53) is equivalent to the following:*

$$\eta_{\mathcal{E}} \in Q\Upsilon, \ D\left(\widetilde{N}\eta_{\mathcal{E}} + \widetilde{\mathbf{r}}\right) \in -\mathcal{N}_{Q\Upsilon}(\eta_{\mathcal{E}}), \tag{54}$$

where $\widetilde{\mathbf{r}} = Q\mathbf{r}$, $\widetilde{N} = QNQ^T$.

Proof

We have that (53) is, by means of the decomposition of H, equivalent to

$$\gamma_{\mathcal{E}} \in \Upsilon, \ Q^T D\left(QNQ^T(Q\gamma_{\mathcal{E}}) + Q\mathbf{r}\right) \in -\mathcal{N}_{\Upsilon}(\gamma_{\mathcal{E}}),$$

which is in turn equivalent to

$$\gamma_{\mathcal{E}} \in \Upsilon, \ D\left(\widetilde{N}(Q\gamma_{\mathcal{E}}) + \widetilde{\mathbf{r}}\right) \in -Q\mathcal{N}_{\Upsilon}(\gamma_{\mathcal{E}}). \tag{55}$$

Using Lemma 1, we obtain that, since $Q^* = Q^T$, and thus $Q^{-*} = Q$,

$$Q\mathcal{N}_{\Upsilon}(\gamma_{\mathcal{E}}) = \mathcal{N}_{Q\Upsilon}(Q\gamma_{\mathcal{E}}).$$

Denoting by $Q\gamma_{\mathcal{E}} = \eta_{\mathcal{E}}$, and replacing it in (55), we prove this result. $\square$

The significance of the previous result is that the nonassociative plasticity cone complementarity problem (52) can now be written with respect to a rotated cone as a cone complementarity problem whose matrix is the diagonally scaled version of a positive semidefinite one.

For the final step leading to the proof of well-posedness of (52), we introduce the following definitions.

Definition 1. *A matrix $T \in \mathbb{R}^{n \times n}$ has the P_0 property if and only if*

$$\max_{i=1,2,\ldots,n} u_i(Tu)_i \geq 0, \forall u \in \mathbb{R}^n .$$

One can immediately see that a positive semidefinite matrix satisfies the P_0 property.

Definition 2. *A matrix $T \in \mathbb{R}^{n \times n}$ has the R_0 property with respect to the cone $\mathcal{K}$ if and only if*

$$u \in \mathcal{K}, Tu \in \mathcal{K}^*, \max_{i=1,2,\ldots,n} u_i(Tu)_i \leq 0 \Rightarrow u = 0 .$$

One can immediately see that a positive definite matrix has both the P_0 property and R_0 property. Unfortunately, the situation is not that simple for our case, as we want to allow for N being positive semidefinite, not being positive definite, as is the case when there is a high contact density, in addition to rescaling with a diagonal matrix. The key needed properties, however, are given by the following result.

Proposition 1. *Assume N is symmetric positive semidefinite. Then,*

(i) *the matrix $D\widetilde{N}$ has the P_0 property;*

(ii) *if the matrix N satisfies $\gamma^T N \gamma > 0$ for every $0 \neq \gamma \in \Upsilon$, then $D\widetilde{N}$ has the R_0 property with respect to the cone $Q\Upsilon$.*

Proof For part (i), assume that the result does not hold. Then, from Definition 1 we have that there exists $u \in \mathbb{R}^n$ such that

$$u_i(D\widetilde{N}u)_i < 0, \forall i = 1, 2, \ldots, n.$$

Since $u_i(D\widetilde{N}u)_i = d_i u_i(\widetilde{N}u)_i$ and $d_i > 0$, we have

$$u_i(\widetilde{N}u)_i < 0, \forall i = 1, 2, \ldots, n.$$

Thus $0 > u^T \widetilde{N} u = (Q^T u)^T N (Qu)$, which contradicts the fact that N is positive semidefinite. This proves (i).

