Computers & Structures Vol. 42, No. 1, pp. 97–116, 1992
Printed in Great Britain.

0045-7949/92 $5.00 + 0.00
Pergamon Press plc

AN AUGMENTED LAGRANGIAN TREATMENT OF CONTACT PROBLEMS INVOLVING FRICTION

J. C. SIMO and T. A. LAURSEN
Division of Applied Mechanics, Department of Mechanical Engineering, Stanford University, Durand Building, Stanford, CA 94304-4040, U.S.A.

(*Received* 5 *September* 1990)

Abstract—A framework is presented within which the method of *augmented Lagrangians* is readily applied to problems involving contact with friction. This method, which has enjoyed considerable success in the treatment of constrained minimization problems, has been previously applied to problems involving incompressible flow, incompressible elasticity of solids and even *frictionless* contact. An additional challenge to the method is provided by frictional contact problems governed by a Coulomb law, due to the special form taken by the frictional constraint. This paper describes a new extension of the augmented Lagrangian technique to frictional problems which is well-suited to finite element implementation. The proposed treatment inherits the traditional advantages of augmented Lagrangian techniques over penalty methods; namely, *decreased ill-conditioning* of governing equations, and *essentially exact* satisfaction of constraints with *finite* penalties. A set of numerical examples is presented in which the utility of the method is demonstrated even in the presence of finite deformations and inelasticity.

1. INTRODUCTION

The method of augmented Lagrangians, originally proposed by Hestenes [1] and Powell [2] in the context of mathematical programming problems subject to *equality* constraints, has been known for years to provide important advantages over the more traditional Lagrange multiplier and penalty methods. Extensions of the method of augmented Lagrangians to mathematical programming problems involving *inequality* constraints are also well established and go back to work of Rockefeller [3] and others; see e.g., the summary accounts in Bertsekas [4] and Fletcher [5]. More recently, within the context of finite element methods, augmented Lagrangian techniques have been successfully applied to incompressible finite deformation elasticity [6, 8], frictionless contact problems (e.g., [9, 10]), and viscoplasticity [11]. All these problems share a common characteristic; namely, a certain key constraint present in each problem is conveniently enforced by a penalization. The advantages of the penalty approach are obvious: the technique is simple, introduces no additional equations, and is readily interpreted from a physical standpoint. Unfortunately, it is also well-known that penalty methods suffer from ill-conditioning that worsens as penalty values are increased, while constraints are satisfied exactly *only* in the limit of infinite penalty values (see [4] for a more concrete discussion of these ideas). Thus, for many problems it may be desirable or even necessary to consider the augmented Lagrangian technique as an alternative approach capable of circumventing these difficulties.

Contact problems with friction in solid mechanics, the subject of the present investigation, constitute physical examples of *variational inequalities*; see e.g., the classical work of Duvaut and Lions [12] and the recent monograph of Kikuchi and Oden [13]. While instructive, such inequalities are not amenable to most finite element implementations. Typically, Lagrange multipliers or penalty regularizations are introduced to reduce the problem to a variational *equality* which can then be handled using traditional finite element methods, as in [14]. See [13] for a comprehensive review of these alternative approaches.

Indeed, the use of a penalization to accomplish this task is especially attractive in the case of frictional contact, because the resulting equations suggest a 'constitutive law' for the interface which is almost exactly analogous to those traditionally used in the theory of plasticity [15]. Return mapping schemes essential for integrating such equations are also well-known [16]. Examples of implementations using a penalty regularization of this type are to be found in [17] and [18].

In this paper, the natural extension of these ideas to an augmented Lagrangian framework will be given. For simplicity, the friction law considered will be a Coulomb law, with no distinction made between static and kinematic coefficients of friction. While it is recognized that mathematical and empirical difficulties exist with such a characterization [13], it still enjoys a great deal of engineering utility in the opinion of the authors.

The description will be given in the following manner. Sec. 2 will describe the frictionless contact problem in the context of a finite-deformable body in contact with a rigid obstacle. Although essentially a review of known results, this section conveniently provides a foundation for Sec. 3,

in which the frictional contact problem will be posed and treated in the context of a rigid obstacle problem in small deformations. This section clarifies the relation between return mapping algorithms for frictional problems of the type employed in plasticity (see Simo and Hughes [7] for a review), and alternative two-step algorithms of the type reviewed in [13]. New augmented Lagrangian algorithms are developed within these two approaches. While the latter class of algorithms have the attractive feature of leading to symmetric equation systems, the former appears to exhibit more robust performance, particularly in highly nonlinear problems. It is emphasized that the small deformation assumption made in Sec. 3 is only introduced for the sake of convenience; the augmented Lagrangian framework developed therein is readily extended to finite deformations. In fact, Sec. 4, which consists of numerical examples demonstrating the method, is comprised mostly of finite deformation simulations. Details pertaining to the implementation of the proposed methodology are given in two appendices.

2. AUGMENTED LAGRANGIAN TREATMENT OF FRICTIONLESS CONTACT

In the following we outline the treatment of frictionless contact by consideration of the rigid obstacle problem in finite deformations as a model problem. By carefully stating the problem and motivating the augmented Lagrangian treatment through presentation of traditional Lagrange multiplier and penalty techniques, we shall provide a framework within which frictional effects are conveniently introduced.

2.1. *Statement of the obstacle problem in finite deformations*

We consider in this section the problem of finite deformation of a continuum body constrained by the presence of a rigid, immovable obstacle. We shall denote material points in the reference configuration Ω (an open subset of either $\mathbb{R}^2$ or $\mathbb{R}^3$) by $\mathbf{X}$ (see Fig. 1). Points in the current configuration are given by $\mathbf{x} = \varphi(\mathbf{X}, t)$, where φ has the property that $\det(D\varphi(\mathbf{X}, t)) > 0$ for all t, where t is the time variable (which shall hereafter be suppressed). We denote by Γ a section of $\partial\Omega$ which we shall consider to include all prospective points of contact, and γ shall be the image of Γ over φ. Lastly, we define $\mathbb{K}$, an open subset of the ambient space which together with $\partial\mathbb{K}$ comprises the admissible region for the motion of Ω. The remainder of the ambient space is then considered to be occupied by the rigid obstacle. We assume that $\mathbb{K}$ is invariant with respect to time.

We next consider a scalar-valued *gauge function h*, defined on the spatial domain, which has the property that $h < 0$ in $\mathbb{K}$, $h = 0$ on $\partial\mathbb{K}$, and $h > 0$ outside of $\mathbb{K}$. It is assumed for the present that the set $\mathbb{K}$ is convex. While convenient for the development which follows, this restriction is not a major consideration in actual implementations, as we shall later see. The specific form of the gauge function is not crucial to what follows; we simply emphasize that for any admissible point $\mathbf{x}$ of the spatial domain, $h \leqslant 0$.

With this notation in hand, we are ready to state the contact conditions:

For all $\mathbf{X} \in \Gamma$, the admissible deformation $\mathbf{x} = \varphi(\mathbf{X}, t)$ satisfies:

$$h(\mathbf{x}) \leqslant 0, \tag{2.1.1}$$

$$t_N = -\mathbf{n}(\mathbf{x}) \cdot \mathbf{P}\mathbf{N} \geqslant 0, \tag{2.1.2}$$

$$t_N(\mathbf{x})h(\mathbf{x}) = 0, \tag{2.1.3}$$

$$\mathbf{t}_T = \mathbf{P}\mathbf{N} + t_N\mathbf{n} = \mathbf{0}, \tag{2.1.4}$$

where

$\mathbf{P}$:=first Piola—Kirchhoff stress tensor
$\mathbf{n}$:=outward normal in the current configuration
$\mathbf{N}$:=outward normal in the reference configuration.

It is noted that eqn (2.1.1) represents the impenetrability condition, (2.1.2) represents the restriction that the normal component of surface traction be compressive (note the sign change in the definition of t_N), and (2.1.3) is a condition ensuring that t_N

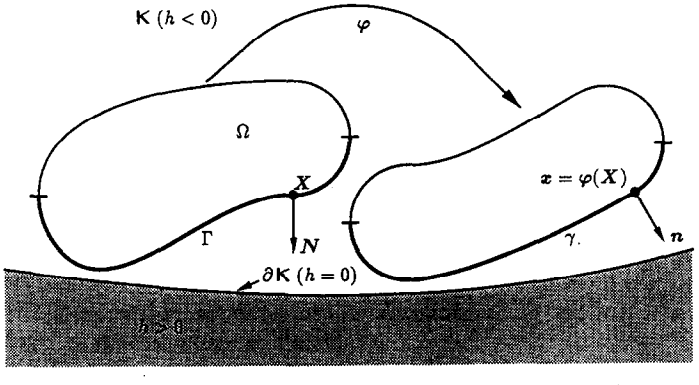

Fig. 1. Notation for the obstacle problem in finite deformations.

may only be nonzero when $h(\mathbf{x}) = 0$. Equations (2.1.1)–(2.1.3) are therefore recognized as the familiar Kuhn–Tucker conditions. Equations (2.1.4) merely asserts that no friction is present.

Accompanying the contact conditions are the remainder of the governing equations

$$\begin{aligned} &\text{Div } \mathbf{P} + \mathbf{f} = \mathbf{0} \text{ in } \Omega, \\ &\mathbf{P}\mathbf{N} = \bar{\mathbf{t}} \text{ on } \Gamma_\sigma, \\ &\varphi = \bar{\varphi} \text{ on } \Gamma_\varphi, \end{aligned} \tag{2.2}$$

where Γ_σ and Γ_φ are the portions of $\partial\Omega$ on which the traction and motion, respectively, are prescribed.

Remark 2.1. In this statement we make *no restriction* on the constitutive law governing **P**. Thus, this statement of the problem allows for either elastic or inelastic response.

We next consider the variational formulation of this problem. In doing so, however, we first remark that admissible variations of the motion are *constrained* by the contact conditions. This restriction takes the form

$$\delta\varphi(\mathbf{X}) \cdot \mathbf{n}(\varphi(\mathbf{X})) \leqslant 0 \quad \text{on} \quad \Gamma \quad \text{if} \quad h(\varphi(\mathbf{X})) = 0, \tag{2.3}$$

where $\delta\varphi(\mathbf{X})$ is an admissible variation of the deformation (implying $\delta\varphi = 0$ on Γ_φ).

Keeping this fact in mind, we convert the strong form of the problem (2.2) into the weak form in the usual manner to obtain

$$\int_\Omega \mathbf{P} \cdot \text{Grad}[\delta\varphi]\, d\Omega - \int_\Omega \mathbf{f} \cdot \delta\varphi \, d\Omega - \int_{\Gamma_\sigma} \bar{\mathbf{t}} \cdot \delta\varphi \, d\Gamma - \int_\Gamma \mathbf{t} \cdot \delta\varphi \, d\Gamma = 0, \tag{2.4}$$

which must hold for all $\delta\varphi$ with $\delta\varphi = 0$ on Γ_φ satisfying (2.3).

We assert that (2.4) is difficult to work with in a finite element setting precisely due to the constraint (2.3) on the admissible variations. For this reason, we shall consider Lagrange multiplier and penalty methods in the next subsection. However, we can further investigate the nature of the variational problem by noting

$$\mathbf{t} \cdot \delta\varphi = -t_N \delta\varphi \cdot \mathbf{n} \quad \text{on} \quad \Gamma. \tag{2.5}$$

Use of (2.3) together with (2.5) leads to

$$\int_\Gamma \mathbf{t} \cdot \delta\varphi \, d\Gamma \geqslant 0, \tag{2.6}$$

which gives

$$G(\varphi, \delta\varphi) := \int_\Omega \mathbf{P} \cdot \text{Grad}[\delta\varphi]\, d\Omega - \int_\Omega \mathbf{f} \cdot \delta\varphi \, d\Omega - \int_{\Gamma_\sigma} \bar{\mathbf{t}} \cdot \delta\varphi \, d\Gamma \geqslant 0, \tag{2.7}$$

which again must hold for all $\delta\varphi$ such that $\delta\varphi = 0$ on Γ_φ and (2.3) is satisfied.

It is noted that eqn (2.7) is a variational inequality precisely of the type considered, for example, in [13].

2.2. *Lagrange multiplier and penalty formulations*

In order to remove the rather inconvenient restriction on the variations given by (2.3), we first consider a formulation in which λ_N, an additional variable, is introduced over Γ. We demand that the following conditions be satisfied over Γ

$$\begin{aligned} &\lambda_N \geqslant 0, \\ &h(\varphi(\mathbf{X})) \leqslant 0, \\ &\lambda_N h(\varphi(\mathbf{X})) = 0. \end{aligned} \tag{2.8}$$

If one considers (2.2) to define the strong form of the boundary value problem, subject to constraints (2.8), the following variational equations, corresponding to the *Lagrange multiplier formulation*, result

$$G(\varphi, \delta\varphi) + \int_\Gamma \lambda_N \delta\varphi \cdot \mathbf{n}(\varphi(\mathbf{X}))\, d\Gamma = 0, \tag{2.9}$$

$$\int_\Gamma [\delta\lambda_N] h(\varphi(\mathbf{X}))\, d\Gamma = 0, \tag{2.9.1}$$

which must hold for all admissible $\delta\varphi$ and $\delta\lambda_N$, where $\delta\lambda_N \geqslant 0$, is the variation of λ_N.

