



HAL
open science

A mixed formulation for frictional contact problems prone to Newton like solution methods

Pierre Alart, Alain Curnier

► **To cite this version:**

Pierre Alart, Alain Curnier. A mixed formulation for frictional contact problems prone to Newton like solution methods. *Computer Methods in Applied Mechanics and Engineering*, 1991, 92 (3), pp.353-375. 10.1016/0045-7825(91)90022-X . hal-04264964

HAL Id: hal-04264964

<https://hal.science/hal-04264964v1>

Submitted on 30 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A mixed formulation for frictional contact problems prone to Newton like solution methods

P. Alart

*Lab. Génie Physique et Mécanique des Matériaux, URA CNRS 793 ENSPG BP 46,
38402 Saint-Martin d'Hères Cedex, France*

A. Curnier

*Lab. Mécanique Appliquée, Département de Mécanique, Ecole Polytechnique Fédérale de Lausanne,
1015 Lausanne, Suisse*

A mixed penalty–duality formulation of the frictional contact problem, inspired from an augmented Lagrangian approach is proposed. The continuity of the resulting conewise linear operator is used to establish a uniqueness condition on the coefficient of friction. Modified and generalized Newton methods are examined and sufficient conditions for their convergence conjectured. A cylindrical frictional contact problem assesses the stability of the method. Mixed penalty–duality methods are found more accurate and stabler than penalty methods and as economical as them.

Introduction

Contact problems involving friction are difficult to formulate and even more difficult to solve, because they are governed by a *multivalued* tribological law which does *not derive* from a natural *potential* (even non-differentiable). As such, they cannot be formulated as standard optimization problems (with inequality constraints).

In [1] a rigorous formulation of the frictional contact problem was obtained using *penalty* methods to circumvent the multivalued character of the law and constructing a non-smooth (yet continuous) *unsymmetric operator* directly, for want of deriving it from a lacking potential. Newton's method was then generalized to solve the resulting non-linear equation. However, the solutions carried slight systematic errors inherent to penalty methods and special damping techniques were necessary to ensure the convergence of Newton's method.

In this paper, an alternative formulation of the frictional contact problem is developed using a mixed *penalty–duality* approach (i.e. adding Lagrange multipliers) inspired from the augmented Lagrangian method to treat the multivalued aspect and deriving the relevant unsymmetric operator from a '*quasi–augmented Lagrangian*' to palliate the absence of a genuine one. Modified and generalized Newton methods are then examined to solve the resulting system of equations. The solutions satisfy the contact constraints and friction criteria exactly and Newton's method converges without any stabilizing damping.

The augmented Lagrangian method was initially introduced by Hestenes [2] and Powell [3]

for solving non-linear programming problems with equality constraints. It was extended to treat convex differentiable optimization problems with inequality constraints (such as the frictionless *contact* problem) by Rockafellar [4]. The same year, Fortin [5] gave an application to a non-differentiable case (such as the pure *friction* problem). More recently, an incomplete formulation was proposed by Landers and Taylor [6] for frictionless contacts and an augmented Lagrangian merit function was introduced by Barthold and Bischoff [7] for stabilizing their frictional contact algorithm. However, a complete mixed formulation of the *frictional contact* problem and a thorough discussion of its crucial features such as the lack of a genuine augmented Lagrangian, the conewise linearity of the resulting unsymmetric operator and, above all, its *continuity* across the border lines separating the (gap, stick and slip) regions of linearity are yet lacking. In this paper, we exhibit these properties and use them to establish a necessary and sufficient condition on the friction coefficient for the solution to curved, discrete, small slip contacts to be *unique* and to guarantee the *convergence* of Newton's method.

A brief outline of the paper follows. In Section 1, a contact mechanics summary provides the necessary background and motivation. A 'brut' variational formulation of the discrete incremental problem is stated in Section 2, as a starting point. In Section 3, we show how to construct the frictional contact operator following an augmented Lagrangian approach. Section 4 is devoted to Newton type solution methods. The cylindrical contact of a disk in a bore is used to illustrate the stability of the algorithm in Section 5.

1. Contact mechanics background

The theory of contact mechanics is less developed than that of continuum mechanics. The choice of appropriate kinematic and dynamic variables to describe contact processes is not well settled yet.

Moreover, normal contact and tangential friction tribological laws are unusual multivalued functions and, worse, they cease to derive from a (pseudo-, i.e. non-differentiable) potential when the two components are treated conjointly, as in a non-associated Coulomb's model. The reader is referred to [1] for a more complete introduction to the subject.

1.1. Kinematics and statics

First of all, it is necessary to define kinematic variables for measuring the relative motion of the two bodies candidate to contact and static variables for enforcing their interaction.

To define the relative displacement between two contacting deformable bodies, it is convenient to take one of them, say body-2, for reference and to describe body-1 with respect to it (Fig. 1). Accordingly, the contact surface is identified with a portion Γ_c^2 of the deformed boundary of body-2. We note a^2 a particle of Γ_c^2 and $x^2 = x(a^2)$ its coordinates in some reference frame. To each *master* particle a^2 of Γ_c^2 , we associate a *slave* particle $a^1(a^2)$ of Γ_c^1 such that

$$x^1 = x(a^1) \in \operatorname{argmin}\{\|z - x^2\| \mid z \in x(\Gamma_c^1)\} \quad \forall a^2 \in \Gamma_c^2. \quad (1)$$

A local frame can be attached to the slave particle a^1 : $n = n^1(x^1)$ is the outward unit normal

vector to body-1. We define the *normal contact distance* as a scalar field on Γ_c^2 :

$$d_n(a^2) = n^i \{x(a^2) - x(a^1(a^2))\} \quad \forall a^2 \in \Gamma_c^2. \quad (2)$$

As time is relevant, the vector field of *relative velocities* can be written as

$$\dot{\delta}(a^2, t) = \dot{x}(a^2, t) - \dot{x}(a^1, t), \quad (3)$$

where $\dot{\delta}_t$ is the projection of $\dot{\delta}$ on the tangent vector space to $x(\Gamma_c^1)$ at x^1 and represents then the vector field of *tangential relative velocities*.

The choice of the particle a^1 is not unique in general unless body-1 is convex or a^2 is in contact with Γ_c^1 . If the argmin is a set, anyone of its element is acceptable. In the simplest case, when Γ_c^1 is convex and x^2 lies outside body-1, x^1 is the projection of x^2 on $x(\Gamma_c^1)$ and

$$d = x(a^2) - x[a^1(a^2)] = d_n n, \quad \text{i.e. } d_t = 0 \quad (4)$$

(situations where x^2 penetrates body-1 are not precluded at this stage). If d_n is positive, it is the local gap between the two bodies. $\dot{\delta}_t$ is defined even when contact does not occur and is not equal to \dot{d}_t , which is zero.

To model contact behaviour, it is necessary to introduce a contact stress field λ defined as the stress vector acting on the boundary of body-2. It is convenient to resolve the contact stress into normal and tangential components also

$$\lambda = \lambda_n n + \lambda_t. \quad (5)$$

We are now in a position to define the tribological law. For didactic reasons, we begin with the normal *contact* law (perfect slip), proceed with the pure *friction* law (at constant pressure) and terminate with the coupled *frictional contact* law (Coulomb's non-associated friction).

1.2. Unilateral contact law

Usually, two bodies can either separate-but-not-pull each other *or* they can press-on-but-not-penetrate each other. These two exclusive status: *gap* ($d_n > 0, \lambda_n = 0$) or *contact* ($d_n = 0, \lambda_n < 0$) are classically formulated by an impenetrability condition, a compression condition and a complementarity condition:

$$d_n \geq 0, \quad \lambda_n \leq 0, \quad \lambda_n d_n = 0. \quad (6)$$

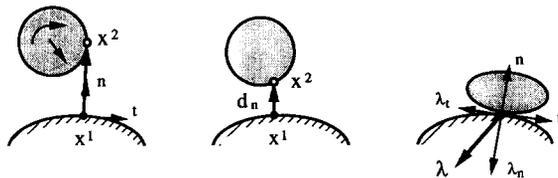


Fig. 1. Local contact kinematics and dynamics.