For (ii), again assume that the result does not hold. Then, from Definition 2 we must have that there exists $0 \neq \eta \in Q\Upsilon$ such that

$$\eta_i(D\widetilde{N}\eta)_i \leq 0, \forall i = 1, 2, \ldots, n.$$

By the same reasoning as in part (i), this implies that

$$\eta_i(\widetilde{N}\eta)_i \leq 0, \forall i = 1, 2, \ldots, n.$$

and thus, that $\vec{\eta}^T \widetilde{N} \vec{\eta} \leq 0$. As $\vec{\eta} \in Q\Upsilon$, it means that $\exists \vec{\gamma} \in \Upsilon, \vec{\gamma} \neq 0, \vec{\eta} = Q\vec{\gamma}$. As $\widetilde{N} = QNQ^T$, we obtain that $\vec{\gamma} N \vec{\gamma} \leq 0$, which is a contradiction. The proof of (ii) is thus complete. $\qquad\square$

We are now ready for our main result.

Theorem 1. *Assume that the CCP problem* (43) *has a problem with symmetric semidefinite N and does not result in nontrivial internal forces, that is,*

$$\gamma \in \Upsilon, N\gamma = 0 \Rightarrow \gamma = 0.$$

Then the skewed complementarity problem modeling nonassociative plastic flow (52) *has a nonempty and bounded solution set.*

Proof Assume that there is $\gamma \in \Upsilon$, $0 \neq \gamma$ such that $\gamma^T N \gamma = 0$. As N is positive semidefinite, it means that γ is a minimum of $u^T N u$ and thus satisfies the optimality condition $N\gamma = 0$, but this contradicts the absence of nontrivial interal forces. We thus have that

$$\gamma \in \Upsilon, 0 \neq \gamma \quad \gamma^T N \gamma > 0,$$

and thus the assumption of Proposition 1 (ii) is satisfied. We can thus invoke Propositions 1 (i) and (ii) to conclude that $D\widetilde{N}$ satisfies both the P_0 and R_0 property. Using now [27, Theorem 3.7], we obtain that the cone complementarity problem (53) has a nonempty and bounded solution set. Thus, from Lemma 2 so does (52). The proof is complete. □

How might one find a solution to problem (52)? While we can extend the Gauss-Seidel proof of convergence from [4] if H is sufficiently close to 1, this may not occur in applications of interest.

An alternative is to use an interior-point approach, which we discuss in the context of problem (53), though it can be immediately transformed through the orthogonal matrix Q in an algorithm for (52). Assume that the cone $Q\Upsilon$ has a strictly convex barrier function $I_{Q\Upsilon}(\eta)$, which is three times continuously differentiable for η in the interior of $Q\Upsilon$ and which satisfies $I_{Q\Upsilon}(\eta) \to \infty$ if η approaches the boundary of $Q\Upsilon$. As in the context discussed here, all cones are rotated, or otherwise scaled second-order cones in three dimensional barriers can be easily found [27]. We then solve the parametric nonlinear equation for $\alpha \in (0, 1]$.

$$D\widetilde{N}\eta + \widetilde{r}(1 - \alpha) + \alpha\widetilde{r}_0 + \alpha\nabla I_{Q\Upsilon}(\eta) = 0.$$

The vector r_0 is chosen such as to have a given interior point of $Q\Upsilon$ as a solution of the problem for $\alpha = 1$. We now find the solution $\eta(\alpha)$ of this problem through continunation in α. For this, we need to find $\frac{d\eta(\alpha)}{d\alpha}$, which can be solved for from the equation derived from the one above:

$$D\widetilde{N}\frac{d\eta(\alpha)}{d\alpha} + \alpha\nabla^2_{\eta\eta}I_{Q\Upsilon}(\eta) - \widetilde{r} + \widetilde{r}_0 + \nabla I_{Q\Upsilon}(\eta) = 0 \;.$$

We note that the matrix of this linear equation is the sum of a P_0 matrix with a positive definite matrix (for $\alpha \neq 0$) and is thus a P matrix, and hence, invertible. So the path $\eta(\alpha)$ is uniquely defined. Since (53) also has the R_0 property, it follows that this path is bounded and that any accumulation point is a solution of (53).