We note that the variation of φ is unconstrained by the contact conditions in this case. However, this comes at the cost of an additional variable and indefinite structure of the resulting matrix problem. Although not insurmountable, these difficulties motivate the consideration of a *penalty regularization*.

We begin this development by defining g, a function over the spatial domain which shall hereafter be referred to as the *gap function* (see Fig. 2). For any point $\mathbf{x}$ in the spatial domain, we define $g(\mathbf{x})$ as follows:

$$g(\mathbf{x}) := \|\mathbf{x} - \bar{\mathbf{x}}\| = \min_{y:h(y)\leqslant 0} \|\mathbf{x} - \mathbf{y}\|. \tag{2.10}$$

Expressed in words, $\bar{\mathbf{x}}$ is the *closest point projection* of $\mathbf{x}$ onto the admissible region, and its determination

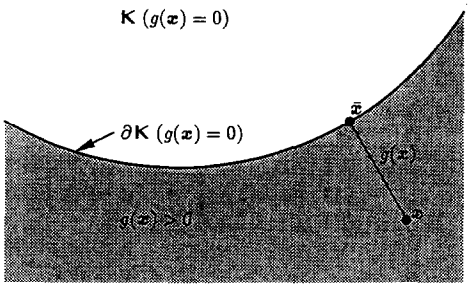

Fig. 2. Definition of the gap function on the spatial domain.

in turn defines $g(\mathbf{x})$. Note that if $\mathbf{x}$ is admissible, $\bar{\mathbf{x}} = \mathbf{x}$. As a consequence of this construction, $g(\mathbf{x}) \geqslant 0$ is nonzero *only* if $\mathbf{x}$ is inadmissible.

Remark 2.2. This definition of the gap function demonstrates why the assumption of convexity of $\mathbb{K}$ was made. As we shall see in a moment, the case in which $\mathbf{x}$ is inadmissible is the only case of interest with regard to the penalty regularization. If the admissible region were not convex, the definition of $\bar{\mathbf{x}}$ could be nonunique. In practice, we are really only concerned with *local* convexity, and steps can be taken in the algorithmic setting to remedy the situation when even this weaker condition is violated. See Appendix A for further elaboration.

Remark 2.3. It is important to note that although the *gauge function h* may be defined such that it is identical to *gap function g* in the inadmissible region, in general this need not be the case. For example, one might conceive of situations in practice where it is advantageous to use a gauge function h distinct from g in order to define the searching algorithm which detects contact. Subsequently, g could be used in the finite element equations to characterize the contact conditions once contact is detected. In the present case, the introduction of g converts the inequality constraint $(h \leqslant 0)$ to an equality constraint $(g = 0)$ on Γ.

With this definition in hand, the penalty regularization is achieved by the replacement of eqns (2.8) by the following

$$\lambda_N := \epsilon_N g(\varphi(\mathbf{X})) \quad \text{on} \quad \Gamma. \tag{2.11}$$

Remark 2.4. ϵ_N is known as the *penalty parameter*. As $\epsilon_N \to \infty$, $g \to 0$ and λ_N is bounded. Thus, as $\epsilon_N \to \infty$, the constraint is increasingly well-satisfied.

Remark 2.5. Note that (2.11) is essentially a (*Yoshida*) *regularization*. The structure of the regularization here is remarkably similar to that of the viscoplastic regularization used in treatments of rate dependent plasticity (see [7] for further discussion).

The variational equation for the penalty method is now easily obtained by substitution of (2.11) into (2.9.1)

$$G(\varphi, \delta\varphi) + \int_\Gamma \epsilon_N g(\varphi(\mathbf{X}))\delta\varphi \cdot \mathbf{n}(\varphi(\mathbf{X}))\, d\Gamma = 0, \tag{2.12}$$

which must hold for all $\delta\varphi$ such that $\delta\varphi = 0$ on Γ_φ. We note that (2.12) now only involves the variable φ, and no constraints of the nature of (2.3) are present on the admissible variations. These facts make (2.12) extremely attractive for finite element implementations. However, we know that the constraint $h(\varphi(\mathbf{X})) \leqslant 0$ on Γ is satisfied only in the limit as $\epsilon_N \to \infty$, and we further recall the previously-discussed difficulty that ill-conditioning increases as ϵ_N increases. It is these considerations that lead us to turn to the *method of augmented Lagrangians*.

2.3. *Augmented Lagrangian formulation*

The concept of the method is remarkably simple. Starting with the variational equation (2.9.1), we append a penalty regularization which renders the following

$$G(\varphi, \delta\varphi) + \int_\Gamma (\lambda_N + \epsilon_N g(\varphi(\mathbf{X})))\delta\varphi \cdot \mathbf{n}(\varphi(\mathbf{X}))\, d\Gamma = 0. \tag{2.13}$$

We note that (2.13) is a penalization of the Lagrange multiplier problem which is exact if the multipliers are the correct ones [corresponding to the solution of (2.9.1)]. We can see this as follows. If λ_N is the correct multiplier, then $g = 0$ on Γ. Thus, in the case where the multipliers are correct, (2.13) attains exactly the same form as (2.9.1), making it an exact penalization (see [19]).

The crucial idea in the method of augmented Lagrangians is to regard λ_N as a *fixed current estimate* of the correct Lagrange multiplier, and solve the problem

$$G(\varphi, \delta\varphi) + \int_\Gamma \langle \lambda_N^{(k)} + \epsilon_N g(\varphi(\mathbf{X})) \rangle \, \delta\varphi \cdot \mathbf{n}(\varphi(\mathbf{X}))\, d\Gamma = 0, \tag{2.14}$$

where $\lambda_N^{(k)} \geqslant 0$ denotes the *fixed* estimate of the correct λ_N. The superscript $(\cdot)^{(k)}$ reflects the fact that the search for the correct λ_N is an *iterative* process. One notes that the term $\langle \lambda_N^{(k)} + \epsilon_N g(\varphi(\mathbf{X})) \rangle$ plays the role of the exact Lagrange multiplier in (2.14). One would suppose, then, that it is a good approximation to the correct multiplier, which motivates the update formula

$$\lambda_N^{(k+1)} = \langle \lambda_N^{(k)} + \epsilon_N g \rangle. \tag{2.15}$$

Remark 2.6. $\langle\cdot\rangle$ is called the Macauley bracket, defined as $\langle x\rangle = \frac{1}{2}[x + |x|]$. Its appearance in (2.14) is consistent with the interpretation of $\lambda_N^{(k)} + \epsilon_N g$ as the normal contact pressure, which must be positive [see (2.1.2)].

Remark 2.7. A slight change in the definition of g is made at this juncture, which accounts for the slight difference in form between (2.13) and (2.14). According to the definition of g made in the context of the penalty regularization, $g \geqslant 0$ is nonzero only in the inadmissible region. The effect of this, as can be seen in (2.12), is that a point of Γ only has a contribution to the integral over the contact surface if its current mapping lies in the inadmissible region. Examination of (2.14) will show that this is not necessarily the case in the augmented Lagrangian formulation due to the presence of λ_N. In fact, one observes from (2.15) that if the old definition for g were retained, the update formula for λ_N would dictate that $\lambda_N^{(k+1)}$ be unamended (and perhaps nonzero) even if $\mathbf{x}$ were admissible. Thus, in order to conform to the augmented Lagrangian recipe for treatment of inequality constraints essentially due to Rockefeller [3] (see, e.g., [19, 5]), we now allow $g(\mathbf{x})$ to be a signed quantity, with magnitude equal to the minimum distance between $\mathbf{x}$ and $\partial\mathbb{K}$. We take $g(\mathbf{x})$ to be positive if $\mathbf{x}$ is inadmissible, and negative otherwise. With this definition, it is noted that g is now a suitable choice for the gauge function h, and it is this definition of g that is used in (2.14). Since this intricacy in the definition of g is a fairly minor point in practice, the old notation for g will be retained, but the reader should bear in mind the changed interpretation in the current context.

It is important to notice that (2.14) is a nonlinear equation due to the contact conditions, geometric nonlinearity, and (perhaps) inelasticity. In general, then, it will be necessary to solve (2.14) in an iterative manner. One can easily envision two different solution schemes; one in which update (2.15) is performed concurrently with the iterations necessary to solve (2.14), and another in which (2.14) is solved completely, before update (2.15) is performed. In this latter scheme, (2.14) and (2.15) are solved recursively (and completely separately) until convergence is attained. The first approach, which we shall denote *simultaneous iteration*, is the one considered for solution of incompressible Navier–Stokes equations by Fortin and Fortin [21], and has subsequently been considered for frictionless contact by [10] and by [9]. The second approach, which we shall refer to as *nested iteration*, is more closely related to ideas advanced independently by Hestenes [1] and Powell [2], and will be the approach focused upon in this paper. Although both techniques have advantages, the nested scheme has the attractive property of preserving quadratic convergence of the inner loop [i.e., solution of (2.14)] when a Newton–Raphson solution scheme is utilized.

Table 1. Nested augmented Lagrangian algorithm for frictionless contact

1. Initialization:

   set $\lambda_N^{(0)} = \langle \lambda_N + \epsilon_N g\rangle$ from the last time step,
   $k = 0$.

2. Solve (using a nonlinear solution strategy) for $\varphi^{(k)}$:

$$G(\varphi^{(k)}, \delta\varphi) + \int_\Gamma \langle \lambda_N^{(k)} + \epsilon_N g(\varphi^{(k)}(\mathbf{X}))\rangle \, \delta\varphi \cdot \mathbf{n}(\varphi^{(k)}(\mathbf{X}))\, d\Gamma.$$

3. Check for constraint satisfaction:

   IF ($g(\mathbf{x}) \leqslant$ TOL for all $\mathbf{x} \in \gamma$) THEN
   Converge. EXIT.
   ELSE
   Augment:
   $\lambda_N^{(k+1)} = \langle \lambda_N^{(k)} + \epsilon_N g(\varphi^{(k)})\rangle$
   $k \leftarrow k + 1$
   GOTO 2.
   ENDIF

Remark 2.8. In the schemes considered in this paper, ϵ_N is *fixed* throughout the procedure (although this need not be the case). In practice, ϵ_N is chosen to be as large as practical without inducing ill-conditioning. The advantage of the current treatment over the penalty method is that satisfaction of the constraints can be improved even if ϵ_N is undersized through repeated application of the augmentation procedure. Since these augmentations only change $\lambda_N^{(k)}$, which is fixed with regard to solution of (2.14), no additional ill-conditioning of the resulting matrix problem is induced.

To close the section, we present the nested augmented Lagrangian algorithm for frictionless contact in Table 1.

3. AUGMENTED LAGRANGIAN TREATMENT OF CONTACT WITH FRICTION

In this section we examine the extension of the framework developed in Sec. 2 to accommodate frictional effects. In order to simplify the presentation, we shall limit the discussion in this section to the rigid obstacle problem in *small* deformations. The ideas developed herein, however, are readily extendible to large deformations and contact of multiple deformable bodies, as shall be demonstrated in Sec. 4 and in the appendices.

3.1. *Statement of the frictional obstacle problem in small deformations*

We now consider the deformation of a body constrained by a rigid, immovable obstacle, but assume

that the displacements of points in the body are very small. We denote points in the body Ω (an open subset of either $\mathbb{R}^2$ or $\mathbb{R}^3$) as $\mathbf{x}$ (see Fig. 3). As in Sec. 2, we define $\mathbb{K}$ as the interior of the admissible region, and we designate Γ as a portion of $\partial\Omega$ containing all prospective points of contact. We denote the displacement field over Ω by $\mathbf{u}$, and mention that although $\mathbf{u}$ is in general time dependent, the time argument will be suppressed except when needed. Outward normals to the body Ω shall be denoted as $\mathbf{n}(\mathbf{x})$.