The gap and contact status are clearly distinguished on the graph of the relationship between λ_n and d_n as shown in Fig. 2. The multivalued contact law $\lambda_n[d_n]$ and its inverse $d_n[\lambda_n]$ can be shown to derive from two conjugate pseudo- (i.e. non-differentiable) potentials, in the form of two similar inclusions,

$$\lambda_n \in \partial \psi_{\mathbb{R}^+}(d_n) \quad \text{or} \quad d_n \in \partial \psi_{\mathbb{R}^-}(\lambda_n), \quad (7)$$

where $\psi_{\mathbb{R}^+}$ is the indicator function of the positive half line, $\partial \psi_{\mathbb{R}^+}$ its sub-differential and $\psi_{\mathbb{R}^-}$ its conjugate (the indicator function of the polar cone \mathbb{R}^- of \mathbb{R}^+) [8].

1.3. Pure friction law

By pure friction, we mean friction at constant pressure as it occurs between two bodies in steady contact (such as a dead weight on an inclined board for instance). Two bodies in steady contact either stick-to-and-shear each other *or* they slip-and-rub-on each other along a same direction. These two exclusive status:

$$\text{stick: } (\|\dot{\delta}_t\| = 0, \|\lambda_t\| < k),$$

$$\text{slip: } (\|\dot{\delta}_t\| > 0, \|\lambda_t\| = k, \lambda_t / \|\lambda_t\| = \dot{\delta}_t / \|\dot{\delta}_t\|),$$

are traditionally expressed by a slip rule, a friction criterion and a complementarity condition:

$$\dot{\delta}_t = \|\dot{\delta}_t\| \frac{\lambda_t}{\|\lambda_t\|}, \quad \|\lambda_t\| \leq k, \quad \|\dot{\delta}_t\|(\|\lambda_t\| - k) = 0, \quad (8)$$

where k is the friction threshold: $k = -\mu\lambda_n = \text{constant}$ (since λ_n is constant). The 2D graph of the resulting relationship between λ_t and $\dot{\delta}_t$ (or $d\delta_t$) is shown in Fig. 3 where the three status: stick, backward slip and forward slip are easily sorted out. As for unilateral contact, the multivalued slip law $\dot{\delta}_t[\lambda_t]$ and its inverse, the multivalued friction law $\lambda_t[\dot{\delta}_t]$, can both be shown to derive from two conjugate pseudo-potentials, in the form of two different inclusions:

$$\dot{\delta}_t \in \partial \psi_C(\lambda_t) \quad \text{or} \quad \lambda_t \in k \partial \|\dot{\delta}_t\|, \quad (9)$$

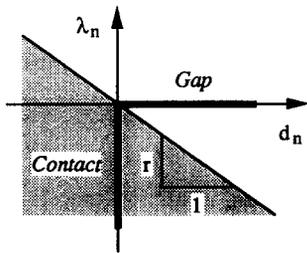


Fig. 2. Unilateral contact (graph and mixed operator).

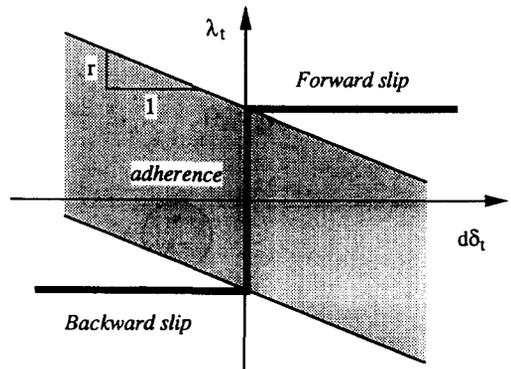


Fig. 3. Friction (graph and mixed operator).

where $k\|\dot{\delta}_t\| = \psi_C^*(\dot{\delta}_t)$ is the Fenchel conjugate of the indicator function ψ_C of the convex disc C of radius k centered at the origin

$$C = \{ \lambda_t / \|\lambda_t\| \leq k \} . \quad (10)$$

Because C is a disk (not a cone) ψ_C^* is a norm (not an indicator function). The conjugate function of an indicator function being always positive homogeneous of degree one ($\psi_C^*(\alpha\dot{\delta}_t) = \alpha\psi_C^*(\dot{\delta}_t) \forall \alpha \geq 0$) the shear-stress slip-rate law $\lambda_t[\dot{\delta}_t]$ is positive homogeneous of degree zero, which expresses the experimental evidence that friction is rate-independent in first approximation. Anisotropic friction can be accommodated by taking an elliptical disc instead of a circular one. Thus, general steady contact, rate independent friction laws deriving from a pseudo-potential can be written as [8]

$$\dot{\delta}_t \in \partial\psi_C(\lambda_t) \quad \text{or} \quad \lambda_t \in \partial\psi_C^*(\dot{\delta}_t) , \quad (11)$$

where C is a convex set containing the origin.

1.4. Frictional contact law

Contrary to what the two previous (didactic) paragraphs may infer, normal contact and tangential friction are strongly coupled phenomena: according to Coulomb's classical law, the friction shear is proportional to the contact pressure. The trouble is that this coupling is not symmetric: the contact pressure is independent of the friction shear. Keeping this lack of symmetry in mind, general and realistic frictional contact laws can still be constructed by combining the normal contact and tangential friction laws as follows:

$$\lambda_n \in \partial\psi_{R^+}(d_n) \quad \text{or} \quad d_n \in \partial\psi_{R^-}(\lambda_n) , \quad (12a)$$

$$\lambda_t \in \partial\psi_{C(\lambda_n)}^*(\dot{\delta}_t) \quad \text{or} \quad \dot{\delta}_t \in \partial\psi_{C(\lambda_n)}(\lambda_t) , \quad (12b)$$

where the friction criterion $C(\lambda_n)$ now *depends* on the pressure. Coulomb's classical law of isotropic friction consists in taking for $C(\lambda_n)$ the closed disk centered at the origin with radius equal to the product of $(-\lambda_n)$ by the friction coefficient μ . $C(\lambda_n)$ is thus a section of Coulomb's cone:

$$C(\lambda_n) = \{ \lambda_t \mid \|\lambda_t\| \leq -\mu\lambda_n \} . \quad (13)$$

In case of *gap* ($d_n > 0, \lambda_n = 0$), the friction convex set is reduced to the singleton $\{0\}$ and $\partial\psi_{\{0\}}(0)$ is equal to the tangent vector space. Then the expressions (12b) involving the sliding velocity remain meaningful for the relative velocity, thus the notation overlap.

In case of *stick contact* ($d_n = 0, \lambda_n < 0; \|\dot{\delta}_t\| = 0, \|\lambda_t\| < -\mu\lambda_n$), the shear stress is undetermined.

In case of *slip contact* ($d_n = 0, \lambda_n < 0; \|\dot{\delta}_t\| > 0, \|\lambda_t\| = -\mu\lambda_n$), the slip rule implied in $\dot{\delta}_t[\lambda_t]$ is associated (i.e. normal) to Coulomb's disc but not associated to Coulomb's cone.

We propose to call the indicator function $\psi_{C(\lambda_n)}$ and its conjugate $\psi_{C(\lambda_n)}^*$ '*quasi*'-potentials instead of pseudo-potentials to stress the dependence of the convex set on the pressure.

Formally, the dependence of the convex set on the solution is a minor change, but fundamentally it is of major importance. In the special case of Coulomb's disc, the friction quasi-potential is simply the norm of the sliding velocity scaled by the shear stress magnitude:

$$\psi_{C(\lambda_n)}^*(\dot{\delta}_t) = -\mu\lambda_n \|\dot{\delta}_t\|, \quad \text{thus } \lambda_t \in -\mu\lambda_n \partial \|\dot{\delta}_t\|. \quad (14)$$

2. Variational formulation of the discrete incremental problem

2.1. Quasi-static evolution

In quasi-static cases, the slip rate $\dot{\delta}_t$ in the slip rule (8) or the friction law (11) must be approximated by increments $d\delta_t$. For this time discretisation, a fully implicit algorithm has long proven the most stable (and accurate) scheme, both in plasticity [9, 8, 10] and in friction [1, 11]. Due to the homogeneity of the normal cone $\partial\psi_{C(\lambda_n)}$, this operation simply consists in replacing $\dot{\delta}_t$ by a slip increment $d\delta_t$ in (9) and (11) provided the local frame $(n-t)$ is assumed fixed over the increment

$$d\delta_t \in \partial\psi_{C(\lambda_n)}(\lambda_t), \quad \lambda_t \in \partial\psi_{C(\lambda_n)}^*(d\delta_t). \quad (15)$$

The second expression is particularly suitable for implementing contact treatment in existing finite element programs based on the incremental loading (continuation) technique.