Therefore interior-point algorithms can be used to solve such problems.

10. Implementation

In this section we discuss some practical issues related to an efficient implementation of a solver.

10.1. The VI Solver

When the convexity assumpion can be accepted and associated flow is involved, which happens in most cases, efficient solution schemes can be used. Rather than directly solving the VI of Eq. (42), we elected to solve the QP of Eq. (44) using a spectral projected gradient (SPG) iterative method [28]. The SPG method solves the convex-constrained optimization problems by performing at each step a gradient projection of an efficient Barzilai-Borwein iteration and enforcing a Grippo-Lampariello-Lucidi line search. A benefit of the SPG method is that it relies only on two computational primitives: matrix-vector multiplication and projection on convex sets (the yield surfaces, in our case). Requiring only matrix-vector operations, SPG relies on software constructs that are ideally suited for computation on the graphics processing unit (GPU) cards [23]. The SPG method performs better than our previous solver that was based on projected fixed-point iterations [8].

10.2. On Warm Starting

When infinitely stiff contact models are used in DVI formulations, the solution of the VI problem may exhibit slow convergence. In these cases, warm starting the solver with the values of the dual variables from the previous timestep does not improve convergence since contacts are often subject to a random pattern of activation that makes warm starting ineffective. However, we noticed that a positive side effect of introducing compliant contacts was that warm starting became effective in accelerating the VI solver. For

instance, Figure 6 shows the scenario where pressure caused by a large mass of a metric ton is applied over thousands of spheres packed in a 1m×0.7m box with a drawer that is pulled on the side; in this class of problems the introduction of warm starting allowed a reduction of two to three times on the number of iterations without affecting the precision, thus contact forces (Fig. 7) can be found faster than with infinitely stiff particles.
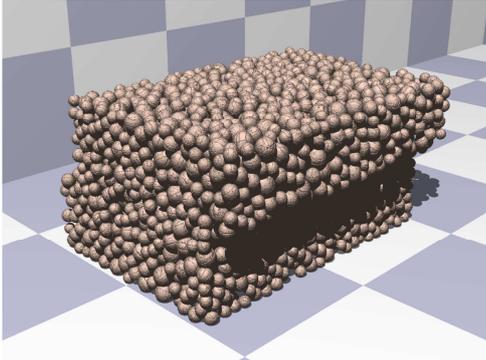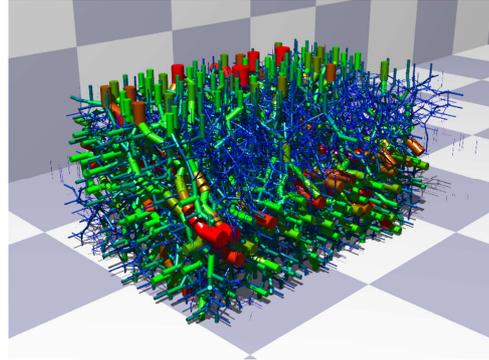

Figure 6: Shear test for granular assemblies.

Figure 7: Normal contact forces.

10.3. Notes on Computing Contact Parameters

In continuum mechanics, yield limits have the dimensions of pressure, whereas in the above discussed model we must define these limits (such as the d cohesive parameter in Figure 4) on a per contact basis, hence, with the dimensions of a force. When solving in terms of measures as in (29) in order to allow impulsive events, such limits are also multiplied by h and get the dimension of impulses. It is not always possible to automatically compute these per contact yield limits starting from the yield limits as stresses (those that are available in literature) because one would need at least some information on the contact areas represented by each contact. Indeed, the problem is that for singular contact cases, such as two stacked boxes, the number of contact points is not even uniquely defined: it might depend on the type of the collision detection algorithm, and it could even change erratically from one timestep to another. A similar difficulty arises in defining the contact compliance. However, an interesting exception is the frequent case of contacts between spheres, that produce only single contact points between pairs, so yield limits and compliances are easily determined and can be defined before the simulation starts.