We shall state the normal contact conditions slightly differently in the current context than was done earlier by a direct appeal to the concept of a (signed) gap function. Suppose a point $\mathbf{x}$ in Γ has an an initial distance $g_0 \geqslant 0$ from its closest point on $\partial\mathbb{K}$ (we assume all points $\mathbf{x}$ are in the admissible region initially). Then g_0 is a scalar-valued function defined over Γ. The contact conditions in the linear theory are now given for all $\mathbf{x} \in \Gamma$ by:

$$\begin{aligned} g(\mathbf{u}) &:= \mathbf{u}\cdot\mathbf{n} - g_0(\mathbf{x}) \leqslant 0, \\ t_N(\mathbf{u}) &:= -\mathbf{n}\cdot\boldsymbol{\sigma}(\mathbf{u})\mathbf{n} \geqslant 0, \\ t_N(\mathbf{u})g(\mathbf{u}) &= 0. \end{aligned} \tag{3.1}$$

Analogous to (2.2) in the previous section, the statements of local momentum balance and boundary condition prescription take the form

$$\begin{aligned} \operatorname{div}\boldsymbol{\sigma} + \mathbf{f} &= \mathbf{0} \text{ in } \Omega, \\ \boldsymbol{\sigma}\mathbf{n} &= \bar{\mathbf{t}} \text{ on } \Gamma_\sigma, \\ \mathbf{u} &= \bar{\mathbf{u}} \text{ on } \Gamma_u. \end{aligned} \tag{3.2}$$

Remark 3.1. Again, we make *no restrictions* on the constitutive law governing $\boldsymbol{\sigma}$, except that it must be within the scope of the infinitesimal theory.

We next state the form of the Coulomb friction law presumed to hold here. We define

$$\begin{aligned} \mathbf{t}_T(\mathbf{u}) &:= -\boldsymbol{\sigma}\mathbf{n} - t_N\mathbf{n} \\ \text{and} \quad \mathbf{u}_T &:= \mathbf{u} - (\mathbf{u}\cdot\mathbf{n})\mathbf{n} \text{ on } \Gamma \end{aligned} \tag{3.3}$$

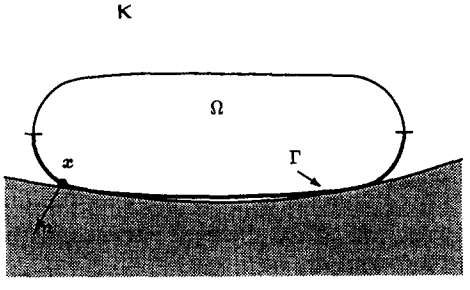

Fig. 3. Notation for the frictional obstacle problem in small deformations.

as the tangential components of the traction (note the sign change) and displacement, respectively. This provides the necessary notation to define the *Kuhn–Tucker conditions for Coulomb friction*

$$\Phi := \|\mathbf{t}_T\| - \mu t_N \leqslant 0, \tag{3.4.1}$$

$$\dot{\mathbf{u}}_T = \xi \frac{\partial}{\partial \mathbf{t}_T}\Phi, \tag{3.4.2}$$

$$\xi \geqslant 0, \tag{3.4.3}$$

$$\xi\Phi = 0. \tag{3.4.4}$$

We recognize (3.4.1) as the familiar Coulomb friction condition, with $\mu > 0$ being the *coefficient of friction*. Equations (3.4.2) and (3.4.3) mandate that slip occur in the direction opposite that of the applied tangential traction, and eqn (3.4.4) enforces the condition that slip may only occur when $\Phi = 0$, i.e., when $\|\mathbf{t}_T\| = \mu t_N$. If $\|\mathbf{t}_T\| < \mu t_N$, then *perfect stick* occurs, and $\dot{\mathbf{u}}_T = 0$.

Remark 3.2. In interpreting (3.4.1)–(3.4.4), the analogy of *rigid–perfectly plastic response* from the theory of plasticity may prove useful. Here, Φ plays the role of the yield criterion, and $\dot{\mathbf{u}}_T$ plays the role of the (deviatoric) strain rate. Note that *all* of the deformation that occurs is inelastic.

Remark 3.3. An important distinction between the current formulation and usual treatments of metal plasticity is also noted. We remark that flow rule (3.4.2) is *non-associated* since there is no irreversible slip allowed in the normal direction, as would be mandated by using Φ itself as the flow potential (see [15]). Thus, the extraction of a maximum dissipation principle appears impossible except in the case of t_N = constant. In this sense, the current formulation bears a close resemblance to a Drucker–Prager type of constitutive law.

To construct the weak form, we first note that the admissible variations are again *constrained*

$$\delta\mathbf{u}\cdot\mathbf{n} \leqslant 0 \text{ on } \Gamma \text{ if } g(\mathbf{u}) = 0. \tag{3.5}$$

Proceeding in the usual manner then yields the following

$$\begin{aligned} G(\mathbf{u}, \delta\mathbf{u}) &:= \int_\Omega \boldsymbol{\sigma} \cdot \operatorname{grad}[\delta\mathbf{u}]\, d\Omega \\ &\quad - \int_\Omega \mathbf{f}\cdot\delta\mathbf{u}\, d\Omega - \int_{\Gamma_\sigma} \bar{\mathbf{t}}\cdot\delta\mathbf{u}\, d\Gamma \\ &= \int_\Gamma [-t_N\mathbf{n}\cdot\delta\mathbf{u} - \mathbf{t}_T\cdot\delta\mathbf{u}_T]\, d\Gamma \end{aligned} \tag{3.6}$$

which must hold for all $\delta \mathbf{u}$ such that $\delta \mathbf{u} = 0$ on Γ_u and (3.5) is satisfied. We note that t_N and $\mathbf{t}_T$ in the right-hand side of (3.6) are determined through eqns (3.1) and (3.4.1)–(3.4.4).

As was the case in frictionless contact, (3.6) is difficult to work with due to restriction (3.5). This motivates the use of a penalty regularization.

3.2. *Penalty regularization of the frictional obstacle problem*

One could follow the procedure suggested in Sec. 2 and introduce *independent* variables λ_N and $\boldsymbol{\lambda}_T$ on the frictional interface Γ. This procedure, corresponding to the *Lagrange multiplier* treatment of the frictional obstacle problem, would result in

$$G(\mathbf{u}, \delta \mathbf{u}) = \int_\Gamma [-\lambda_N \mathbf{n} \cdot \delta \mathbf{u} - \boldsymbol{\lambda}_T \cdot \delta \mathbf{u}_T] \, d\Gamma \quad (3.7)$$

subject to, for all $\mathbf{x} \in \Gamma$

$$g(\mathbf{u}) \leqslant 0 \quad (3.8.1)$$

$$\lambda_N \geqslant 0 \quad (3.8.2)$$

$$\lambda_N g = 0 \quad (3.8.3)$$

$$\Phi = \|\boldsymbol{\lambda}_T\| - \mu \lambda_N \leqslant 0 \quad (3.8.4)$$

$$\dot{\mathbf{u}}_T = \xi \frac{\partial}{\partial \boldsymbol{\lambda}_T} \Phi \quad (3.8.5)$$

$$\xi \geqslant 0 \quad (3.8.6)$$

$$\xi \Phi = 0. \quad (3.8.7)$$

This approach, while viable, suffers from possible indefinite structure in the resulting matrix problem. We thus introduce a *penalty regularization* in which the Kuhn–Tucker unilateral constraint conditions (3.8.1)–(3.8.3), and the slip rule (3.8.5) are replaced as follows:

$$t_N = \epsilon_N \langle g(\mathbf{u}) \rangle \quad (3.9.1)$$

$$\Phi = \|\mathbf{t}_T\| - \mu t_N \leqslant 0 \quad (3.9.2)$$

$$\dot{\mathbf{u}}_T - \xi \frac{\partial}{\partial \mathbf{t}_T} \Phi = \frac{1}{\epsilon_T} \dot{\mathbf{t}}_T \quad (3.9.3)$$

$$\xi \geqslant 0 \quad (3.9.4)$$

$$\xi \Phi = 0. \quad (3.9.5)$$

We note that (3.9.1) is an *identical* treatment of the normal contact constraint as that given in (2.11), keeping in mind that the appearance of the Macauley bracket in (3.9.1) is due to the fact that g is now a signed quantity. Equation (3.9.3) is a penalization of the constraint suggested by (3.8.5), and is satisfied exactly only in the limit $\epsilon_T \to \infty$.

Remark 3.4. Some authors (see, for example, [18]) have felt it convenient to interpret ϵ_T as a *physical* stiffness corresponding to the stiffness of junctions at the interface. In this interpretation, $(1/\epsilon_T)\, \dot{\mathbf{t}}_T$ is the *elastic* part of the relative velocity at the interface. In our current treatment, however, we wish to allow no such elastic motion on the interface. Thus, in the context now being considered, ϵ_T is best thought of as strictly being a penalization, induced mathematically to ensure that the slip rate equals the relative velocity.

We now wish to consider the integration algorithm necessary to solve equality (3.6) (with variations no longer constrained by the contact conditions), where t_N and $\mathbf{t}_T$ are given by (3.9.1)–(3.9.5). In general, we solve this problem incrementally over the time interval

$$[0, T] = \bigcup_{n=1}^{N} [t_n, t_{n+1}].$$

In each time increment, we start with (3.6) being satisfied at time t_n, and wish to enforce satisfaction of (3.6) at t_{n+1} subject to laws of evolution (3.9.1)–(3.9.5). We shall utilize a backward Euler integration scheme to integrate equations (3.9.1)–(3.9.5) and employ a return mapping strategy; such approaches are now well known and have been used extensively in the theory of plasticity (see [7]). Extension of these ideas to the frictional contact problem is straightforward and has been discussed in detail in [17] and [18]. The following variational problem results, where $\{\mathbf{u}_n, t_{N_n}, \mathbf{t}_{T_n}\}$ are given at t_n

$$G(\mathbf{u}_{n+1}, \delta \mathbf{u}) = \int_\Gamma [-t_{N_{n+1}} \mathbf{n} \cdot \delta \mathbf{u} - \mathbf{t}_{T_{n+1}} \cdot \delta \mathbf{u}_T] \, d\Gamma, \quad (3.10)$$

where the right-hand side of (3.10) is evaluated by setting

$$t_{N_{n+1}} = \epsilon_N \langle g(\mathbf{u}_{n+1}) \rangle \quad (3.11)$$

and by defining the trial 'stick' state as

$$\mathbf{t}^{\text{trial}}_{T_{n+1}} = \mathbf{t}_{T_n} + \epsilon_T (\mathbf{u}_{T_{n+1}} - \mathbf{u}_{T_n}),$$

$$\Phi^{\text{trial}}_{n+1} = \|\mathbf{t}^{\text{trial}}_{T_{n+1}}\| - \mu t_{N_{n+1}}. \quad (3.12)$$

The return mapping is completed by setting

$$\mathbf{t}_{T_{n+1}} = \mathbf{t}^{\text{trial}}_{n+1} - \Delta\xi \frac{\mathbf{t}^{\text{trial}}_{T_{n+1}}}{\|\mathbf{t}^{\text{trial}}_{T_{n+1}}\|}, \quad (3.13)$$

where

$$\Delta\xi = \begin{cases} 0 & \text{if } \Phi^{\text{trial}}_{n+1} \leqslant 0, \\ \dfrac{\Phi^{\text{trial}}_{n+1}}{\epsilon_T} & \text{if } \Phi^{\text{trial}}_{n+1} > 0. \end{cases} \quad (3.14)$$

A simple calculation shows that the above algorithm requires that $\|\mathbf{t}_{T_{n+1}}\| = \mu t_{N_{n+1}}$ if $\Phi_{n+1}^{\text{trial}} > 0$, and $\mathbf{t}_{T_{n+1}} = \mathbf{0}$ if $t_{N_{n+1}} = 0$.

We note, however, that just as in the case of frictionless contact, the kinematic constraint for parts of Γ that stick (the perfect stick requirement) is only accurately satisfied as $\epsilon_T \to \infty$, which engenders ill-conditioning. Thus, as was done in Sec. 2, we shall consider application of the method of augmented Lagrangians to frictional contact.