From here on, we will restrict our attention to the *incremental* problem (any one step of the loading sequence). Within such an increment, a simultaneous treatment of contact and friction laws ((12a) and (12b)) requires an accurate localisation of the slave particle a^1 at impact time and determination of the associated local frame. This is briefly discussed in the next paragraph.

2.2. Contact discretisation

Using the finite element method to discretize the two solids, the discrete contact surface consists in the set of nodes I^2 located on the mesh boundary Γ_{ch}^2 . A contact element is the associated to each node of Γ_{ch}^2 as follows. If a one-to-one correspondence between the boundary nodes of the two bodies can be established and maintained throughout the contact duration (stick or small slip contact), then a node-on-node geometry is quite adequate. Otherwise, a node-on-facet contact must be used to account for initial mismatching as well as subsequent sliding (moderate slip contact). Finally, if body-1 is rigid, a simpler but more accurate node-on-curve treatment is possible (body-1 does not need to be discretized and finite slip problems are tractable).

With the *node-on-node* geometry, the slave node of Γ_{ch}^1 on mesh-1 which is the closest to each master node of Γ_{ch}^2 can be assigned in advance. This assignation can be seen as a coarse discrete version of (1). But this definition supposes that nodes e^2 and $e^1(e^2)$ come in contact exactly one on top of the other which is exceptionally the case. Otherwise some impact localisation strategy must be used. The fixed local frame $(n-t)$ is defined by the nodes adjacent to node e^1 (2 nodes in 2D, 4 in 3D) [12, 1]. The impact location of node e^2 on body-1

is then taken as the intersection of its linear trajectory (over the increment) with the tangent plane to body-1 passing through node e^1 .

With the *node-on-facet* geometry, the slave particle location x^1 (which is unknown a priori) on the slave facet (which is known a priori) can be expressed in terms of the facet corner coordinates by means of the interpolation functions used for body-1 [13, 14, 1]. The local frame is determined by the facet geometry in this case, and the impact point localized accordingly.

Finally, with the *node-on-curve* geometry, the slave particle position x^1 on the rigid surface Γ_c^1 can be localized exactly (at least in principle) at any time. Over a time increment, two instants are commonly used (Fig. 4):

- the beginning of the increment, corresponding to an explicit type of projection algorithm;
- the impact time, whose determination requires an approximation of the master particle trajectory, typical of a predictor strategy.

Depending on the approach, both the slave particle x^1 and the local frame ($n - t$) will differ significantly for strongly curved obstacles.

In the three fore-mentioned cases, a *global* normal distance vector d_n can be defined as the vector composed of the discrete normal distances ${}^e d_n$, $e = 1, p$ – for convenience, we denote discrete nodal variables by the same letters as the corresponding continuous variables:

$$d_n = \{ {}^1 d_n, \dots, {}^p d_n \}^t, \quad (16)$$

where p is the number of master nodes, i.e. of contact elements. Moreover d_n can be written as a linear function of u :

$$d_n = A_n u + d_n^o, \quad {}^e d_n = n_e^t u + {}^e d_n^o, \quad (17)$$

where ${}^e d_n^o$ denotes the initial normal contact distance at the beginning of the increment. The global slip increment vector $d\delta_i$ is constructed in the same way in 2D discretisation:

$$d\delta_i = A_i u + \delta_i^o, \quad (18)$$

where δ_i^o is the initial tangential contact gap at the beginning of the increment. Accordingly, λ

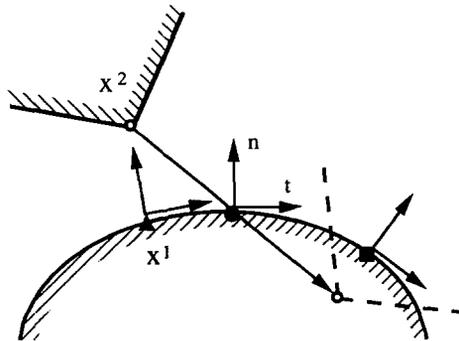


Fig. 4. Explicit projection (▲), impact detection (●) and implicit projection (■) methods.

becomes the global contact force vector whose components are the contact forces concentrated at the master nodes,

$$\lambda = \{ \dots, {}^e\lambda_n, {}^e\lambda_t, \dots \}^t. \quad (19)$$

The rotation and assembly of this force vector (effected by the transposed matrices A_n^t and A_t^t) will be discussed later.

2.3. Variational formulation

The following formulation is based on the non-differentiable functional expressions (7), (11) and (12) of the contact and friction laws. For convenience, we will *also* assume the existence of a differentiable strain energy functional to characterize the elastic response of the (two) separated bodies, but it is emphasized that most of the following results will not depend on this assumption. Accordingly, the equilibrium of the two discretized bodies in the absence of contact is written as

$$u = \operatorname{argmin} [\varphi(v)] \quad \Leftrightarrow \quad \varphi(u) \leq \varphi(v) \quad \forall v, \quad (20)$$

where $\varphi(v)$ is the (elastic) strain energy potential in terms of the global displacement vector v . A necessary condition, which becomes sufficient if φ is strictly convex, for this equilibrium to hold, is

$$G(u) = \nabla\varphi(u) = K(u) - q = 0, \quad (21)$$

where ∇ denotes the gradient operator, $K(u)$ is the internal force vector and q the external one.

If one treats contact and friction *separately*, then two new optimisation problems are obtained.

Unilateral contact problem:

$$u = \operatorname{argmin} \{ \varphi(v) + \psi_{\mathbb{R}_+^p}(d_n(v)) \}, \quad (22)$$

where \mathbb{R}_+^p is the positive cone of \mathbb{R}^p .

Pure friction problem:

$$u = \operatorname{argmin} \{ \varphi(v) + \psi_{C_n^p}^*(d\delta_i(v)) \}, \quad (23)$$

where C_n^p is the cartesian product of the p local discs C of fixed radius k . C_n^p in (23) does not depend on the solution u through λ_n . This assumption still allows to define the friction problem (23) as an optimisation problem.

Classically, the solution u of the above problems is characterized by variational inequalities (VI) [15, 16]. *Formally*, coupling contact and friction reduces to adding the two corresponding pseudo-potentials to the elastic potential φ .

Frictional contact problem:

$$u = \operatorname{argmin} \{ \varphi(v) + \psi_{\mathbb{R}_+^p}(d_n(v)) + \psi_{C_n^p[\mu]}^*(d\delta_r(v)) \}, \quad (24)$$

where C_n^p is now the cartesian product of the local discs $C(\lambda_n)$ of variable radius $-\mu\lambda_n$. It must be stressed that the above problem is not a standard optimisation problem but only a ‘quasi’-optimisation problem because the convex set C_n^p depends on the solution u through λ_n . The solution u of this coupled problem is now characterized by quasi-variational inequalities (QVI) [17]. We coined the term quasi-optimization by analogy. The absence of a genuine potential for this frictional contact problem reflects the non-associated character of the slip rule.

3. Mixed operators by augmented Lagrangian approach

The purpose of this section is to construct frictional contact operators which are smooth enough to use Newton like solution methods. Newton’s method can be formally extended to Lipschitz continuous operators [1] by using the notion of generalized Jacobian [18]. This possibility has been investigated in [20] using penalty methods. But the resulting algorithms are subject to troublesome numerical instabilities (illustrated therein), calling for alternatives.

In what follows, we propose a *mixed penalty–duality approach* based on the augmented Lagrangian method. For didactic reasons, we first derive the pure contact and friction operators associated to problems (22) and (23) from their respective augmented Lagrangian. We then show how to construct a Lipschitz continuous unsymmetric operator for the frictional contact problem (24), starting from a ‘quasi-’ rather than a genuine augmented Lagrangian.

