Constitutive Model and Finite Element Procedure for Dilatant Contact Problems

By Michael E. Plesha,[1] Associate Member, ASCE, Roberto Ballarini,[2] Associate Member, ASCE, and Atul Parulekar[3]

Abstract: A constitutive law for dilatant frictional behavior is reviewed. It is developed by distinguishing between the macrostructural and microstructural features of a material discontinuity. Macrostructural considerations provide the general form of the constitutive equations, while microstructural considerations allow the inclusion of an appropriate surface idealization. The result is an incremental relation between contact stresses (traction) and relative surface deformation that accounts for phenomena such as surface damage due to wear and arbitrary cyclic sliding. A quadratic-displacement-isoparametric finite element is derived that permits modeling of curved-contact surfaces and crack surfaces terminating at a tip with a surrounding medium that is modeled with quarter-point quadratic elements. Emphasis is on the use of established finite-element-solution methodologies and program architecture for material-nonlinear problems. Several examples are considered. The resulting methodology is useful for modeling geologic discontinuities, crack-shear transfer in concrete, and dilatancy-induced mixed-mode fracture mechanics.

Introduction

Material interfaces are common in mechanical systems and media and often have a substantial influence on response. The behavior of a material discontinuity is complex and involves frictional sliding, possible contact-surface separation, sometimes dilatancy, and usually various types of surface damage that affect subsequent behavior of the discontinuity. Because quantitative expressions for such behavior have been lacking, some of these phenomena have gone unaccounted for in analyses, and most often, a discontinuity has been idealized as being smooth with simple Coulomb friction. Even with simple Coulomb friction, because of nonlinearity, finite-element-solution methodologies are still not advanced to the point where contact-friction capabilities are included in general-purpose programs, and in most cases, special-purpose programs are used.

The contact problems considered in this paper have surface roughness that is small compared with the macroscopic contact area and have well-defined normal and tangent directions to the macroscopic contact surface. In addition, we restrict attention to problems where the initial mating, or correlation, between the contact surfaces is close. Such a situation is shown in Fig. 1(*a*), which is characteristic of naturally generated material discontinuities, such as crack surfaces, which propagate through an initially continuous medium. Examples include cracks in polycrystalline and aggregate materials,

[1]Assoc. Prof., Dept. of Engrg. Mech., Univ. of Wisconsin, Madison, WI 53706.
[2]Asst. Prof., Dept. of Civ. Engrg., Case Western Reserve Univ., Cleveland, OH 44106.
[3]Res. Asst., Dept. of Civ. Engrg., Case Western Reserve Univ., Cleveland, OH.

Note. Discussion open until May 1, 1990. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on September 12, 1988. This paper is part of the ***Journal of Engineering Mechanics***, Vol. 115, No. 12, December, 1989. ©ASCE, ISSN 0733-9399/89/0012-2649/$1.00 + $.15 per page. Paper No. 24117.

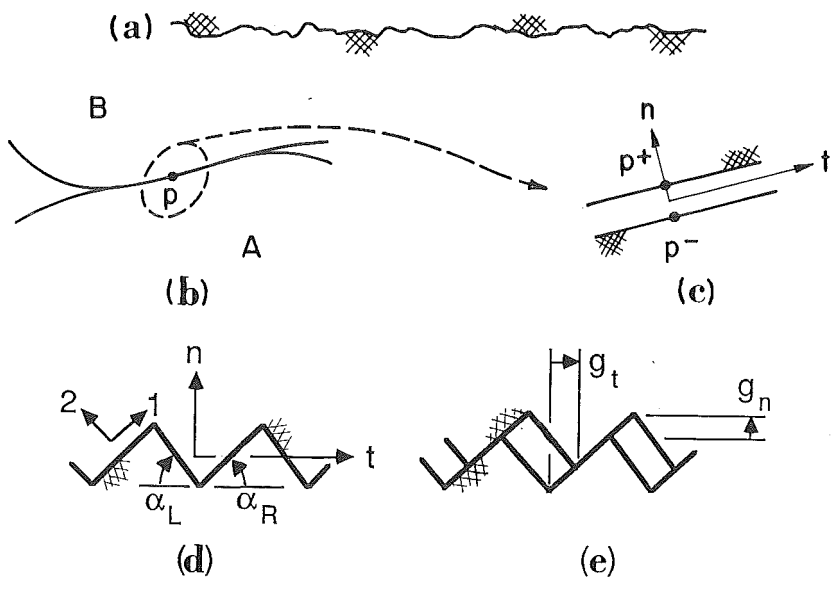

FIG. 1. (*a*) Contact Surface Profile with Very Close Initial Mating of Asperities; (*b*) Two-Body Contact in Two Dimensions along Macroscopically Smooth Surface; (*c*) Definition of Coordinate System at Point p (Surfaces are Shown Separated for Clarity and Surface Roughness Is Not Shown); (*d*) Possible Microstructural Idealization of Surface Profile Shown in Fig. 1(*a*); and (*e*) Geometry of Surfaces after Deformation when There is No Asperity Damage

such as ceramics and concrete, and geologic discontinuities, such as rock joints and faults. The most important behavioral feature that these contact problems display is dilatancy—the coupling between normal and tangential relative displacements due to asperities of one surface riding up on those of the other surface.

In this paper, a modern two-dimensional incremental constitutive law for contact-friction, analogous to the incremental theory of plasticity, is reviewed, and a finite element spatial-discretization procedure is developed; extension to three-dimensions is straightforward. The constitutive theory was fully developed in Plesha (1987) in the spirit of the original work of Seguchi et al. (1974), Fredriksson (1976), Michalowski and Mroz (1978), Curnier (1984), and Cheng and Kikuchi (1985). A finite element spatial discretization for a two-dimensional contact region is presented with emphasis on straightforward numerical implementation, using standard finite-element-solution procedures and program architecture. The element derived is particularly useful for modeling curved-contact surfaces and for mixed-mode fracture mechanics problems. Several examples showing the performance of this modeling approach are also presented.