3.3. *Augmented Lagrangian formulation*

We begin discussion of the augmented Lagrangian formulation for frictional problems by rewriting the equation to be solved for contact problems involving friction

$$G(\mathbf{u}_{n+1}, \delta\mathbf{u}) = \int_\Gamma [-t_{N_{n+1}}\mathbf{n}\cdot\delta\mathbf{u} - \mathbf{t}_{T_{n+1}}\cdot\delta\mathbf{u}_T]\,d\Gamma. \tag{3.15}$$

In the present context, we recognize that t_N and $\mathbf{t}_T$ must be redefined such that they include contributions due to both the penalization and the Lagrange multipliers. In the case of t_N this is easily performed by using the augmentation scheme already discussed in Sec. 2. In view of eqn (2.14) we thus write

$$t_N = \langle \lambda_N + \epsilon_N g \rangle. \tag{3.16}$$

In the case of $\mathbf{t}_T$, we turn our attention to eqn (3.9.3). We recall that this equation amounts to a penalization of the constraint requiring the computed slip rate to equal the tangential velocity. In the case that $\Phi < 0$, this equation penalizes the constraint that $\dot{\mathbf{u}}_T = \mathbf{0}$. In the present case, we alter (3.9.3) such that only the penalized part of $\dot{\mathbf{t}}_T$ appears on the right hand side. Assuming that $\mathbf{t}_T$ decomposes additively into its penalty and Lagrange multiplier parts, and denoting the Lagrange multiplier part of $\mathbf{t}_T$ by $\boldsymbol{\lambda}_T$, we write down the augmented Lagrangian statement of the friction law

$$t_N = \langle \lambda_N + \epsilon_N g \rangle, \tag{3.17.1}$$

$$\Phi = \|\mathbf{t}_T\| - \mu t_N \leqslant 0, \tag{3.17.2}$$

$$\dot{\mathbf{u}}_T - \xi \frac{\partial}{\partial \mathbf{t}_T}\Phi = \frac{1}{\epsilon_T}(\dot{\mathbf{t}}_T - \dot{\boldsymbol{\lambda}}_T), \tag{3.17.3}$$

$$\xi \geqslant 0, \tag{3.17.4}$$

$$\xi\Phi = 0. \tag{3.17.5}$$

Again we are interested in integrating eqns (3.17.1)–(3.17.5) between t_n and t_{n+1} in the algorithmic setting. Application of a backward Euler scheme to (3.17.3), and subsequent utilization of the return mapping scheme alluded to previously gives

$$t_{N_{n+1}} = \langle \lambda_{N_{n+1}} + \epsilon_N g(\mathbf{u}_{n+1}) \rangle \tag{3.18.1}$$

$$\mathbf{t}_{T_{n+1}} = \mathbf{t}_{T_n} + \Delta\boldsymbol{\lambda}_T + \epsilon_T\left(\Delta\mathbf{u}_T - \Delta\xi \frac{\mathbf{t}_{T_{n+1}}^{\text{trial}}}{\|\mathbf{t}_{T_{n+1}}^{\text{trial}}\|}\right), \tag{3.18.2}$$

where

$$\Delta\boldsymbol{\lambda}_T = \boldsymbol{\lambda}_{T_{n+1}} - \boldsymbol{\lambda}_{T_n},$$

$$\Delta\mathbf{u}_T = \mathbf{u}_{T_{n+1}} - \mathbf{u}_{T_n},$$

$$\mathbf{t}_{T_{n+1}}^{\text{trial}} = \mathbf{t}_{t_n} + \Delta\boldsymbol{\lambda}_T + \epsilon_T \Delta\mathbf{u}_T, \tag{3.19}$$

and $\Delta\xi$, the consistency parameter, is given by

$$\Delta\xi = \begin{cases} 0 & \text{if } \Phi_{n+1}^{\text{trial}} \leqslant 0 \\ \dfrac{\Phi_{n+1}^{\text{trial}}}{\epsilon_T} & \text{if } \Phi_{n+1}^{\text{trial}} > 0 \end{cases} \tag{3.20}$$

with $\Phi_{n+1}^{\text{trial}}$ simply being $\Phi(\mathbf{t}_{T_{n+1}}^{\text{trial}}, t_{N_{n+1}})$.

In examination of these expressions, we note that $\mathbf{t}_{T_n}$ is in practice the converged sum of the penalty and Lagrange multiplier contributions from the last time step, and is completely fixed with regard to determination of the state at t_{n+1}. We further assert that is the context of *nested* augmented Lagrangian schemes, $\Delta\boldsymbol{\lambda}_T$ is fixed with regard to the solution phase (as is λ_N). Thus, the use of equations (3.18.1)–(3.20) in the course of solving (3.15) constitutes determination of $\mathbf{t}_{T_{n+1}}$ and $t_{N_{n+1}}$ within a displacement-driven framework just as in the case of the normal penalty regularization, with the appropriate consideration of the multiplier contributions.

Remark 3.5. It is noteworthy that an increment of $\boldsymbol{\lambda}_T$, $\Delta\boldsymbol{\lambda}_T$, appears conveniently in (3.19) rather than the multiplier itself. This seems intuitively reasonable due to the fact that the constraint involved here is expressed in terms of time derivatives of quantities, rather than in terms of undifferentiated quantities (as is the case with the normal contact constraint). It is this increment in the multiplier, rather than $\boldsymbol{\lambda}_{T_{n+1}}$, which is actually stored and augmented in the implementation proposed here.

In order to complete the description of the algorithm the update formulas must be prescribed. The treatment of the normal contact is identical to that already discussed in the frictionless case, and the update formula for λ_N is given in (2.15). In order to determine the update for $\Delta\boldsymbol{\lambda}_T$ we first note that if the multipliers are the correct ones, $\boldsymbol{\lambda}_{T_{n+1}} = \mathbf{t}_{T_{n+1}}$ and $\boldsymbol{\lambda}_{T_n} = \mathbf{t}_{T_n}$. Then $\mathbf{t}_{T_{n+1}} = \mathbf{t}_{T_n} + \Delta\boldsymbol{\lambda}_T$, with $\Delta\boldsymbol{\lambda}_T$ now denoting the exact change in the multiplier from t_n to

t_{n+1}. Comparison with (3.18.2) suggests the proper update formula for the tangential multipliers

$$\Delta\boldsymbol{\lambda}_T^{(k+1)} = \Delta\boldsymbol{\lambda}_T^{(k)} + \epsilon_T\left(\Delta\boldsymbol{u}_T^{(k)} - \Delta\xi^{(k)}\frac{\mathbf{t}_{T_{n+1}}^{\text{trial}(k)}}{\|\mathbf{t}_{T_{n+1}}^{\text{trial}(k)}\|}\right) \quad (3.21)$$

where $\Delta\mathbf{u}_T^{(k)}$, $\Delta\xi^{(k)}$, and $\mathbf{t}_{T_{n+1}}^{\text{trial}(k)}$ are given in (3.19) and (3.20), computed using $\mathbf{u}_{n+1}^{(k)}$. It is important to note that this update is done (with a return mapping, if necessary) so that the Coulomb condition is satisfied by $\lambda_{N_{n+1}}^{(k+1)}$ and $\boldsymbol{\lambda}_{T_{n+1}}^{(k+1)}$.

As was the case in frictionless contact, the solution of (3.15) involves solving a highly nonlinear set of equations. As a result, the issue of whether *simultaneous* or *nested* iteration should be used becomes important. We shall herein consider only the technique of nested iteration, but remark in passing that schemes based upon simultaneous iteration are easily conceived in the present framework. Summarizing the information discussed above, we present the (primary) nested augmented Lagrangian for frictional contact contact in Table 2.

Remark 3.6. Note that as a consequence of the nested approach, the multipliers are completely fixed during the solution phase. Transmittal of information from the penalty terms to the multipliers occurs during the augmentation procedure, creating new 'best estimates' of the correct multipliers. In such a way, we see that in the first augmentation iteration during a time step, all information about the change in contact tractions during the step is contained in the penalty terms. As augmentations continue, however, this information is transferred into the multipliers via the augmentation procedure, so that when convergence is achieved the multipliers are the contact tractions, while the penalty terms are essentially zero.

Remark 3.7. Note that the conditions for constraint satisfaction given in step 3 of Table 2 amount to checks of the impenetrability condition and the perfect stick condition (satisfaction of the Coulomb condition is guaranteed by the construction of $\mathbf{t}_T$). In a computational environment, one could preset the indicated tolerances and check these constraints automatically in the manner indicated. In practice, it is convenient (particularly in the case of interactive computing) to simply continue the augmentation process until successive augmentations yield little or no discernible change in the solution.

We note once more that the advantage of the algorithm lies in the fact that ϵ_N and ϵ_T (which are considered to be fixed) are chosen as large as possible without inducing ill-conditioning. Improvement of constraint satisfaction is attained through the augmentation procedure, without significantly altering the conditioning of the matrix equations emanating from (3.15). Furthermore, if the solution in step 2 of Table 2 is carried out in a Newton–Raphson scheme, utilizing a consistent linearization of the equations, the rate of convergence within step 2 will be asymptotically quadratic, as in the case of the

Table 2. Primary nested augmented Lagrangian algorithm for frictional contact

1. Initialization:

   set $\lambda_N^{(0)} = \langle\lambda_N + \epsilon_N g\rangle$ from last time step,
   $\Delta\boldsymbol{\lambda}_T^{(0)} = \mathbf{0}$,
   $k = 0$.

2. Solve (using a nonlinear solution strategy) for $\mathbf{u}_{n+1}^{(k)}$:

$$G(\mathbf{u}_{n+1}^{(k)}, \delta\mathbf{u}) + \int_\Gamma [\langle\lambda_N^{(k)} + \epsilon_N g(\mathbf{u}_{n+1}^{(k)})\rangle\delta\mathbf{u}\cdot\mathbf{n} + \mathbf{t}_T(\mathbf{u}_{n+1}^{(k)})\cdot\delta\mathbf{u}_T]\,d\Gamma = 0$$

   where $\mathbf{t}_T(\mathbf{u}_{n+1}^{(k)})$ is given by (3.18.1)–(3.20) using $\Delta\boldsymbol{\lambda}_T^{(k)}$ for $\Delta\boldsymbol{\lambda}_T$.

3. Check for constraint satisfaction:

   IF $(g(\mathbf{u}_{n+1}^{(k)}) \leqslant \text{TOL1}$ for all $\mathbf{x}\in\Gamma$ AND $\|\mathbf{u}_{T_{n+1}} - \mathbf{u}_{T_n}\| \leqslant \text{TOL2}$ for all $\mathbf{x}\in\Gamma$ such that $\|\mathbf{t}_T\| < \mu(\langle\epsilon_N g + \lambda_N\rangle)$ THEN
   Converge. EXIT.
   ELSE
   AUGMENT (for all $\mathbf{x}\in\Gamma$):
   $\lambda_N^{(k+1)} = \langle\lambda_N^{(k)} + \epsilon_N g(\mathbf{u}_{n+1}^{(k)})\rangle$

$$\Delta\boldsymbol{\lambda}_T^{(k+1)} = \begin{cases} \Delta\boldsymbol{\lambda}_T^{(k)} + \epsilon_T\Delta\mathbf{u}_T^{(k)} \text{ if } \|\mathbf{t}_{T_n} + \Delta\boldsymbol{\lambda}_T^{(k)} + \epsilon_T\Delta\mathbf{u}_T^{(k)}\| \leqslant \mu\lambda_N^{(k+1)} \\ \dfrac{\mathbf{t}_{T_n} + \Delta\boldsymbol{\lambda}_T^{(k)} + \epsilon_T\Delta\mathbf{u}_T^{(k)}}{\|\mathbf{t}_{T_n} + \Delta\boldsymbol{\lambda}_T^{(k)} + \epsilon_T\Delta\mathbf{u}_T^{(k)}\|}\mu\lambda_N^{(k+1)} - \mathbf{t}_{T_n} \text{ otherwise} \end{cases}$$

   $k \leftarrow k+1$
   GOTO 2.
   ENDIF

penalty regularization. In fact, one further notes that if no augmentations were performed in any time step, the method reduces to the penalty regularization.

Despite these advantages, we assert that solution of (3.15), due to return mapping (3.18.1)–(3.20), inherits a somewhat unsavory characteristic of the penalty regularization; namely, nonsymmetric tangent stiffnesses are obtained in the consistent linearization of (3.15) in the case where part of Γ slips (see, e.g., [18]). This nonsymmetry arises from linearization of the return map, so one alternative might be to eliminate the return map in the solution of (3.15). This motivates the consideration of an alternative scheme for augmented Lagrangian treatment of contact.

The variation of the algorithm which we shall consider resembles, in some ways, algorithms discussed in [13], and presented in more detail in [20] and [22]. In these treatments, a return mapping is not performed using the current normal tractions, thus eliminating the troublesome nonsymmetric nature of the return map emanating from the nonassociated flow rule discussed earlier. Instead of this return mapping, the Coulomb condition is enforced with respect to a previously-obtained estimate of the normal tractions, which is held fixed in the solution phase. Based upon the solution to this problem, these algorithms compute new estimates of the normal tractions, which are then used as data in order to define the friction law for the next solution phase. The key idea of these algorithms is that strict enforcement of the friction law is not made in the solution phase alone; a second step is utilized to provide a correction which better enforces the Coulomb criterion. In effect, these algorithms render the same results as the algorithm discussed in the previous subsection, since they are based on a penalty regularization. They render symmetric equations, but at the cost of a two-step algorithm which is iterative.