3.1. Unilateral contact operator

Problem (22) can be formulated as a classical constrained convex optimisation problem: minimize $\varphi(v)$ over all v such that $d_n(v) \in \mathbb{R}_+^p$, more synthetically written as

$$\inf_{d_n \in \mathbb{R}_+^p} \varphi(v). \quad (25)$$

The standard Lagrangian associated to this problem is

$$L(v, \gamma_n) = \varphi(v) + (d_n(v), \gamma_n), \quad (26)$$

where $(,)$ denotes the scalar product in \mathbb{R}^p and γ_n the Lagrange multipliers (dual variables). Problem (25) may be replaced by the equivalent minimax problem [19]

$$\inf_v \sup_{\gamma_n \in \mathbb{R}_+^p} L(v, \gamma_n). \quad (27)$$

Thus, the constraints have been transferred from the primal variables v to the dual variables γ_n , using \mathbb{R}_+^p the polar cone of \mathbb{R}_+^p in the sense of Moreau. The augmented Lagrangian is more

complicated [16]:

$$L_r(v, \gamma_n) = \varphi(v) - \frac{1}{2r} \|\gamma_n\|^2 + \frac{1}{2r} \text{dist}^2(\gamma_n + rd_n(v), \mathbb{R}_+^p), \quad (28)$$

where r is a positive ‘penalty’ factor and $\text{dist}(x, C)$ denotes the distance from x to C . The augmented Lagrangian includes a Lagrange multiplier term as well as a penalty term and the method is called a mixed or penalty–duality method. Remark that when γ_n is zero, one recovers a penalty method

$$L_r(v, 0) = \varphi(v) + \frac{1}{2}r \text{dist}^2(d_n(v), \mathbb{R}_+^p). \quad (29)$$

In order to obtain additional insight in the augmented Lagrangian expression (28), it is useful to look at the particular case of a single contact element ($p = 1$) and to split the distance term according to the contact status:

$$L_r(v, \gamma_n) = \begin{cases} \varphi(v) - \frac{1}{2r} \gamma_n^2 & \text{if } \gamma_n + rd_n > 0 \quad (\text{gap}), \\ \varphi(v) + \gamma_n d_n(v) + \frac{r}{2} d_n^2(v) & \text{if } \gamma_n + rd_n \leq 0 \quad (\text{contact}). \end{cases} \quad (30)$$

Thus, when contact occurs, the standard augmented Lagrangian [2, 3] for equality constraints is recovered, whereas, when there is a gap, we discover a concave term ensuring the *continuity* (and differentiability) along the border line $\gamma_n + rd_n = 0$. Another expression of L_r uses \mathbb{R}_-^p instead of \mathbb{R}_+^p :

$$L_r(v, \gamma_n) = L(v, \gamma_n) + \frac{r}{2} \|d_n(v)\|^2 - \frac{1}{2r} \text{dist}^2(\gamma_n + rd_n(v), \mathbb{R}_-^p). \quad (31)$$

This alternative form will serve as a reference later. The minimax problem associated to the augmented Lagrangian (28) or (31) turns out to be *unconstrained*:

$$\inf_v \sup_{\gamma_n} L_r(v, \gamma_n). \quad (32)$$

This is a feature of the augmented Lagrangian technique. Moreover, L_r is continuously differentiable with respect to v and γ_n as soon as φ is continuously differentiable with respect to v as already inferred. Consequently, the unique (due to strict convexity of φ) saddle-point (u, λ_n) verifies the stationary condition

$$\nabla_{v, \gamma_n} L_r(u, \lambda_n) = 0. \quad (33)$$

Using (17), (33) can be resolved into two equations

$$\nabla_v L_r(u, \lambda_n) = \nabla \varphi(u) + A_n^t P(\lambda_n + rd_n(u)) = 0, \quad (34a)$$

$$\nabla_{\gamma_n} L_r(u, \lambda_n) = -\frac{1}{r} \{\lambda_n - P(\lambda_n + rd_n(u))\} = 0, \quad (34b)$$

where $P(x) := \text{proj}_{\mathbb{R}_-^p}(x)$ is the projection of x on \mathbb{R}_-^p . From a mechanical standpoint (34a) is

the *equilibrium equation*, and (34b) the *contact equation*. For clarity it is useful to separate the contact part F_n from the structural part G in (34):

$$G(u) + F_n(u, \lambda_n) = 0, \quad (35)$$

where

$$F_n(u, \lambda_n) = \begin{bmatrix} A_n^t P(\lambda_n + r d_n(u)) \\ -\frac{1}{r} \{ \lambda_n - P(\lambda_n + r d_n(u)) \} \end{bmatrix}; \quad (36a)$$

$$(36b)$$

F_n is called a *mixed contact operator*; it is *conewise linear* and *continuous* because of the properties of the projection on a cone. It is instructive to reduce again F_n to one contact element ($p = 1$) as follows:

$$F_n(u, \lambda_n) = \begin{cases} \begin{bmatrix} 0 \\ -(1/r)\lambda_n \end{bmatrix} & \text{if } \lambda_n + r d_n(u) > 0 \quad (\text{gap}), \\ \begin{bmatrix} A_n^t(\lambda_n + r d_n(u)) \\ d_n(u) \end{bmatrix} & \text{if } \lambda_n + r d_n(u) \leq 0 \quad (\text{contact}). \end{cases} \quad (37a)$$

$$(37b)$$

Thus another feature of the augmented Lagrangian technique is that contact is detected with a linear combination of primal and dual variables: $\lambda_n + r d_n$. This is particularly crucial for programming purposes. Linearity regions of F_n are delimited by a *line* as illustrated in Fig. 2. Each region corresponds to a specific contact status (contact, gap) and contains the half-line of the graph of the contact law (7), associated to this status. Note that (34b) conveys the complementary condition $\lambda_n d_n = 0$. Finally, the operator F_n being piecewise linear, it is not differentiable everywhere. At singular points, it possesses a weak notion of derivative called the *generalized Jacobian* [18]. Locally ($p = 1$) it is the convex hull of the classical Jacobians on the adjacent regions: the gap Jacobian J_g and the contact one J_c . The application of this notion is investigated in [20, 21].

$$J_g = \begin{bmatrix} 0 & 0 \\ 0 & -1/r \end{bmatrix}, \quad J_c = \begin{bmatrix} r n n^t & n \\ n^t & 0 \end{bmatrix}. \quad (38)$$

3.2. Pure friction operator

The non-differentiable convex optimisation problem (23) can be treated in almost the same way as (22) upon replacing $d_n(u)$ by $d\delta_t(u)$, λ_n by λ_t , A_n by A_t and \mathbb{R}_-^p by C_n^p . The major difference is that the only expression of the augmented Lagrangian which remains valid in this case is the one analog to (31), because C_n^p is no longer a cone:

$$L_r(v, \gamma_t) = L(v, \gamma_t) + \frac{r}{2} \|d\delta_t(v)\|^2 - \frac{1}{2r} \text{dist}^2(\lambda_t + r d\delta_t(v), C_n^p). \quad (39)$$

The friction operator deriving from this augmented Lagrangian and analog to (36) is

$$F_t(u, \lambda_t) = \begin{bmatrix} A_t^t P_c(\lambda_t + r d\delta_t(u)) \\ -(1/r) \{ \lambda_t - P_c(\lambda_t + r d\delta_t(u)) \} \end{bmatrix}, \quad (40a)$$

$$(40b)$$

where $P_c(x) = \text{proj}_{C_n^p}(x)$. Analytic expressions for this projection are available only for specific forms of the friction convex set C_n , and these are dimension dependent, as illustrated below. *2D anisotropic friction*: The convex set C_n is a segment $[-R_1, R_2]$, where R_1 and R_2 are the non-negative backward and forward thresholds. The projection is trivial. The operator F_t is then piecewise linear. The local linearity regions are drawn in Fig. 3. The local slip contact and adherent contact Jacobians J_s and J_a can be deduced from the gap and contact Jacobians J_g and J_c , respectively, by simply replacing n by t in (38).