10.4. Notes on Plastic Flow

When plasticization is taken into account, the plastic flow $\mathbf{y}_P^{i,l}$ at the beginning of the l-th step must be known for each i-th contact; this means that at the end of the timestep computation the updated $\mathbf{y}_P^{i,l+1}$ value must be stored somewhere, to be used for the next step. Such data can be saved in a persistent contact manifold; that is, it can be associated with the i-th contact if it survives across timesteps. This solution works fine if contacts never separate, as in dense granular packings, and for small displacements of contacts: in fact if the contact experiences large tangential motion during plasticization (for example, as in the case of a sliding sphere that digs a trench over another sphere), this is not sufficient, and the plastic flow must be stored in a fine mesh that covers the surfaces, similarly to what happens in continuum mechanics with FEM. This would introduce a large amount of data that, expecially when replicated for the thousands of particles in granular matter, would hamper the benefit of using the fast DVI method instead than a full FEM description of the moving parts; hence one might consider storing plastic flow on a very coarse mesh for each particle, which would be sufficient to capture the macroscopic effects of collective plasticization without requiring much detail per each particle. As a side note, for the special case of irreversible cohesion, given the small amount of displacement and given the fact that once contacts break apart, the plastic flow is not used anymore, thus storing the plastic flow in the contact manifold is enough.

11. Example

We present the example of the compaction and shear test of a granular media because in perspective, after this type of calibration, one could use the method to study advanced cases of earth-moving machines (bulldozers, agricultural vehicles) interacting with soil.

A soil sample is put in a box with base 0.1 m $\times$ 0.1 m and height 0.2 m and is cut in two stacked parts. The upper part of the box can shift horizontally, and the force reacting to the shear displacement is measured [29].

In the simulation of this example, a mass of 100 kg is applied over the granular matter. This corresponds approximately to a pressure of 10 Pa over a surface of 0.1 m $\times$ 0.1 m.

The simulation is performed considering 720 spherical bodies, whose properties are listed in Table 1, resulting in a compliant cohesionless granular matter compacted under the load mass as shown in Figure 8. Normal and

tangential compliances are 0.9×10^{-7} m/N and the damping coefficient is $\alpha_K = 0.1$.

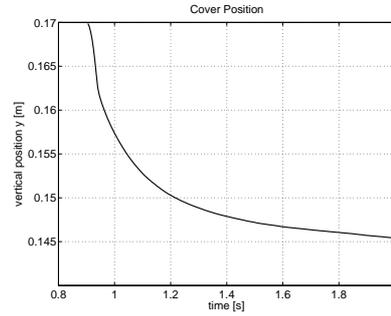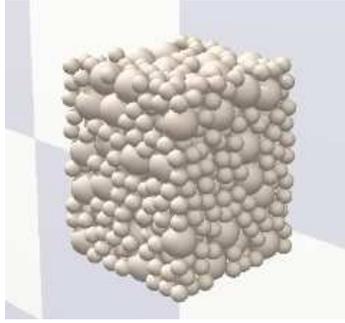| Number	Density	Diameter	Friction
90	1700 kg/m^3	0.020 m	0.5
630	1700 kg/m^3	0.010 m	0.5

Table 1: Families definition of granular material

Figure 8: Shear test for granular assemblies in a shear box of 0.1 m × 0.1 m.

Figure 9: Vertical position of the load over the granular material.

At the beginning of the simulation the upper mass is released over the granular material and falls down from a small height, as shown in Figure 9, compressing the particles until the expected pressure is reached.