Table 3. Alternative (symmetrized) nested augmented Lagrangian algorithm for frictional contact

1. Initialization:

$$\text{set } \lambda_N^{(0)} = \langle \lambda_N + \epsilon_N g \rangle \text{ from last time step,}$$
$$\Delta\boldsymbol{\lambda}_T^{(0)} = \mathbf{0},$$
$$\Gamma_{\text{stick}}^{(0)} = \Gamma,$$
$$k = 0$$

2. Solve (using a nonlinear solution strategy) for $\mathbf{u}_{n+1}^{(k)}$:

$$G(\mathbf{u}_{n+1}^{(k)}, \delta\mathbf{u}) + \int_\Gamma [\langle \lambda_N^{(k)} + \epsilon_N g(\mathbf{u}_{n+1}^{(k)}) \rangle \delta\mathbf{u} \cdot \mathbf{n} + \mathbf{t}_T(\mathbf{u}_{n+1}^{(k)}) \cdot \delta\mathbf{u}_T] \, d\Gamma = 0$$

where

$$\mathbf{t}_T(\mathbf{u}_{n+1}^{(k)}) = \begin{cases} \mathbf{t}_{T_n} + \Delta\boldsymbol{\lambda}_T^{(k)} + \epsilon_T \Delta\mathbf{u}_T^{(k)} \text{ if } \mathbf{x} \in \Gamma_{\text{stick}}^{(k)} \\ \Delta\boldsymbol{\lambda}_T^{(k)} + \mathbf{t}_{T_n} \text{ otherwise} \end{cases}$$

3. Check for constraint satisfaction:

IF $(g(\mathbf{u}_{n+1}^{(k)}) \leqslant \text{TOL1}$ for all $\mathbf{x} \in \Gamma$ AND $\|\mathbf{u}_{T_{n+1}} - \mathbf{u}_{T_n}\| \leqslant \text{TOL2}$

for all $\mathbf{x} \in \Gamma$ such that $\|\mathbf{t}_T\| < \mu \langle \epsilon_N g + \lambda_N^{(k)} \rangle$ AND
$\|\mathbf{t}_T\| \leqslant (1 + \text{TOL3}) \mu \langle \epsilon_N g + \lambda_N^{(k)} \rangle$ for all $\mathbf{x} \in \Gamma$) THEN
Converge. EXIT.

ELSE
Augment and update Γ_{stick} (for all $\mathbf{x} \in \Gamma$):

$$\lambda_N^{(k+1)} = \langle \lambda_N^{(k)} + \epsilon_N g(\mathbf{u}_{n+1}^{(k)}) \rangle$$

IF $(\|\mathbf{t}_{T_n} + \Delta\boldsymbol{\lambda}_T^{(k)} + \epsilon_T \Delta\mathbf{u}_T^{(k)}\| \leqslant \mu\lambda_N^{(k+1)})$ THEN

$$\Delta\boldsymbol{\lambda}_T^{(k+1)} = \Delta\boldsymbol{\lambda}_T^{(k)} + \epsilon_T \Delta\mathbf{u}_T^{(k)},$$

$$\mathbf{x} \in \Gamma_{\text{stick}}^{(k+1)}$$

ELSE

$$\Delta\boldsymbol{\lambda}_T^{(k+1)} = \frac{\mathbf{t}_{T_n} + \Delta\boldsymbol{\lambda}_T^{(k)} + \epsilon_T \Delta\mathbf{u}_T^{(k)}}{\|\mathbf{t}_{T_n} + \Delta\boldsymbol{\lambda}_T^{(k)} + \epsilon_T \Delta\mathbf{u}_T^{(k)}\|} \mu\lambda_N^{(k+1)} - \mathbf{t}_{T_n}$$

$$\mathbf{x} \notin \Gamma_{\text{stick}}^{(k+1)}$$

ENDIF
$k \leftarrow k + 1$
GOTO 2.
ENDIF

In the algorithm we now consider, we wish to exploit the idea that the presence of augmented Lagrangians mandates an iterative procedure already, without regard to satisfaction of the Coulomb criterion. Thus, we consider an algorithm which will cause the solution phase to involve only symmetric equations, by eliminating the return map altogether from this phase. We then use the augmentation phase not only to augment the multipliers, but to enforce the Coulomb criterion as well. We do this by defining Γ_{stick}, a portion of Γ which is treated in the solution phase as being the part of the interface where perfect stick occurs, without regard to the current tangential and normal tractions. The remainder of Γ is treated as a slip region, and has *fixed* tangential tractions which satisfy the Coulomb condition defined by the current Lagrange multiplier λ_N. In the first iteration in a given time step, Γ_{stick} is taken to be all of Γ (i.e., the entire interface is assumed to stick), and further amendments to Γ_{stick} are made in the augmentation phase, according to the current state of the multipliers. Thus, the Coulomb condition is only enforced in the augmentation phase, which causes the equations of the solution phase to be symmetric. The complete algorithm is given in Table 3.

We reemphasize that the motivation for this alternative algorithm is to symmetrize the equations. It does not emanate from the penalty regularization as naturally as does the primary algorithm, but exploits the iterative nature of the augmented Lagrangian process in the manner in which it enforces the Coulomb friction conditions. More will be said about the performance of these two approaches in the next section.

4. NUMERICAL EXAMPLES

In this section we present some numerical examples intended to demonstrate the utility and performance of the proposed augmented Lagrangian treatment of frictional contact. In these simulations, the finite element code FEAP was utilized, which has most recently been described in [23]. The discretization and subsequent implementation of the contact conditions employed in these examples were performed within the framework discussed by [18], and is discussed in more detail in Appendices A and B. The important point to note is that the ideas discussed in the previous sections are directly applicable to problems involving finite-deformation response and contact of multiple deformable bodies, as is demonstrated through the following numerical examples and the Appendices.

4.1. *Sliding of an elastic block*

As a first demonstration of the performance of the techniques discussed in this paper we consider the problem of an elastic block sliding against a rough rigid foundation. This problem, which has been considered previously by [22] and [18] is rather simple and idealized, but serves to highlight the manner in which the proposed algorithms work in practice.

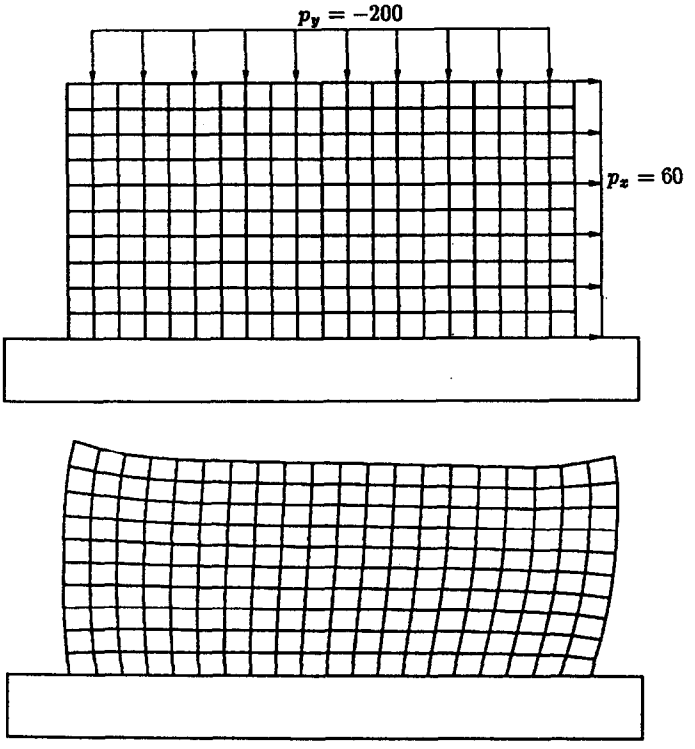

Fig. 4. Undeformed and deformed geometries for the elastic block problem.

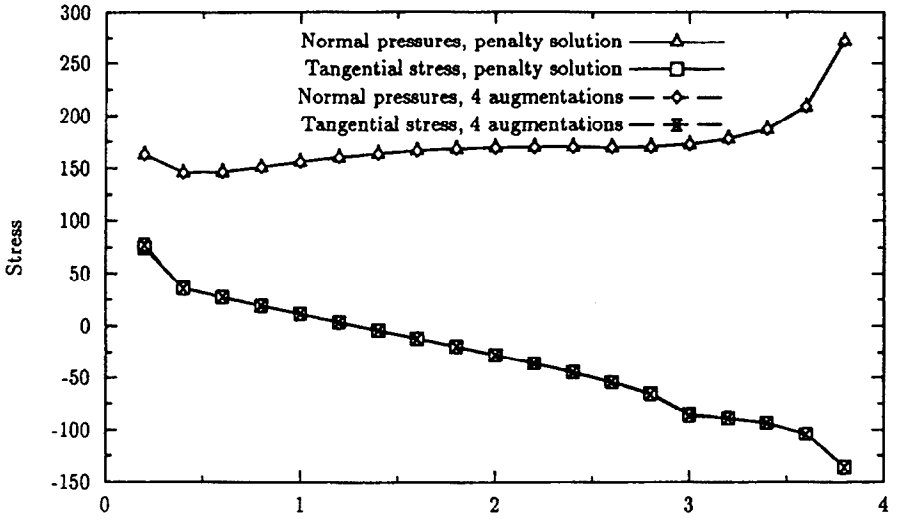

Fig. 5. Computed contact tractions for the elastic block problem, for the standard regularization (no augmentations) and the primary augmented Lagrangian algorithm (four augmentations).

In this problem, an elastic block is simultaneously pushed into the foundation and pulled along it, resulting in frictional sliding response at the interface. The finite element discretization and computed response are to be found in Fig. 4. The block, which has an elastic modulus of 1000 and Poisson's ratio of 0.3, has been discretized using 200 four-node linear isotropic elastic elements. A Coulomb friction law is presumed to hold between the block and foundation, with $\mu = 0.5$. As reported in [18] the solution to this problem may be obtained in only one load increment when using the standard penalization of the Coulomb friction law. To conform to the solution of these authors, no frictional stress is allowed to develop at the first and last nodes of the contact surface.

In examining this problem the first calculation done was one in which the standard penalty method was used, with $\epsilon_N = 10^8$ and $\epsilon_T = 10^4$. These values exactly correspond to those used by [18] in their simulation. The computed contact tractions on the frictional interface are shown in Fig. 5, where the results of a calculation using these same penalty values but employing the (primary) augmented Lagrangian algorithm are also shown. It is to be noted that these tractions are not the nodal projections of the element stresses, but are merely the nodal reactions normalized by the element lengths. This characterization of the contact tractions was performed to conform to the technique used by the above authors.

It is to be noted that these results are in accordance with those reported by [22] and [18]. It is also apparent from Fig. 5 that the tractions remain essentially unchanged when the augmentations are

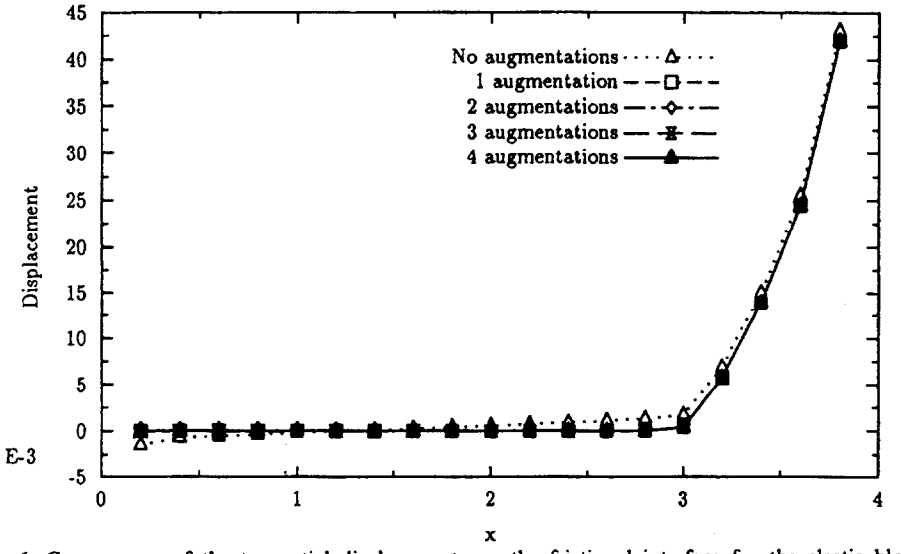

Fig. 6. Convergence of the tangential displacements on the frictional interface for the elastic block problem, using the primary augmented Lagrangian algorithm.