3D isotropic friction: C_n is a disk of radius k . The projection may be written as follows:

$$P_c(x) = \text{proj}_c(x) = \begin{cases} x & \text{if } \|x\| \leq k, \\ kx/\|x\| & \text{if } \|x\| > k. \end{cases} \quad (41a)$$

$$(41b)$$

The operator F_t is thus piecewise differentiable and the Jacobian matrices are similar to the 2D expressions, with the unit slip vector equal to

$$t = (\lambda_t + r \, d\delta_t) / \|\lambda_t + r \, d\delta_t\|. \quad (42)$$

3.3. Frictional contact operator

As in the previous cases, the incremental problem (24) can be *formally* written as an optimisation problem:

$$\inf_v \{ \varphi(v) + \psi_{\mathbb{R}_+^p}(d_n(v)) + \psi_{C^p[u]}^*(d\delta_t(v)) \}. \quad (43)$$

Remember that the solution u enters the objective function as a parameter though. To begin with, we define a standard Lagrangian,

$$L(v, \gamma_n, \gamma_t) = \varphi(v) + (d_n(v), \gamma_n) + (d\delta_t(v), \gamma_t). \quad (44)$$

Given the fact that the friction convex set C^p depends on the solution u through the normal contact forces λ_n (identified to dual variables) we obtain a ‘*quasi*’-minimax problem in the form

$$\inf_v \sup_{\gamma_n \in \mathbb{R}_+^p} \sup_{\gamma_t \in C^p(\lambda_n)} L(v, \gamma_n, \gamma_t). \quad (45)$$

At this stage, we postulate a ‘*quasi*’-augmented Lagrangian for the frictional contact problem:

$$\begin{aligned} L_r(v, \gamma_n, \gamma_t) &= \varphi(v) + (d_n(v), \gamma_n) + (d\delta_t(v), \gamma_t) \\ &\quad + \frac{r}{2} \|d_n(v)\|^2 - \frac{1}{2r} \text{dist}^2\{\gamma_n + r d_n(v), \mathbb{R}_+^p\} + \frac{r}{2} \|d\delta_t(v)\|^2 \\ &\quad - \frac{1}{2r} \text{dist}^2\{\gamma_t + r \, d\delta_t(v), \hat{C}[\lambda_n + r d_n(u)]\}, \end{aligned} \quad (46)$$

where, unlike the standard convex $C^p(\lambda_n)$ which is defined only for λ_n belonging to \mathbb{R}_+^p , the ‘*augmented*’ convex set $\hat{C}[\lambda_n + r d_n(u)]$ is defined for all λ_n and d_n as follows:

$$\hat{C}[\lambda_n + rd_n(u)] := C^P \{ \text{proj}_{\mathbb{R}_+^p} [\lambda_n + rd_n(u)] \}. \quad (47)$$

In words, we replace the standard multiplier λ_n by the augmented one $\lambda_n + rd_n(u)$ and project it on \mathbb{R}_+^p . This operation consists in prolongating the multivalued function $C(\lambda_n)$ graph (which is the friction cone) by the positive half-line, as illustrated in Fig. 5 (this operation is rather technical but it is the price to pay for getting an unconstrained problem).

A solution $(u, \lambda_n, \lambda_t)$ of the quasi-minimax problem is then characterized by the system of stationarity conditions:

$$\nabla_v L_r(u, \lambda_n, \lambda_t) = \nabla \varphi + A_n^t P(\lambda_n + rd_n(u)) + A_t^t \hat{P}(\lambda_t + r d\delta_t(u)) = 0, \quad (48a)$$

$$\nabla_{\lambda_n} L_r(u, \lambda_n, \lambda_t) = -\frac{1}{r} \{ \lambda_n - P(\lambda_n + rd_n(u)) \} = 0, \quad (48b)$$

$$\nabla_{\lambda_t} L_r(u, \lambda_n, \lambda_t) = -\frac{1}{r} \{ \lambda_t - \hat{P}(\lambda_t + r d\delta_t(u)) \} = 0, \quad (48c)$$

where \hat{P} is the projection on the ‘augmented’ convex set $\hat{C}(\lambda_n + rd_n)$. Note that (48b) implies that $\hat{C}(\lambda_n + rd_n) = C(\lambda_n)$ at the solution λ_n . The ‘augmented’ projection is crucial to ensure the continuity of the frictional contact operator at hand. Moreover, this operator is conewise linear (CL) in 2D and (provided friction is assumed isotropic) raywise linear (RL) in 3D. As such it is Lipschitz continuous. We refer to [21, 20] for the development of these notions.

Once again, it is instructive to write the contribution of one contact element ($p = 1$) to the global frictional contact operator. We can use this opportunity to obtain an intrinsic formulation by aggregating λ_n and λ_t in a vector $\lambda = \lambda_n n + \lambda_t$ (thus $\nabla_\lambda \lambda_t = I - nn^t$) and derive $L_r(v, \gamma)$ with respect to it. Observe that the unilateral contact and friction constraints are now compacted in a single line. To simplify the expressions we note $\sigma_n = \lambda_n + rd_n$ and $\sigma_t = \lambda_t + r d\delta_t$. The expression of $F(u, \lambda)$ in each case is

$$\begin{bmatrix} 0 \\ -(1/r)\lambda \end{bmatrix} \text{ if } \sigma_n \geq 0 \text{ (gap)}, \quad (49a)$$

$$\begin{bmatrix} \sigma_n n + (I - nn^t)\sigma_t \\ d_n n + (I - nn^t) d\delta_t \end{bmatrix} \text{ if } \|\sigma_t\| + \mu\sigma_n < 0 \text{ (stick)}, \quad (49b)$$

$$\begin{bmatrix} \sigma_n n - \mu\sigma_n(I - nn^t)t \\ d_n n - (1/r)\{(I - nn^t)(\lambda_t + \mu\sigma_n t)\} \end{bmatrix} \text{ if } \|\sigma_t\| + \mu\sigma_n \geq 0 \text{ (slip)}, \quad (49c)$$

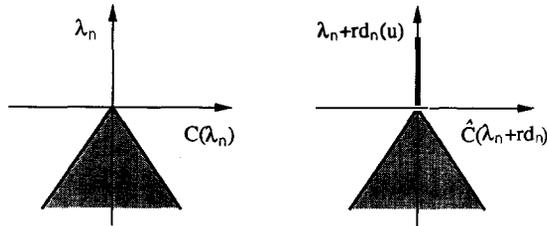


Fig. 5. Prolongation of Coulomb's cone.

where t , defined in (42), is the slip direction unit vector. Due to the identity $(nn^t)t = (n^t)n = 0$, $F(u, \lambda)$ can be further simplified into

$$\begin{bmatrix} 0 \\ -(1/r)\lambda \end{bmatrix} \quad \text{if } \sigma_n \geq 0 \quad (\text{gap}), \quad (50a)$$

$$\begin{bmatrix} \sigma \equiv \sigma_n n + \sigma_t \\ \delta \equiv d_n n + d\delta_t \end{bmatrix} \quad \text{if } \|\sigma_t\| + \mu\sigma_n < 0 \quad (\text{stick}), \quad (50b)$$

$$\begin{bmatrix} \sigma_n(n - \mu t) \\ d_n(n - \mu t) - (1/r)(\lambda_t + \mu\lambda_n t) \end{bmatrix} \quad \text{if } \|\sigma_t\| + \mu\sigma_n \geq 0 \quad (\text{slip}). \quad (50c)$$

The Jacobian matrices associated to the operator (49) are, according to the contact status,

$$J_g = \begin{bmatrix} 0 & 0 \\ 0 & -(1/r)\mathbf{I} \end{bmatrix} \quad \text{gap}, \quad (51a)$$

$$J_a = J_g + \begin{bmatrix} r\mathbf{I} & \mathbf{I} \\ \mathbf{I} & (1/r)\mathbf{I} \end{bmatrix} \quad \text{stick}, \quad (51b)$$

$$J_s = J_g + \begin{bmatrix} rM & M \\ M & (1/r)M \end{bmatrix} \quad \text{slip}, \quad (51c)$$

where $M = (n - \mu t)n^t + \rho (\mathbf{I} - nn^t - tt^t)$ and ρ is a scaling factor defined by

$$\rho = -\mu(\lambda_n + rd_n) / \|\lambda_t + r d\delta_t\| \in [0, 1]. \quad (52)$$

It is checked that the slip contact Jacobian is not symmetric (except if $\mu = 0$). The projection scaling factor ρ comes from a correct linearisation of the projection on a disk [1]. Its influence disappears in 2D discretisation because $\mathbf{I} - nn^t - tt^t = 0$, but we must distinguish backward and forward slip matrices in that case.

3.4. Uniqueness condition

Investigation of the mathematical notions mentioned above (conewise and raywise linearity, generalized Jacobian) leads to a necessary and sufficient condition for *uniqueness* of a solution of the problem (45) [21].

The main conclusion is that positiveness of the determinant of all possible classical Jacobians gives a uniqueness condition on the friction coefficient. Readers interested in a precise statement of these theorems are referred to [20], where their proofs can also be found. To understand the discussion about convergence in the following section, it is helpful however to illustrate this condition on a simple example. Consider the 2D truss depicted in Fig. 6, with its only free node in grazing contact with a rigid plane. K denotes the stiffness matrix of the truss condensed at the free node. Adding the constraint equations, the Jacobian matrix of G defined in (21) is $\nabla_{v,\gamma} G = \begin{bmatrix} K & 0 \\ 0 & 0 \end{bmatrix}$. Then the matrices $\nabla G + J_s$ associated to slip contact status have a positive determinant if and only if

$$\mu < n^t K^{-1} n / |n^t K^{-1} t|. \quad (53)$$

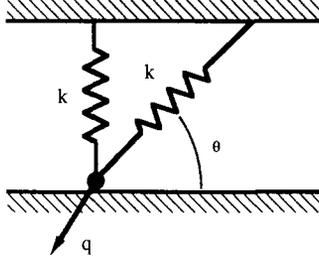


Fig. 6. 2D truss example.