Among the possible outputs, in this example we focus on the vertical force on the floor and on the shear resistance on the upper part of the box, which are related respectively to the normal pressure and to the shear stress of the soil.

The trend of the vertical load on the box floor is plotted in Figure 10. One can spot the instant when the upper mass reaches the granular matter, corresponding to the sudden increase of the load of almost 1000 N, as expected. At the same moment one can see a spike in shear forces of Figure 11, probably due to a lock-in effect between the masses. After 0.1 s from this event, the shear test begins and the upper part of the box shifts; hence the shear force increases.

This test provides a small shear speed, allowing one to measure the shear resistance in quasistatic conditions (reaching the breakup stress would require longer simulation periods, but it was not the goal of this example).
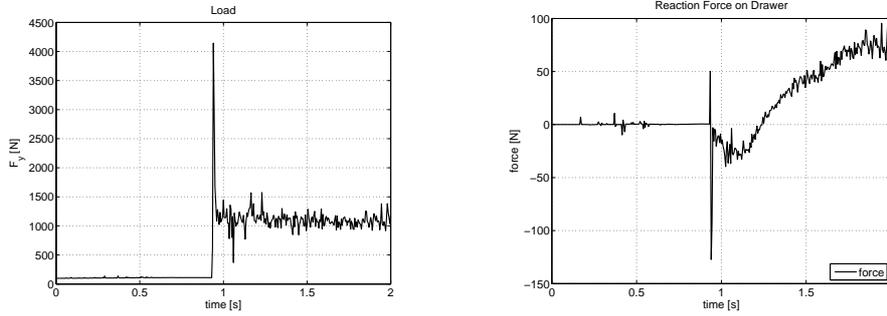
25

Figure 10: Vertical load on the box floor. Figure 11: Reaction force on the drawer box.

Figures 12 and 13 show the modulus and the direction of the pebbles within the granular matter at $t = 1.2$ s. The predominant component is due to the downward motion of bulk under the compaction effect of the upper load, whereas the shift displacement due to the shift of the upper part of the box is less significant, as the shear speed is very low.
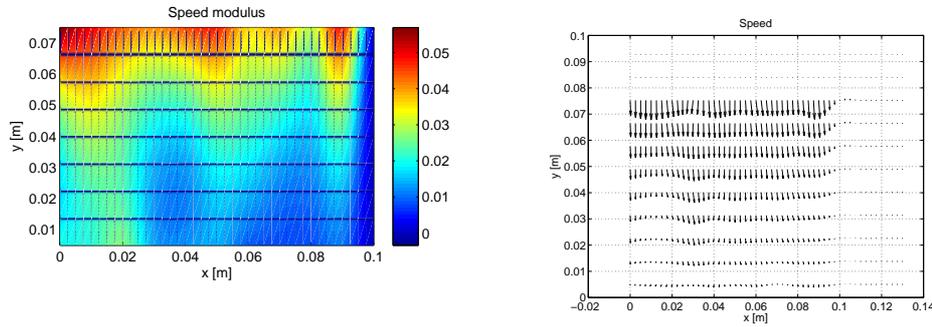

Figure 12: modulus of the particles speed under the vertical load. Figure 13: Particles speed direction under the vertical load.

12. Conclusions

This work discusses how to extend a DVI method used to capture frictional contact in rigid bodies dynamics problems to accommodate also local compliance, viscous damping and plasticity in contact points. The benefit

26

is that nonsmooth phenomena such as unilateral frictional contact, cohesion and piecewise plasticity are solved directly at each time step as a single complementarity problem using a VI or QP solver.

Future work must address the following issues: improving the convergence of the VI solver in case of very large systems (millions to billions of contact events), finding an efficient way of storing the plastic flow in different zones of the surface, and evaluating the possibility of adopting interior-point methods.