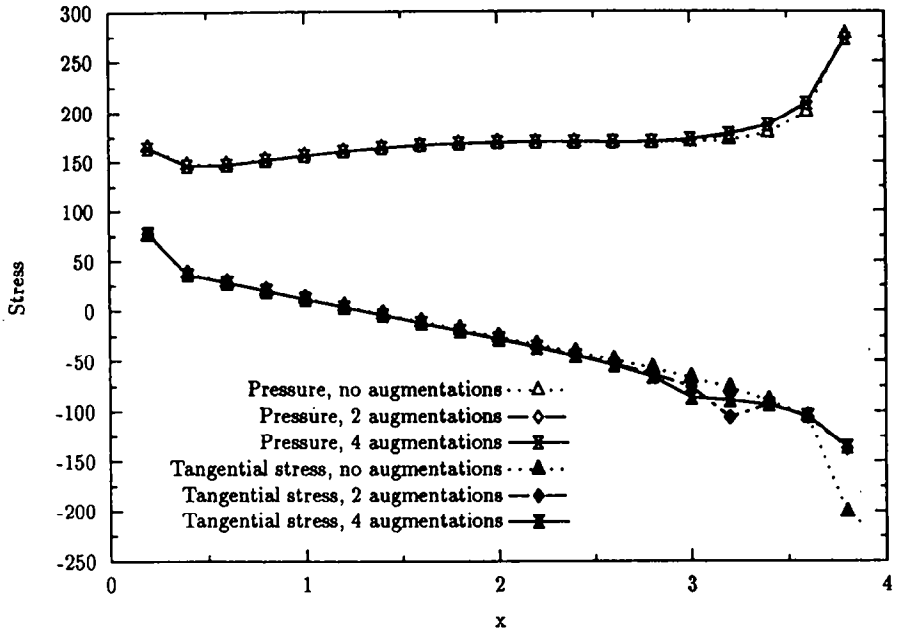

Fig. 7. Convergence of the contact tractions for the elastic block problem, using the alternative augmented Lagrangian algorithm.

performed, suggesting that the penalties used are adequately enforcing the constraints. Although for practical purposes this is certainly the case, it is instructive to also examine the convergence of the tangential displacements on the interface as the augmentations are performed. This information is reported in Fig. 6.

As one can see from the figure, the augmentations actually produce a noticeable change in the tangential displacements. The curve corresponding to the penalty solution (no augmentations) is seen to include nonzero nodal displacements where stick is to occur, amounting to a slight violation of the tangential constraint. As can also be seen from the figure, however, only one augmentation is required to correct this situation, with subsequent augmentations producing no discernible change in the displacements on the interface. Although this is a fairly minor point for this problem, this example shows that successive applications of the augmentation procedure do indeed improve satisfaction of the constraints. In the final converged solution, the stick and slip regions of the interface are easily discerned from Fig. 6, with the last five nodes slipping and the remainder sticking.

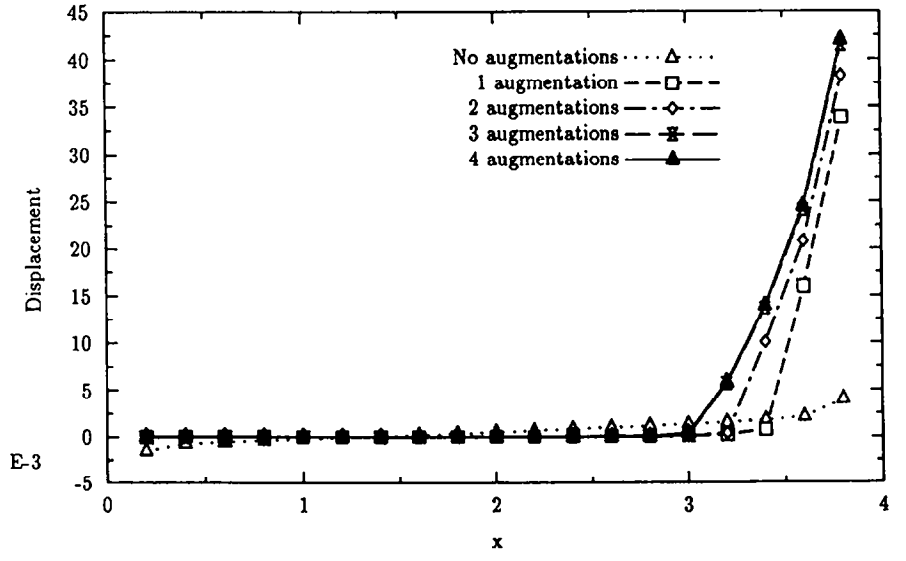

Fig. 8. Convergence of tangential displacements on the frictional interface for the elastic block problem, using the alternative augmented Lagrangian algorithm.

This example is also convenient for examination of the alternative augmented Lagrangian algorithm discussed in Sec. 3.3. Figures 7 and 8 show the convergence of the contact stresses and tangential displacements, respectively, as the result of augmentations performed according to the alternative algorithm (using the same penalties as above). Note that although the early solutions are clearly inadequate due to the initial assumption of stick on the entire interface, the algorithm converges quite rapidly and after four augmentations produces the correct result. The advantage of this algorithm in the current situation is that due to the fact that the block is linear elastic and the contact area does not change from iteration to iteration, the elimination of the return map from the solution phase makes the problem entirely linear, requiring only one iteration for solution for each augmentation. Solution of either the standard regularized system or the equations of the primary augmented Lagrangian algorithm, on the other hand, are nonlinear due to the return mapping present in the frictional response. Thus, for problems of this type, the alternative algorithm has a strong appeal, and has the advantage that upon convergence, the contact constraints will be essentially exactly satisfied.

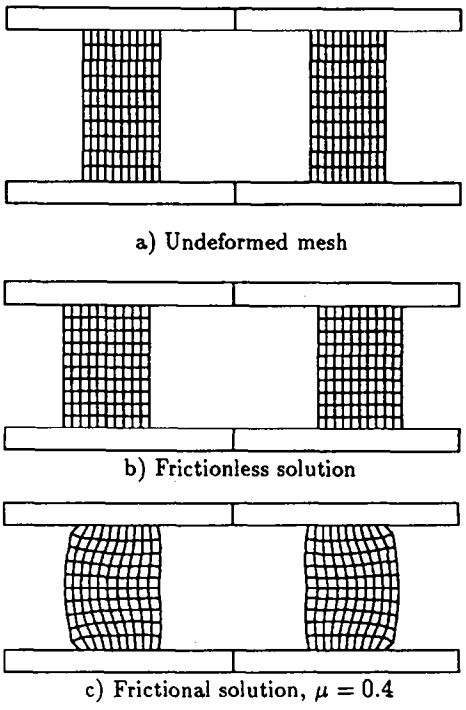

Fig. 9. Initial and deformed geometries for the ring upsetting problem, showing both frictionless and frictional solutions.

4.2. *Upsetting of a ring*

We next consider the problem of the upsetting of a ring. In doing so, we wish to further explore the ability of augmented Lagrangians to correct underpenalized solutions, as well as to explore the performance of the method in a finite deformation context. The problem is displayed pictorially in Fig. 9. In this axisymmetric problem, a ring of inner radius 10, outer radius 20, and initial height 20 is compressed by a rigid plate to 80% of its initial height. Due to the symmetry of the problem, only one quarter of the mesh shown in Fig. 9 is actually modelled. The ring is discretized using four-node finite strain elastoplastic elements of the type discussed in [24], and has the following material properties: $K = 166670$, $G = 76920$, $\sigma_Y = 300$, and H (linear hardening rate) $= 700$.

As indicated in the figure, the problem has been run with both frictionless contact and Coulomb friction assumed between the sample and the rigid plate. It is seen that the response in the frictionless

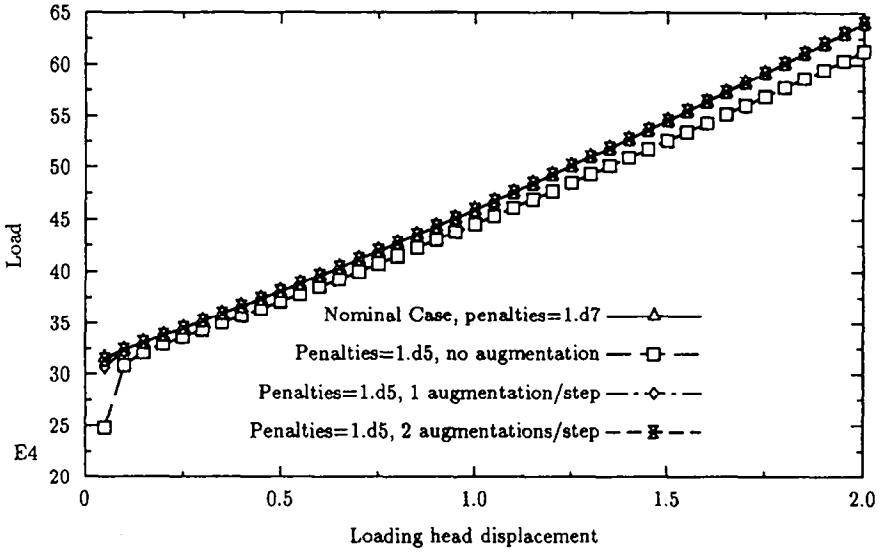

Fig. 10. Convergence of the load-displacement curve for the frictional ring upsetting problem, using the primary augmented Lagrangian algorithm.

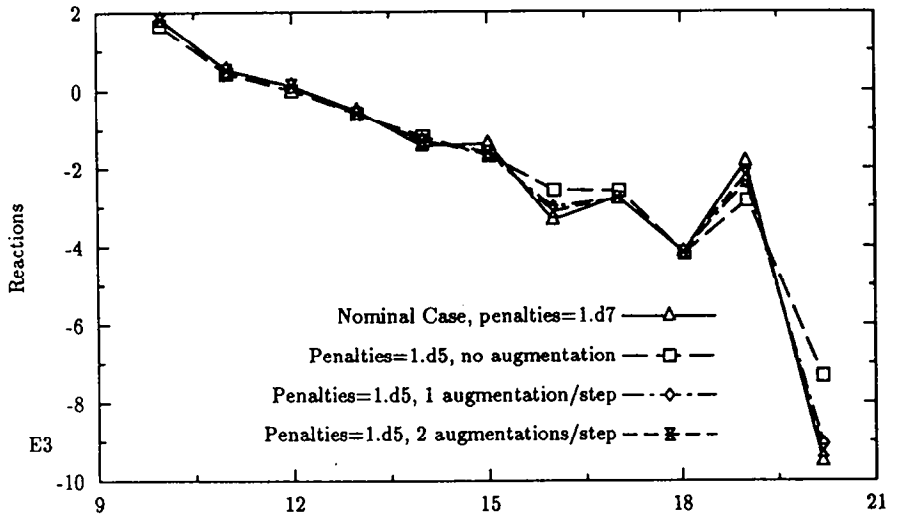

Fig. 11. Convergence of the tangential nodal reactions on the contact surface at the final state for the frictional ring upsetting problem, using the primary augmented Lagrangian algorithm.

case is completely homogeneous; no barreling of the sample occurs and the net movement of the material comprising the ring is outward. In the case with friction, however, barreling of the sample is apparent, and material flows both outward and inward during the process. It is seen then that the introduction of friction in this problem has an important effect upon response, and thus proper enforcement of constraints on the interface is important for a correct solution. It is thus the frictional case (with $\mu = 0.4$) upon which we shall concentrate in assessing the augmented Lagrangian algorithm.

The calculation shown in Fig. 9(c) was performed using $\epsilon_N = \epsilon_T = 10^7$, using 40 equal increments in the prescribed motion of the rigid plate. This combination leads to a proper enforcement of the contact and frictional constraints. With this information in hand, subsequent calculations were done in which the contact conditions were intentionally underpenalized (using $\epsilon_N = \epsilon_T = 10^5$), followed by application of the augmentation procedure in order to improve the solution. Rather than checking the criteria in each step, as suggested in Table 2, these calculations were done by simply performing a fixed number of augmentations in each time step and then proceeding to the next step (as before, the loading was applied in 40 equal increments). The results of these calculations (using the primary augmented Lagrangian algorithm) are given in Figs 10–12. In these figures, the convergence of the solution with respect to the number of augmentations per step is displayed through examination of the load displacement curve for the entire process (Fig. 10) and the contact reactions on the interface at the final state (Figs 11 and 12).