4. Newton type methods for non-smooth operators

The algorithms usually applied to augmented Lagrangian problems in optimisation (e.g. Uzawa's scheme) are stable but very slow because they are based on an alternate treatment of the primal and dual variables. Here, we will show that a simultaneous treatment of both variables by Newton's method converges much faster and that (unlike for penalty based operators [20]) the method remains stable without any damping scheme. We begin our presentation with a modified Newton method, because its stability analysis is simpler, and then we proceed with Newton's method.

To start, note that the contact, friction and frictional contact problems formulated in the previous section are all three governed by a non-linear (non-smooth) equation in the form

$$G(u) + F(u, \lambda) = 0, \quad (54)$$

where G and F are defined in (21)–(35) and (50), respectively.

4.1 Modified Newton algorithm

The basic idea underlying the Zincenko modified Newton method [22] is to resolve (54) in two parts, a differentiable part \tilde{G} and a non-differentiable part \hat{F} ,

$$\tilde{G}(u, \lambda) + \hat{F}(u, \lambda) = 0, \quad (55)$$

where

$$\tilde{G}(u, \lambda) = G(u) - \frac{1}{r} \begin{bmatrix} 0 \\ \lambda \end{bmatrix} \quad \text{and} \quad \hat{F}(u, \lambda) = F(u, \lambda) + \frac{1}{r} \begin{bmatrix} 0 \\ \lambda \end{bmatrix}.$$

The next step is to use the only well defined Jacobian $\tilde{K} = \text{grad } \tilde{G}$, as a slope matrix. If x denotes $(u, \lambda)^t$, Zincenko's method consists in generating the following sequence of approximations:

$$x^{i+1} = x^i - (\tilde{K}_i)^{-1} \{ \tilde{G}(x^i) + \hat{F}(x^i) \}. \quad (56)$$

For the uncoupled problems (22) and (23), this method can be identified with a classical duality algorithm as follows.

4.1.1. Modified Newton algorithm for uncoupled contact and friction problems

Firstly, assume that φ is quadratic, i.e. $\varphi(u) = \frac{1}{2}u^tKu - q^t u$. The contact and friction operators have the same form:

$$\hat{F}(u, \lambda) = \begin{bmatrix} A^t \\ 1/r \end{bmatrix} P_C(\lambda + rAu),$$

where C is the global non-positive orthant cone for the contact operator and the global friction convex set for the friction operator. As \tilde{K} is ‘bloc diagonal’, the i th iteration of Zincenko’s algorithm can be divided into two subiterations identified with Uzawa’s method applied to the standard minimax problem:

Given $\lambda^0 \in C$,

(i) Solve $Ku^i = q - A^t\lambda^i$,

(ii) Project $\lambda^{i+1} = P_C(\lambda^i + rAu^i)$. (57)

If φ is not quadratic but only convex, the above algorithm is a duality method with incomplete resolution of subiteration (i). This equivalence between the two methods provides a stability condition for Zincenko’s method, classically known for Uzawa’s method [19]: the convergence of the Zincenko modified Newton method (56) occurs for r satisfying

$$0 < r < 2\lambda_{\min}(K) / \|A\|, \quad (58)$$

where $\lambda_{\min}(K)$ denotes the smallest eigenvalue of K and $\|A\|$ is the Euclidian matrix norm of A (equal to 1 in our cases). The rate of convergence is often very poor however, because it is governed by the biggest to smallest eigenvalue ratio of K which is very big.

4.1.2. Modified Newton algorithm for frictional contact problem

The algorithm (56) can no longer be identified with a duality method because no classical minimax problem can be associated with it. As a consequence, a convergence condition cannot be given rigorously. However, the particular form of \tilde{K} suggests to decompose the iteration (56) in three subiterations and to classify them:

Given $\lambda_n^0 \in R_-^p$ and $\lambda_t^0 \in C^p(\lambda_n^0)$,

(i) Solve $Ku^i = q - A_n^t\lambda_n + A_t^t\lambda_t^i$,

(ii) Project $\lambda_n^{i+1} = P(\lambda_n^i + rA_n u^i)$,

(iii) Project $\lambda_t^{i+1} = P_{C(\lambda_n^{i+1})}(\lambda_t^i + rA_t u^i)$. (59)

The friction projection in (iii) is effected using the convex set evaluated at the end of subiteration (ii) associated with unilateral contact. This algorithmic hierarchisation of the contact conditions is physically sound.

4.2. Generalized Newton method (GNM)

A rigorous and straightforward extension of Newton's method to non-differentiable but continuous equations such as (54) consists in

$$x^{i+1} = x^i - (K_i + J_i)^{-1} \{G(x^i) + F(x^i)\}, \quad J_i \in \partial F(x^i), \quad (60)$$

where $\partial F(x^i)$ is the generalized Jacobian of F at x^i [18]. This notion is not pursued further because x^i always falls inside a region of linearity in practice and $\partial F(x^i)$ is thus reduced to a classical Jacobian. Unlike the modified algorithm, the GNM cannot be connected to other classical methods for obtaining convergence results. However, a heuristic approach leads to the following conjectures.

4.2.1. GNM applied to uncoupled problems

We ought to distinguish the contact problem and the friction problem. The GNM applied to the unilateral contact problem (22) always converges in practice. Without having a convergence theorem, partial results can be obtained [23]:

- The path of the iterates does not depend on the factor r and the GNM performs better than Lemke's method applied to linear complementarity problems [24]. If p denotes the number of contact elements, the maximum number of iterations to obtain the solution is 2^{p-1} (2^p for Lemke's method [25, 26]) if convergence occurs).
- No simple cycle (order 2) can occur.

When applied to the pure friction problem, the GNM may be unstable. For instance the method may cycle between forward and backward slip (when solution is stick) even on the simple problem as represented in Fig. 7 with $q_t = 0$. In this case, we can determine easily how to choose the factor r to avoid cycling:

$$r < 2k. \quad (61)$$

It is the same condition as (58) for the modified Newton method restricted to this example, for which $\|A\| = 1$. The following inequality is thus proposed to ensure convergence of GNM applied to the friction problem (23).

$$0 < r < 2\lambda_{\min}(K). \quad (62)$$

A study of more complicated examples leads to sharper conditions [23], which validate the relation (62).

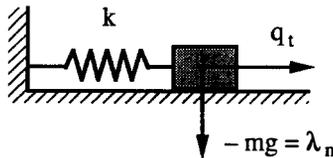


Fig. 7. Elementary friction pad problem.

4.2.2. GNM applied to frictional contact problems

Using a penalty method, we showed in [20] that Newton's method can cycle on simple problems. The order of the cycle increases fastly with the number of contact elements. But the local resolution of the elementary cycle, forward slip–backward slip, is always sufficient in practice. This fact justified the slip reversal control technique presented in [1]. Based on this experience, we limit the present study to a simple model containing only one contact element (master node) to assess the GNM convergence when applied to mixed operators. Reconsider the truss of Fig. 6 for example. The worst loads for convergence are $q_n = n^t q \leq 0$, $q_t = t^t q = 0$. Starting from the backward slip status, the first iterate $x^1 = (u^1, \lambda^1)^t$ satisfies

$$Ku^1 + n^t \lambda^1 (\mu t + n) = q_n n = q, \quad n^t u^1 = 0, \quad t^t \lambda^1 = \mu n^t \lambda^1.$$

Solving for λ_n yields

$$\lambda_n^1 = \left(1 + \mu \frac{n^t K^{-1} t}{n^t K^{-1} n} \right)^{-1} q_n.$$

λ_n is non-positive in view of the uniqueness condition (53). The status associated to x^1 will be stick (and cycling will be prevented) if

$$t^t (\lambda^1 + r u^1) \leq -\mu n^t \lambda^1.$$

This leads to the *sufficient stability conjecture*

$$r \leq 2(t^t K^{-1} t - (n^t K^{-1} t / n^t K^{-1} n)^2)^{-1} = r_{\max}. \quad (63)$$

The following bounds show that the above restriction is a relaxed version of the condition (62) for the friction operator

$$r_{\max} \geq 2(t^t K^{-1} t) \geq 2\lambda_{\min}(K). \quad (64)$$

Moreover, (63) degenerates exactly into (61) for the 2D truss of Fig. 6 because $r_{\max} = 2t^t K t$. Condition (62) seems thus sufficient to enforce convergence of the GNM applied to frictional contact problems. Unlike the damping method and the control technique presented in [1], the above approach reduces to a good choice of the factor r , while preserving the fast rate of convergence of Newton's method. This is the most important characteristic of this method for problem solving.