Acknowledgements

Appendix A: Formulation of D vectors

We assume that the vector of generalized velocities $\mathbf{v}$ contains the speeds of the centers of mass of the bodies $\dot{\mathbf{x}}^{(W)}$ expressed in absolute coordinates (W) and the angular velocities $\omega^{(i)}$ expressed in the local coordinates of the i-th body, as $\mathbf{v} = \left[\dot{\mathbf{x}}_1^{(W)}, \omega_1^{(1)}, \dot{\mathbf{x}}_2^{(W)}, \omega_2^{(2)}, \ldots, \right]^T$.

Given a contact between a pair of two rigid bodies A and B, we define the positions of the two contact points with respect to the centers of mass, expressed in the coordinate systems of the two bodies, as $\mathbf{s}_A^{(A)}$ and $\mathbf{s}_B^{(B)}$. The absolute rotations of the coordinate systems of the bodies are $R_A^{(W)}, R_B^{(W)} \in \mathsf{SO}(\mathbb{R}, 3)$ and the absolute rotation of the contact plane is $R_P^{(W)} \in \mathsf{SO}(\mathbb{R}, 3) = [\mathbf{n}, \mathbf{u}, \mathbf{w}]$. Thus, assuming A as the reference body, and assuming small distance $\mathbf{s}_A^{(W)} - \mathbf{s}_B^{(W)}$, the vectors $\mathbf{D}_{\gamma_n}$, $\mathbf{D}_{\gamma_u}$, $\mathbf{D}_{\gamma_w}$ have the following simple expression $D_\gamma = [\mathbf{D}_{\gamma_n}, \mathbf{D}_{\gamma_u}, \mathbf{D}_{\gamma_w}] \in \mathbb{R}^{3 \times m_v}$:

$$D_\gamma^T = \begin{bmatrix} 0, & \ldots & -R_P^{(W)^T}, & R_P^{(W)^T} R_A^{(W)} \widetilde{s}_A^{(A)}, & \ldots, \\ 0, & \ldots & R_P^T, & -R_P^{(W)^T} R_B^{(W)} \widetilde{s}_B^{(B)}, & \ldots, & 0 \end{bmatrix} \tag{56}$$

where $\widetilde{s}$ is the skew symmetric matrix such that $\widetilde{s}\mathbf{x} = \mathbf{s} \wedge \mathbf{x}$.

We remark that, because of the extreme sparsity of (56), only the two following 3×6 matrices need to be stored per each contact.

$$D_{\gamma,A}^T = \left[-R_P^{(W)T}, \quad R_P^{(W)T} R_A^{(W)} \widetilde{s}_A^{(A)} \right] \tag{57}$$

$$D_{\gamma,B}^T = \left[R_P^{(W)T}, \quad -R_P^{(W)T} R_B^{(W)} \widetilde{s}_B^{(B)} \right] \tag{58}$$

Bibliography

[1] P. A. Cundall, O. D. L. Strack, A discrete numerical model for granular assemblies, Geotechnique 29 (1979) 47–65.

[2] F. Pfeiffer, C. Glocker, Multibody Dynamics with Unilateral Contacts, John Wiley, New York City, 1996.

[3] D. E. Stewart, J. C. Trinkle, An implicit time-stepping scheme for rigid-body dynamics with inelastic collisions and Coulomb friction, International Journal for Numerical Methods in Engineering 39 (1996) 2673–2691.

[4] M. Anitescu, A. Tasora, An iterative approach for cone complementarity problems for nonsmooth dynamics, Computational Optimization and Applications 47(2) (2010) 207–235.

[5] A. Tasora, M. Anitescu, A convex complementarity approach for simulating large granular flows, Journal of Computational and Nonlinear Dynamics 5 (2010) 1–10.

[6] M. Renouf, P. Alart, Conjugate gradient type algorithms for frictional multi-contact problems: applications to granular materials, Computer Methods in Applied Mechanics and Engineering 194 (2005) 2019–2041.