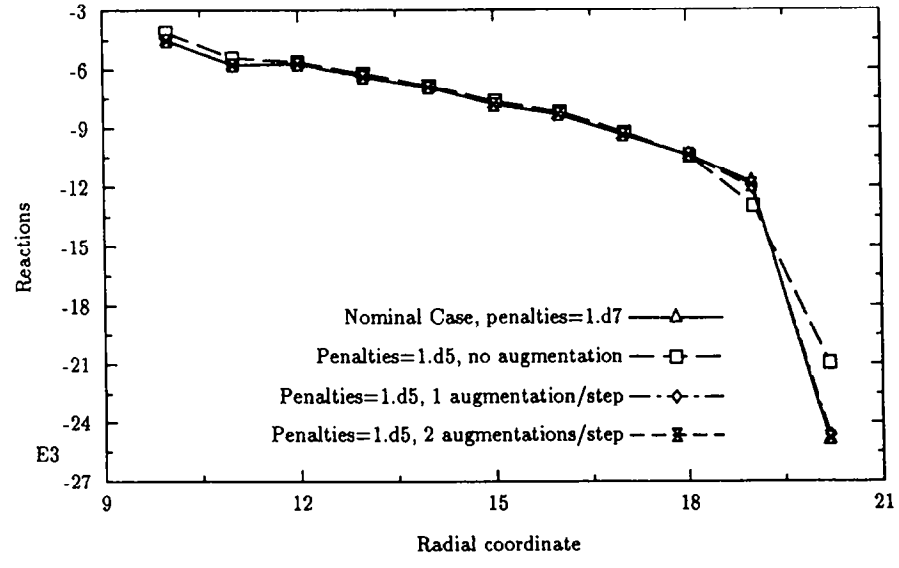

Fig. 12. Convergence of the normal nodal reactions on the contact surface at the final state for the frictional ring upsetting problem, using the primary augmented Lagrangian algorithm.

As one can see from the figures, the application of the augmentation procedure to an underpenalized system corrects a clearly inadequate solution using a modest number of augmentations. The benefit of the augmented Lagrangian scheme in a situation of this type is the added robustness it provides the standard penalty method; it enables a satisfactory solution to be obtained over a much larger range of penalty parameters than does the penalty method alone.

4.3. *Extrusion of an aluminum cylinder*

As a final example demonstrating the augmented Lagrangian technique for frictional problems, we consider the frictional extrusion of an aluminum cylinder into a rigid, conical die. This problem, also considered in [13], is one in which the augmentation technique actually enables easier solution of the equations than does the standard penalty method. In this axisymmetric problem, an aluminum billet of radius 5.08 cm and initial length 25.4 cm is pushed (using displacement control) a total distance of 17.78 cm into a conical die with wall angle 5°. The mesh, as well as the deformed geometry at various stages of the simulation, is shown in Fig. 13. The discretization of the billet was performed using 80 4-node finite strain elastoplastic elements of the type mentioned previously, with the material properties of the billet being: $K = 63.84$ GPa, $G = 26.12$ GPa, $\sigma_Y = 31$ MPa and H (the linear hardening rate) $= G/100$. The coefficient of friction between the billet and the die walls was prescribed as 0.1.

The interesting feature of this problem is that due to the highly plastic response of the billet and the shearing near the die walls due to friction, the introduction of high penalties on the contact surface causes difficulties in the solution of the equations. On the other hand, the use of undersized penalties causes erroneous predictions of stick regions on the contact surface, where a properly penalized solution displays no such behavior. The meaning of this is that for the case of the standard penalty regularization, the problem is extremely sensitive to the choice of the

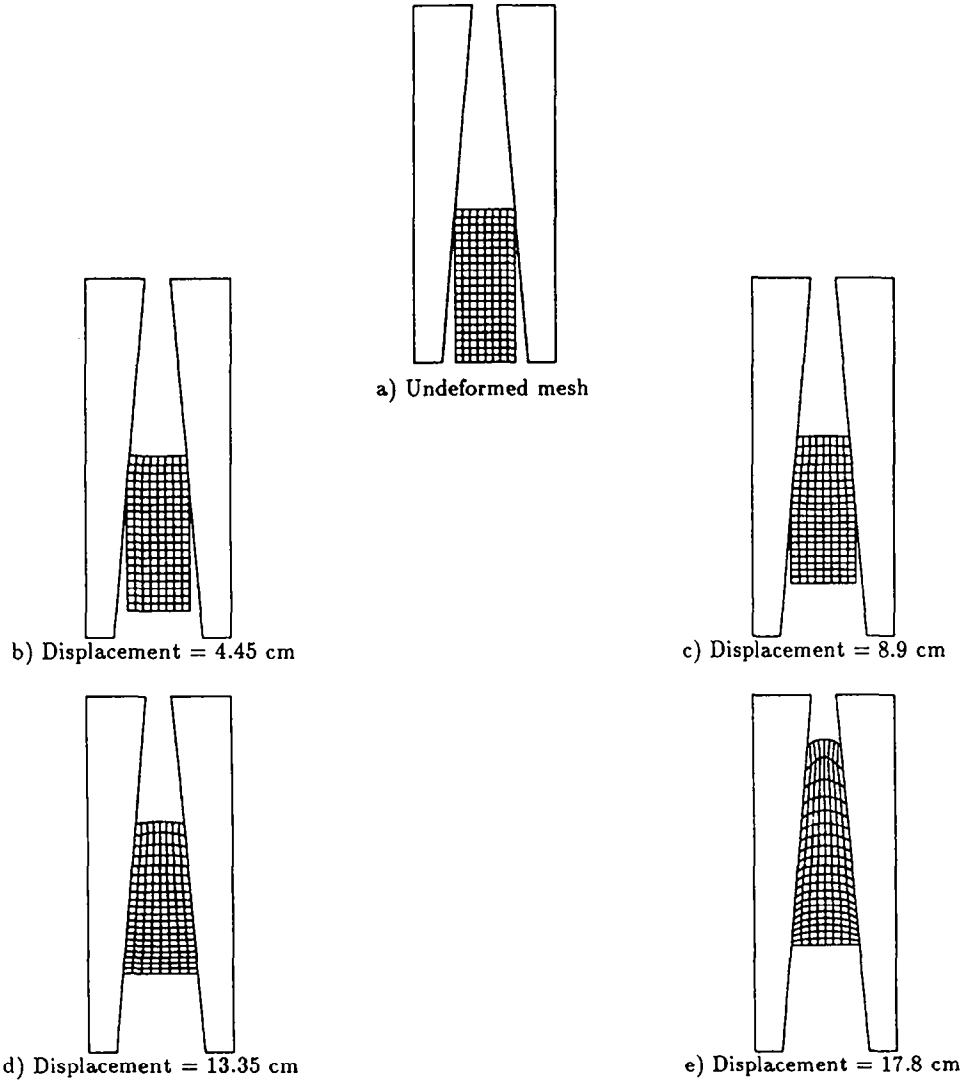

Fig. 13. Undeformed and deformed geometries for the aluminum extrusion problem.

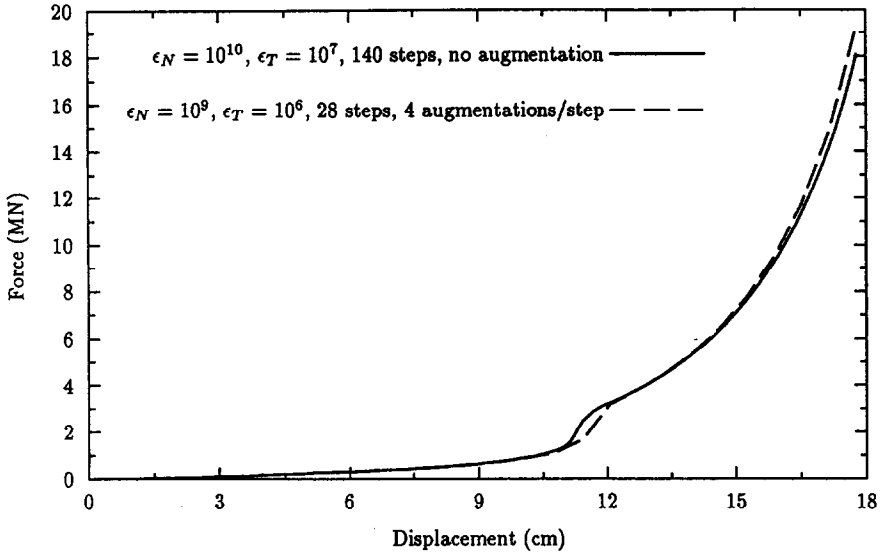

Fig. 14. Load–displacement curves for the conical extrusion problem, using the standard regularization and the primary augmented Lagrangian algorithm.

penalties. After much trial and error involving experimentation with these penalties, the optimum choice of the penalties seemed to be $\epsilon_N V = 10^{10}$ and $\epsilon_T = 10^7$, which yielded the solution shown in Fig. 13 in 140 equal time steps. This choice of load increment appears to be near optimal as well; for this choice of penalties larger time steps failed to converge at some point in the calculation.

In application of the augmented Lagrangian technique (the primary algorithm) to this problem, we exploit the opportunity to lower the penalties, increase the time step, and correct the solution via the augmentation procedure. The results of doing this are displayed graphically in the plot of force vs displacement given in Fig. 14. In this figure, the response computed using the nominal penalties given above is compared with the response obtained using $\epsilon_N = 10^9$ and $\epsilon_T = 10^6$ with the primary augmented Lagrangian algorithm. As can be seen, essentially the same solution is obtained using these softer penalties, with 4 augmentations per time step, *in only one-fifth as many time steps*. The advantage of the augmentation in this problem is that it circumvents a difficulty in solving the equations that is unrelated to the mechanics of the problem but arises only due to the penalty regularization of the friction law. This example displays again the enhanced robustness afforded the penalty method by the method of augmented Lagrangians.

5. SUMMARY AND CONCLUSIONS

In this paper, we have presented a framework within which the application of the method of augmented Lagrangians to frictional contact problems is a natural extension of the more classical case of frictionless contact. This has been done by considering the constraint in the tangential direction to be an equality constraint, between the total relative tangential velocity and the slip rate computed from the laws of evolution for the Coulomb friction law. Stated another way, we require that no elastic tangential displacement take place on the interface. We have also remarked on some alternative implementations of the method, involving both different enforcements of the Coulomb friction law (see Tables 2 and 3) and different augmentation procedures (simultaneous vs nested). Altough the latter scheme is the one focused upon in this paper, simultaneous iteration may also prove to be fruitful for some problems.

Numerical examples have demonstrated that the method is applicable in both small and large deformations. They have shown that the method is useful in correcting underpenalized solutions, and that it can in some cases obtain solutions more efficiently than the penalty method. As a practical matter, it is also pointed out that the technique can be used during interactive solution of a contact problem to check whether a penalized solution is adequate; if it is, application of the augmented procedure will produce very little change to the residual and subsequent solution vector, which is readily observed by the user. It is felt that the chief benefit of the augmented Lagrangian procedure for contact problems is the added robustness it provides the penalty method, while at the same time being a simple procedure which introduces no additional equations to the discrete system.

Acknowledgements—We thank Dr T. O'Sullivan and Dr M. Doerner of IBM for their continued interest and support of this research. Support for this work was provided by IBM under contract no. IBM-2DJA-0348 with Stanford University; this support is gratefully acknowledged.

REFERENCES

1. M. R. Hestenes, Multiplier and gradient methods. *J. Optimiz. Theory Applic.* **4**, 303–320 (1969).
2. M. J. D. Powell, A method for nonlinear constraints in minimization problems. In *Optimization* (Edited by R. Fletcher). Academic Press, New York (1969).
3. R. T. Rockefeller, Augmented Lagrange multiplier functions and duality in non-convex programming. *SIAM J. Control* **12**, 268–285 (1974).
4. D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press, New York (1982).
5. R. Fletcher, *Practical Methods of Optimization*, 2nd Ed. John Wiley, New Delhi (1989).
6. R. Glowinski and P. Le Tallec, Finite elements in nonlinear incompressible elasticity. In *Finite Elements, Vol. V: Special Problems in Solid Mechanics*, (Edited by J. T. Oden and G. F. Carey). Prentice-Hall, Englewood Cliffs, NJ (1984).
7. J. C. Simo and T. J. R. Hughes, *Elastoplasticity and Viscoplasticity: Computational Aspects*. Springer, Berlin (1991).
8. J. C. Simo and R. L. Taylor, Quasi-incompressible finite elasticity in principal stretches. Continuum basis and numerical algorithms. *Comput. Meth. appl. Mech. Engng* **85**, 273–310 (1991).
9. J. A. Landers and R. L. Taylor, An augmented Lagrangian formulation for the finite element solution of contact problems. NCEL Contract Report, Naval Civil Engineering Laboratory, Port Hueneme, CA (1986).
10. P. Wriggers, J. C. Simo and R. L. Taylor, Penalty and augmented Lagrangian formulations for contact problems. In *Proceedings of the NUMETA '85 Conference*. Elsevier (1985).
11. R. Glowinski and P. Le Tallec, *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*. SIAM Studies in Applied Mathematics, Philadelphia (1989).
12. G. Duvaut and J. L. Lions, *Les Inequations en Mecanique et en Physique*. Dunod, Paris (1972).
13. N. Kikuchi and J. T. Oden, *Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Methods*. SIAM, Philadelphia (1988).
14. T. J. R. Hughes, R. L. Taylor, J. L. Sackman, A. Curnier, and W. Kanoknukulchai, A finite element method for a class of contact-impact problems. *Comput. Meth. appl. Mech. Engng* **8**, 249–276 (1976).
15. R. Michalowski and Z. Mroz, Associated and non-associated sliding rules in contact friction problems. *Arch. Mech.* **30**, 259–276 (1978).
16. J. C. Simo and R. L. Taylor, Consistent tangent operators for rate-independent elastoplasticity. *Comput. Meth. appl. Mech. Engng* **48**, 101–118 (1985).
17. A. E. Giannakopoulos, The return mapping method for the integration of friction constitutive relations. *Comput. Struct.* **32**, 157–167 (1989).
18. P. Wriggers, T. Vu Van and E. Stein, Finite element formulation of large deformation impact-contact problems with friction. *Comput. Struct.* **37**, 319–331 (1990).
19. D. G. Luenberger, *Linear and Nonlinear Programming*, 2nd edn. Addison-Wesley, Reading, MA (1984).
20. L. T. Campos, J. T. Oden and N. Kikuchi, A numerical analysis of a class of contact problems with friction in elastostatics. *Comput. Meth. appl. Mech. Engng* **34**, 821–845.
21. M. Fortin and A. Fortin, A generalization of Uzawa's algorithm for the solution of the Navier-Stokes equations. *Comms appl. Numer. Meth.* **1**, 205–208 (1985).
22. J. T. Oden and E. B. Pires, Algorithms and numerical results for finite element approximations of contact problems with non-classical friction laws. *Comput. Struct.* **19**, 137–147 (1984).
23. O. C. Zienkiewicz and R. L. Taylor, *The Finite Element Method. Volume 1: Basic Formulation and Linear Problems*. McGraw-Hill, London (1989).
24. J. C. Simo, A framework for finite strain elastoplasticity based on maximum plastic dissipation and the multiplicative decomposition. Part II: Computational aspects. *Comput. Meth. appl. Mech. Engng* **68**, 1–31 (1988).
25. D. J. Benson and J. O. Hallquist, A single surface contact algorithm for the post-buckling analysis of shell structures. *Comput. Meth. appl. Mech. Engng* **78**, 141–163 (1990).