5. Cylindrical contact problem with friction

The contact between a *disk* and a *bore* of comparable radius with Coulomb's friction at the interface is a good problem to test the capacity of the formulation and algorithm to cope with curved contacts since analytical solutions to the frictionless case [27] and the frictional case [28] are available. The main objectives are to determine the *pressure* and *shear* distributions

$p(\Phi)$ and $q(\Phi)$ at the interface and in particular the half *contact angle* α and the half *adherence angle* β . The numerical solution is obtained with the contact analysis program TACT [12].

5.1. Definition of the problem

a. Analytical data

An elastic disk of unit thickness is pressed, by means of a concentrated load applied along its axis, against a cylindrical bore of almost the same diameter, drilled in a plate of unit thickness too and infinite extent, as shown in Fig. 8. The mechanical and material data are specified besides. The coefficient of friction is taken equal to $\mu = 0.4$.

b. Numerical data

The mesh, which takes advantage of the symmetry is shown in Fig. 9. A total number of 16 contact nodes proves adequate to reach a satisfactory accuracy. Contact nodes between 0 and 60° are evenly distributed (5° apart). The disk and the half space are modeled with 96 and 84 four-node linear isoparametric plane stress elements, respectively. An arc of 21 infinite elements is used to simulate the half-space influence.

5.2. Solution discussion

a. Loading case

The closed form solution found by Klang [28] is rather complicated. It involves integrals which must be evaluated numerically. Instead of reproducing the formulas, we prefer to tabulate and display the pressure and shear distributions for the selected data, cf. Table 1 and

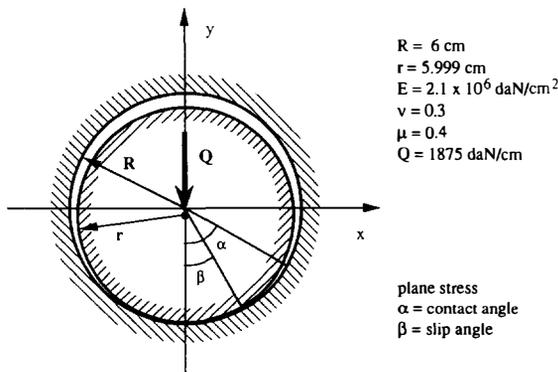


Fig. 8. Cylindrical contact with friction on a disc in a bored plate.

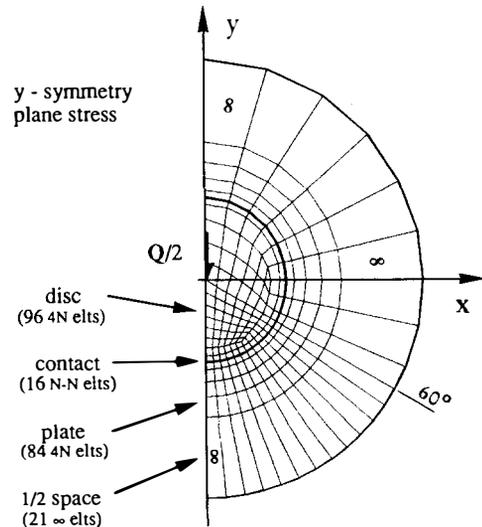


Fig. 9. Finite element mesh.

Table 1
Exact and FEM results to cylindrical contact problem with friction

Ex angle	Ex pressure	Ex press 0.4	Ex shear 0.4	FE angle	FE pressure	FE press 0.4	FE shear 0.4
0	232.89	178.13	0.00	0	235.40	191.70	0.00
3	232.57	178.13	7.00	5	234.50	191.20	7.65
6	231.59	177.50	14.06	10	231.80	189.30	15.10
9	229.62	176.88	21.22	15	227.00	186.30	22.00
12	227.34	175.94	28.50	20	220.10	181.80	26.80
15	224.39	174.70	35.94	25	210.90	175.50	36.50
18	220.46	173.12	43.75	30	199.10	167.60	48.10
21	215.87	171.25	51.88	35	184.60	158.00	55.20
24	210.65	169.06	60.30	40	166.90	145.70	58.30
27	204.43	166.56	66.56	45	145.20	129.40	51.70
30	197.57	163.43	65.31	50	118.40	107.60	43.10
33	189.72	159.37	63.75	55	79.89	73.84	29.60
36	180.89	154.06	61.56	60	6.58	6.20	2.48
39	170.75	147.81	59.06				
42	159.63	140.31	55.94				
45	147.19	130.94	52.50				
48	132.48	119.69	47.81				
51	115.47	105.93	42.50				
54	94.86	88.12	35.31				
57	67.71	63.75	25.46				
60	0.00	0.00	0.00				

Fig. 10, respectively. The pressure distribution is nearly elliptical. The maximum pressure decreases as the coefficient of friction increases, because the shear stresses carry increasing parts of the load. The contact arc is divided into an *inner stick* region and an *outer slip* region (the disc slips ‘down’). The *contact angle* is $\alpha = 60^\circ$ and the *slip angle* $\beta = 26.2^\circ$. Because the contact is *advancing* under a monotonic loading and because friction is path dependent but rate independent, the analytical solution is independent of the loading rate. The numerical solution is tabulated besides the exact one in Table 1 and plotted over it in Fig. 10. The frictionless pressure is in good agreement with the exact one. The pressure and shear distributions are not as accurate when friction is present because the solution is very sensitive to nodal impact location accuracy. In particular, the numerical solution depends on the incremental loading rate, due to the finite accuracy of the impact detection strategy. The solution displayed in Fig. 10 was obtained after applying the load in four uneven increments, following the parabolic law $Q(t) = t^2 Q$, i.e. $Q_n = nQ/16$, $n = 1, \dots, 4$. The contact angle $\alpha \approx 60^\circ$ is well approximated but the slip angle $\beta \approx 37^\circ$ is far from being precise. On the contrary, the algorithm *stability* is excellent. Convergence occurs in a few iterations (less than 5) regardless of the size of the increment and the value of the friction coefficient (the value $\mu = 0.7$ was also used as a check).

b. Unloading case

We do not know of any closed form solution to the unloading *receding* contact problem. The numerical solution shown in Fig. 11 corresponds to a load equal to 75% of the original

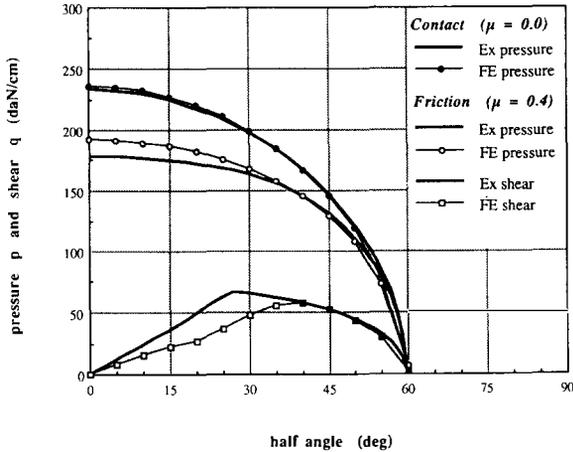


Fig. 10. Pressure and shear distributions.

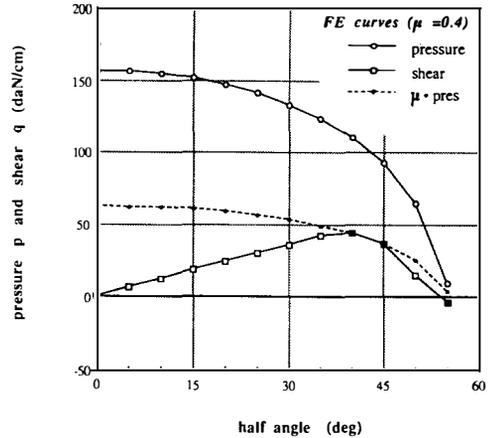


Fig. 11. Pressure and shear distributions after unloading to 3/4.

load. For this moderate unloading the contact angle decreases from $\alpha = 60^\circ$ to $\alpha = 55^\circ$. The inward slip angle $\beta \approx 37^\circ$ is virtually unchanged. The solution is characterized by another *upward (opposite) slip region* at the contact edge, in analogy with flat and parabolic punch problems solutions [1]. This pattern subsists if the load is further decreased, but the upward slip band reduces to a negative shear zone. The algorithm capacity to cope with severe unloading is verified.