[7] K. Krabbenhoft, A. Lyamin, J. Huang, M. V. da Silva, Granular contact dynamics using mathematical programming methods, Computers and Geotechnics 43 (2012) 165 – 176.

[8] A. Tasora, M. Anitescu, A matrix-free cone complementarity approach for solving large-scale, nonsmooth, rigid body dynamics, Computer Methods in Applied Mechanics and Engineering 200 (2011) 439 – 453.

[9] K. Krabbenhoft, J. Huang, M. da Silva, A. Lyamin, Granular contact dynamics with particle elasticity, Granular Matter 14 (2012) 607–619.

[10] C. Lacoursière, M. Servin, A. Backman, Fast and stable simulation of granular matter and machines, in: Proceedings of the Fifth International Conference on Discrete Element Methods (DEM5).

[11] P. Song, J.-S. Pang, V. Kumar, A semi-implicit time-stepping model for frictional compliant contact problems, International Journal of Numerical Methods in Engineering 60 (2004) 267–279.

[12] R. T. Rockafellar, Convex Analysis, Princeton University Press, 1970.

[13] J. Moreau, Numerical aspects of the sweeping process, Computer Methods in Applied Mechanics and Engineering 177 (1999) 329 – 349.

[14] J. J. Moreau, Unilateral contact and dry friction in finite freedom dynamics, in: J. J. Moreau, P. D. Panagiotopoulos (Eds.), Nonsmooth Mechanics and Applications, Springer-Verlag, Berlin, 1988, pp. 1–82.

[15] M. Anitescu, F. A. Potra, D. Stewart, Time-stepping for three-dimensional rigid-body dynamics, Computer Methods in Applied Mechanics and Engineering 177 (1999) 183–197.

[16] D. E. Stewart, Reformulations of measure differential inclusions and their closed graph property, Journal of Differential Equations 175 (2001) 108–129.

[17] F. Jourdan, P. Alart, M. Jean, A Gauss Seidel like algorithm to solve frictional contract problems, Computer methods in applied mechanics and engineering 155 (1998) 31 –47.

[18] C. G. Christoph Studer, Solving normal cone inclusion problems in contact mechanics by iterative methods, Journal of System Design and Dynamics 1 (2007) 458–467.

[19] D. Stewart, J.-S. Pang, Differential variational inequalities, Mathematical Programming 113 (2008) 345–424.

[20] D. Kinderleher, G. Stampacchia, An Introduction to Variational Inequalities and Their Application, Academic Press, New York, 1980.

[21] D. Stewart, Convergence of a time-stepping scheme for rigid body dynamics and resolution of Painlevé's problems, Archive Rational Mechanics and Analysis 145 (1998) 215–260.

[22] V. I. Bogachev, Measure theory, Springer, New York, 2007.

[23] A. Tasora, D. Negrut, M. Anitescu, Large-scale parallel multi-body dynamics with frictional contact on the graphical processing unit, Journal of Multi-body Dynamics 222 (2008) 315–326.

[24] R. D. Mindlin, H. Deresiewicz, Elastic spheres in contact under varying oblique forces, Journal of Applied Mechanics 20 (1953) 327–344.

[25] M. Yu, G. Ma, H. Qiang, Y. Zhang, Generalized Plasticity, Springer, 2010.

[26] F. Calvetti, R. Nova, Micromechanical approach to slope stability analysis, in: Degradations and instabilities in geomaterials, Springer, New York, 2004.

[27] L. Kong, L. Tunçel, N. Xiu, Homogeneous Cone Complementarity Problems and P Properties, ArXiv e-prints (2009).

[28] E. G. Birgin, J. M. Martínez, M. Raydan, Nonmonotone spectral projected gradient methods on convex sets, SIAM Journal on Optimization 10 (1999) 1196–1211.

[29] ASTM, D3080:2011, Standard test method for direct shear test of soils under consolidated drained conditions, 2011.