APPENDICES

A. Closest point projection to a piecewise linear obstacle

In the following we expand upon the details necessary to implement the definition of gap function discussed in Sec. 2.2 in a finite element setting. In so doing, we limit ourselves to the two-dimensional case, while remarking that generalizations to the 3-D case are similarly conceived.

Suppose that $\partial\mathbb{K}$ is defined by a series of line segments, as shown in Fig. A.1. We then wish, for any $\mathbf{x}$ in the spatial domain, to define the gap function $g(\mathbf{x})$ in a manner consistent with Sec. 2.2. With the usual case of interest in mind, we concentrate on the case where $\mathbf{x}$ is inadmissible. We further note that in practice the admissible region need not, and in general will not, be convex.

As suggested by the figure, we can characterize $\partial\mathbb{K}$ as being defined by $N+1$ nodal points $\mathbf{y}_k$, and we may define the outward normal to a segment between $\mathbf{y}_k$ and $\mathbf{y}_{k+1}$ as $\boldsymbol{\nu}_k$. As also suggested by the figure, if $\mathbf{x}$ is in a position such as the one shown, the expression for $g(\mathbf{x})$ is easy to find and is given by

$$g(\mathbf{x}) = -\boldsymbol{\nu}_k \cdot (\mathbf{x} - \mathbf{y}_k). \tag{A.1}$$

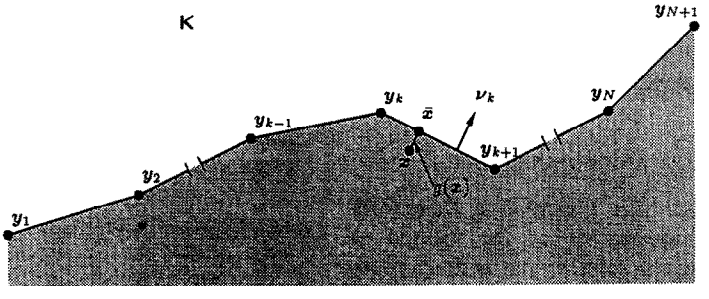

Fig. A.1. Closest point projection to a piecewise linear obstacle.

The trouble in practice, of course, is that we do not know beforehand which k is the appropriate one to use, and we recognize immediately that there are some points $\mathbf{x}$ for which the choice of $\bar{\mathbf{x}}$, the closest point projection of $\mathbf{x}$ on $\partial\mathbb{K}$, is nonunique. In FEAP, we address these issues by first finding the nodal point $\mathbf{y}_k$ closest to $\mathbf{x}$, which we shall denote by $\mathbf{y}_{\bar{k}}$. This is done most reliably by a sort through all nodes $\mathbf{y}_k$, but since this is also the most expensive technique to use, we only perform this type of search once, at the beginning of the problem, for all material points of interest. Subsequent searches for $\mathbf{y}_{\bar{k}}$ for a material point currently at $\mathbf{x}$ are performed using a method of steepest descent, where the search starts at the $\mathbf{y}_{\bar{k}}$ from the last iteration (see [25] for a thorough discussion).

Having found the closest nodal point $\mathbf{y}_{\bar{k}}$ to $\mathbf{x}$, we search the local part of $\partial\mathbb{K}$ (corresponding to the two segments adjacent to $\mathbf{y}_{\bar{k}}$) to find $\bar{\mathbf{x}}$. This local part of $\partial\mathbf{K}$ may be convex or concave with respect to x (see Fig. A.2).

In case (a) of the figure, definition of g is completely unique. If $\mathbf{x}$ lies in region A or B, g is given by the usual form of the gap function suggested in (A.1). If it lies in region C, $\bar{\mathbf{x}} = \mathbf{y}_{\bar{k}}$, and $g = \|\mathbf{x} - \mathbf{y}_{\bar{k}}\|$. Case (b) provides a potential uniqueness problem. In this case, g_1 and g_2 (given by the normal projections onto the two segments) are computed, and $\bar{\mathbf{x}}$ is determined by the one that is least. If $g_1 = g_2$, then $\mathbf{y}_{\bar{k}}$ is (somewhat arbitrarily) chosen as $\bar{\mathbf{x}}$. This is to preclude any bias for one segment over the other in the subsequent solution.

We see that in the case of a piecewise linear obstacle, the closest point projection problem is quite easily handled, by consideration of a local portion of $\partial\mathbf{K}$ adjacent to $\mathbf{x}$. We further assert that this procedure is essentially equivalent to the so-called 'master-slave' characterization of sliding surfaces discussed, for example, in [25] and in [18]. In this interpretation, $\mathbf{x}$ can be thought of as the current position of a *slave node*, and $\partial\mathbf{K}$ is the *master surface*. This approach is readily extended to multiple deformable bodies by considering the facets comprising $\partial\mathbf{K}$ to be the (linear) element edges of the 'master body' in the current configuration. Thus, the master–slave idea may really be interpreted as a particularization of the concept of closest point projection to the special case where the bodies are discretized by bilinear (four-node) elements.

B. Implementation details of augmented Lagrangians

We now consider briefly the manner in which the augmented Lagrangian method has been incorporated into FEAP, producing the numerical examples given in Sec. 4. This implementation is within the framework described in [18], which treats the case of penalized contact with friction.

In order to motivate the treatment, we begin by consideration of the small deformation frictional contact of *two* contacting bodies, one of which we shall consider to be the 'slave' body (1) and the other of which will be denoted as the 'master' body (2). We generalize eqn (3.6) to two bodies, thereby obtaining

$$G^{(1)}(\mathbf{u}^{(1)}, \delta\mathbf{u}^{(1)}) + G^{(2)}(\mathbf{u}^{(2)}, \delta\mathbf{u}^{(2)}) = \int_{\Gamma^{(1)}} [-t_N^{(1)}\mathbf{n}^{(1)} \cdot \delta\mathbf{u}^{(1)} - \mathbf{t}_T^{(1)} \cdot \delta\mathbf{u}_T^{(1)}]\, d\Gamma + \int_{\Gamma^{(2)}} [-t_N^{(2)}\mathbf{n}^{(2)} \cdot \delta\mathbf{u}^{(2)} - \mathbf{t}_T^{(2)} \cdot \delta\mathbf{u}_T^{(2)}]\, d\Gamma, \tag{B.1}$$

where $(\cdot)^{(1)}$ and $(\cdot)^{(2)}$ denote previously described quantities defined over bodies (1) and (2), respectively. Noting that the integrands on the right-hand side of (B.1) are zero in the regions where no contact occurs, and that where contact does occur the tractions on each body must be equal and opposite, we may write

$$G^{(1)}(\mathbf{u}^{(1)}, \delta\mathbf{u}^{(1)}) + G^{(2)}(\mathbf{u}^{(2)}, \delta\mathbf{u}^{(2)}) = \int_{\Gamma_c} [-t_N^{(1)}\mathbf{n}^{(1)} \cdot \delta(\mathbf{u}^{(1)} - \mathbf{u}^{(2)}) - \mathbf{t}_T^{(1)} \cdot \delta(\mathbf{u}_T^{(1)} - \mathbf{u}_T^{(2)})]\, d\Gamma, \tag{B.2}$$

where Γ_c is the portion of $\Gamma^{(1)}$ and $\Gamma^{(2)}$ where contact occurs.

In the usual master-slave implementation (including [18]) the right-hand side integral is replaced by a summation over all slave nodes (at current positions $\mathbf{x}_s$) which takes the following form

$$\int_{\Gamma_c} [-t_N^{(1)}\mathbf{n}^{(1)} \cdot \delta(\mathbf{u}^{(1)} - \mathbf{u}^{(2)}) - \mathbf{t}_T^{(1)} \cdot \delta(\mathbf{u}_T^{(1)} - \mathbf{u}_T^{(2)})]\, d\Gamma \Rightarrow \sum_{s=1}^{n_s} [-t_{N_s}\delta g_{N_s} - t_{T_s}\delta g_{T_s}], \tag{B.3}$$

where

n_s = the number of slave nodes,

g_{N_s} = the normal gap function (denoted as g in Appendix A) for slave node s,

g_{T_s} = the tangential gap function, a new quantity.

We note that the transition suggested in (B.3), while presented in a small-deformation context, results in an expression [the right-hand side of (B.3)] which is suitable for large deformation problems, provided the involved forces and gaps are defined appropriately. Also, we remark that t_{T_s} is now a *scalar*, and δg_{T_s} represents the variation of g_{T_s}, which we consider to be a tangential gap function. As defined in [18], g_{T_s} has the interpretation of being the distance between the current projection of the material point at $\mathbf{x}_s$ on the master surface and the projection of the same

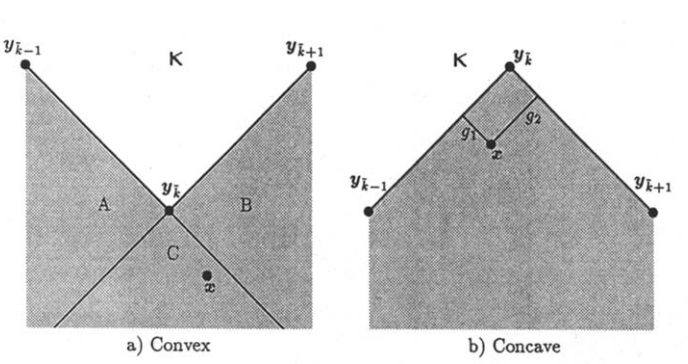

Fig. A.2. Prospective local geometries for closest point determination.

material point on the master surface at the end of the last time step, with this distance between the two projections being evaluated in the converged geometry of the last time step. Put more simply, g_{T_s} is a measure of the *relative* tangential motion between $\mathbf{x}_s$ and the master surface which occurs during the time step, just as (B.2) suggests it should be.

In this framework, implementation of penalty and augmented Lagrangian techniques is very simple, since the constitutive relations governing the interface become *scalar* equations. The normal response is handled exactly as discussed previously, while the tangential (frictional) response is treated analogously to the previous presentations where $\mathbf{t}_T$ was the quantity of interest. Here, t_{T_s} plays the role of $\mathbf{t}_T$, and g_{T_s} plays the role of $\mathbf{u}_{T_{n+1}} - \mathbf{u}_{T_n}$. Thus, Tables 2 and 3 still contain the relevant aspects of the implementation of augmented Lagrangians in this setting. All constitutive equations for the friction law are merely replaced by their scalar counterparts, and these equations are assumed to hold for quantities defined at the slave nodes.

Finally, we remark that the matrix equations emanating from the last expression in (B.3) are obtained by taking the indicated variations. Since the gap functions depend on the local geometries of both bodies, this operation generates the contact forces for both slave and master bodies. The calculations needed to do this and the subsequent linearizations, although lengthy, are standard and are given in detail in [18].