Conclusions

In this article, a mixed *penalty–duality* formulation of the frictional contact problem, inspired from the augmented Lagrangian approach has been proposed. The *continuity* of the resulting *conewise linear* operator has been used to establish a uniqueness condition on the coefficient of friction. Modified and *generalized Newton* methods have been examined and sufficient conditions for their *convergence* conjectured. The cylindrical frictional contact problem of a disk in a bore has shown the stability of the generalized Newton method.

In our opinion, mixed penalty–duality methods are so much more *robust* than primal penalty methods that it is worthwhile undertaking the extra complexity inherent to their formulation. Indeed, mixed penalty–duality methods are both more accurate and stabler than penalty methods. Moreover, contrary to a wide-spread idea, they do not require more storage if that needed to stabilize the penalty algorithm is taken into account. More specifically, the place necessary to store the additional dual multiplier unknowns is not larger than the place taken by the extra primal history variables necessary to stabilize the penalty method, whether they are previous displacements vectors as in a damping method or past slip hidden variables as with the slip reversal control technique [1]. The next step will be to extend these formulations to large slip problems using node-on-facet elements.

Note added in proof

In their recent book of 1988, Kikuchi and Oden [29] give a fairly extensive bibliography on the subject, albeit omitting some major references of the European school such as Moreau, Panagiotopoulos, Klarbring or Ciarlet-Nécas. Therein, they mention (and we agree!) that the frictional contact problem between continuous deformable solids involves formidable mathematical difficulties. They add further on that many of the techniques developed to solve the corresponding discrete problems are based on Lagrange multipliers or penalty methods. Following Martins, they advocate normal compliance models which correspond to penalty methods and which have been suggested by Curnier [30] as early as 1984, extensively used afterwards [12, 1] and recently dropped for their lack of numerical robustness [1, 20]. In their bibliography, Kikuchi and Oden do not mention *augmented Lagrangian* formulations however, nor do they use them in their own treatment of the subject.

If the augmented Lagrangian technique can now be considered as straightforward for the frictionless contact problem [31], its application to the pure friction problem with given constant pressure is not so obvious and the augmented multiplier technique introduced in this paper for the non-associated frictional contact problem with unknown pressure is believed to be original and powerful. Since the submission of the present paper, it has been extended with success to 2D and 3D *large slip* discrete contact problems, without friction in the case of two deformable bodies and with friction in the case of a deformable body against a rigid obstacle.

References

- [1] A. Curnier and P. Alart, A generalized Newton method for contact problem with friction, *J. Méc. Théor. Appl.*, Special Issue: Numerical Methods in Mechanics of Contact Involving Friction (1988) 67–82.
- [2] M. Hestenes, Multiplier and gradient method, *J. Opt. Th. Appl.* 4 (1969) 303–320.
- [3] B.T. Powell, A method for nonlinear constraints in minimization problems, in: R. Fletcher, ed., *Optimization* (Academic Press, London, 1969).
- [4] R.T. Rockafellar, Augmented Lagrangians and applications of the proximal point algorithm in convex programming, *Math. Oper. Res.* 1 (2) (1976).
- [5] M. Fortin, Minimization of some non-differentiable functionals by the augmented Lagrangian method of Hestenes and Powell, *Appl. Math. Optim.* 2 (4) (1976) 236–250.
- [6] J.A. Landers and R.L. Taylor, An augmented Lagrangian formulation for the finite element solution of contact problems, Rept. No. UCB/SESM-85/09, Department of Civil Engineering, University of California, Berkeley, 1985.
- [7] F.J. Barthold and D. Bischoff, Generalization of Newton type methods to contact problems with friction, *J. Méc. Théor. Appl.*, Special Issue: Numerical Methods in Mechanics of Contact Involving Friction (1988) 97–110.
- [8] J.J. Moreau, Application of convex analysis to some problems of dry friction, in: Zorski ed., *Trends of Pure Mathematics Applied to Mechanics*, Vol. II (1979).
- [9] M.L. Wilkins, Calculation of elastic-plastic flow, in: Berni Alder et al., eds., *Methods of Computational Physics*, Vol. 3, *Fundamental Methods in Hydrodynamics* (Academic Press, New York, 1964) 211–263.
- [10] J. Simo and R.L. Taylor, Consistent tangent operators for rate-independent elastoplasticity, *Comput. Methods Appl. Mech. Engrg.* 48 (1985) 101–118.
- [11] M. Jean, Un algorithme numérique simple pour un système d'oscillateurs avec frottement de Coulomb sur un plan, Université des Sciences et Techniques du Languedoc, Lab. de Mécanique Générale et des Milieux Continus, Note Technique 85-2 (Montpellier, 1985).

- [12] A. Curnier, TACT: A contact analysis program, in: Del Piero-Maceri, ed., Proceedings of 3rd meeting on Unilateral Problems in Structural Analysis (1987) 97–116.
- [13] R. Taylor, J. Sackman, W. Kanoknukulchai and A. Curnier, A finite element formulation and solution of contact-impact problems in continuum mechanics, V, Report 78-3, CE-SESM, University of California, Berkeley, 1978.
- [14] C. Simo, P. Wriggers, N. Schweizerhof and R.L. Taylor, Finite deformation post-buckling analysis involving inelasticity and contact constraints, *Internat. J. Numer. Methods Engrg.* 23 (1986) 779.
- [15] G. Duvaut and J.L. Lions, *Les inéquations en mécanique et en physique* (Dunod, Paris, 1972).
- [16] R.T. Rockafellar, Lagrange multipliers and variational inequalities, in: Gianessi, Cottle and Lions, eds., *Variational Inequalities and Complementarity Problems* (Wiley, New York, 1979).
- [17] I. Capuzzo Dolcetta and U. Mosco, Implicit complementarity problems and quasi-variational inequalities, in: Gianessi, Cottle and Lions, eds., *Variational Inequalities and Complementarity Problems* (Wiley, New York, 1979).
- [18] F.H. Clarke, *Optimization and Nonsmooth Analysis* (Wiley, New York, 1983).
- [19] I. Ekeland and R. Temam, *Analyse convexe et problèmes variationnels* (Dunod, Collection Etudes Mathématiques, Paris, 1978).
- [20] P. Alart and A. Curnier, Contact discret avec frottement, document LMA-DME-EPFL, Lausanne, 1987.
- [21] P. Alart, Caractérisation des homéomorphismes non différentiables en dimension finie, in preparation.
- [22] A.I. Zincenko, On a method for approximate solution of equations with nondifferentiable operators, *Dop. Akad. Nank. Ukr. RSR*, No. 2 (1963) 156–161.
- [23] P. Alart, Multiplicateurs “augmentés” et méthode de Newton généralisée pour contact avec frottement, document LMA-DME-EPFL, Lausanne, 1988.
- [24] C.E. Lemke, Bimatrix equilibrium points and mathematical programming, *Management Sci.* 4 (1965) 681–689.
- [25] K.G. Murty, Computational complexity pivot methods, *Math. Prog. St.* 7 (1978) 61–73.
- [26] Y. Fathi, Computational complexity of LCP’s associated with positive definite symmetric matrices, *Math. Prog. St.* 17 (1979) 335–344.
- [27] A. Persson, On the stress distribution of cylindrical elastic bodies in contact, Chalmers University, Gothenburg, 1964.
- [28] M. Klang, On interior contact under friction between cylindrical elastic bodies, Thesis, Linköping University, Linköping, 1979.
- [29] N. Kikuchi and J.T. Oden, *Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Methods* (SIAM, Philadelphia, PA, 1988).
- [30] A. Curnier, A theory of friction, *Internat. J. Solids and Structures* 20 (7) (1984) 637–647.
- [31] R. Glowinski and P. Le Tallec, *Augmented Lagrangian and Operator Splitting Methods in Nonlinear Mechanics* (SIAM, Philadelphia, PA, 1989).