Contact-Friction Constitutive Law

A number of constitutive laws for dilatant contact problems have been proposed for geologic discontinuities (Plesha 1987) and crack surfaces in

concrete (Bažant et al. 1980, 1984; Riggs and Powell 1986; Divakar et al. 1987). Most of these models are deformation-theory models that are valid for unidirectional sliding and have stiffness coefficients (i.e., the $\mathbf{E}^{ep}$ coefficients of Eq. 7) that are determined empirically or by curve-fitting experimental data. The constitutive law adopted in this paper is valid for arbitrary sliding histories and has material parameters that can be determined from conventional direct-shear tests. This constitutive model is analogous to the theory of continuum elastoplasticity and can be termed a continuum theory of friction in the sense that contact area is continuous; hence, the traction components that the theory predicts are also continuous. A detailed description of the theory is given in Plesha (1987), however, a brief description of the model is appropriate. The theory is formulated by distinguishing between macrostructural and microstructural features of an interface as follows.

Macrostructural Considerations

Fig. 1(*b*) shows a macroscopically smooth contact surface in two dimensions with local tangential and normal coordinate directions t and n, respectively, with origin at point p^- which is affixed to body A [Fig. 1(*c*)]. Roughness on the contact surface is not shown in Fig. 1(*c*) and will be discussed subsequently as a microstructural feature. The only requirement of the macroscopic contact surface is sufficient smoothness so that the surface tangent and normal are not ill-defined.

The kinematic variables used in the constitutive law are the relative surface displacements in the tangential and normal directions that are defined as

$$g_t = (\vec{\mathbf{u}}_B - \vec{\mathbf{u}}_A) \cdot \vec{\mathbf{t}} \quad \ldots\ldots (1)$$

$$g_n = (\vec{\mathbf{u}}_B - \vec{\mathbf{u}}_A) \cdot \vec{\mathbf{n}} \quad \ldots\ldots (2)$$

where $\vec{\mathbf{u}}_A$ and $\vec{\mathbf{u}}_B$ = the displacement vectors of the points p^- and p^+ associated with bodies A and B, respectively; and $\vec{\mathbf{t}}$ and $\vec{\mathbf{n}}$ = unit vectors in the tangent and normal directions of the interface at point p. The tangent and normal stresses the interface supports at point p are denoted by σ_t and σ_n with the convention that compressive stresses are negative (in proper terminology, these are traction components σ_{nt} and σ_{nn}, respectively, but we adopt the more conventional nomenclature of tangent and normal stresses).

A basic assumption in the theory is that the deformation can be additively decomposed into

$$g_i = g_i^e + g_i^p \qquad i = t, n \quad \ldots\ldots (3)$$

where superscripts e and p = the elastic (recoverable) and plastic (irrecoverable) parts of the deformation; and $\mathbf{i}$ = a vector component in the tangent or normal direction. There exists experimental evidence supporting this decomposition for almost every class of friction problem that has been carefully studied. Furthermore, it leads to a more convenient numerical implementation compared to frictional idealizations in which a stick condition precedes frictional sliding.

Assumably, the stress supported by an interface relates to the elastic part of Eq. 3 by

$$\dot{\sigma}_i = E_{ij}\dot{g}_j^e \quad \ldots\ldots (4)$$

where the E_{ij} = interface stiffnesses; superposed dots = time differentiation;

and the summation convention is applied to repeated indices. The stiffnesses appearing in Eq. 4 can be interpreted from two points of view. In the first, as discussed by Oden and Campos (1981), Cheng and Kikuchi (1985), and others, the stiffnesses are penalty numbers that approximately enforce contact-surface compatibility consisting of impenetrability and presliding stick constraints. Using this point of view, $E_{tn} = E_{nt} = 0$ and E_{tt} and E_{nn} are taken to be much larger than the stiffness of the contacting materials. For example, impenetrability at point p in Fig. 1(*c*) requires $g_n = \sigma_n/E_{nn} \geq 0$. With the convention that a compressive stress is negative, impenetrability is only satisfied in the limit of infinite E_{nn}. For finite values of E_{nn}, compatibility is violated; by making E_{nn} large in relation to other stiffnesses in the model, the incompatibility is typically insignificant. Kikuchi and Song (1981) have established that solutions exist for finite penalty numbers and convergence, when the penalty numbers become infinite. In the second point of view, these stiffnesses are interpreted as physically significant properties of a material discontinuity. Inspection of a contact indicates that while the contact zone is usually very thin, it does have finite thickness. Results obtained from carefully conducted experiments on a variety of contact problems indicate that reversible deformability of the contact zone occurs at all load levels, including presliding load levels (Goodman 1980). Based on physical considerations, it is appropriate to take $E_{nt} = E_{tn} = 0$ so that changes of stress in the tangential and normal directions are unrelated to changes of elastic deformation in the normal and tangential directions, respectively. Data extracted from tests for E_{tt} and E_{nn} typically provide relatively large stiffnesses; hence, small negative values of g_n, are possible and are considered to be physically realizable deformations. Based on experiment, stiffnesses are stress-dependent, although for practical applications it is not known if or when such detail is warranted. In the following, we assume uniform stiffnesses.

The plastic deformations arise from sliding and sliding-associated damage, and are assumed to be given by the sliding rule

$$\dot{g}_i^p = 0 \quad \text{if } F < 0 \quad \text{or } \dot{F} < 0 \qquad (5a)$$

$$\dot{g}_i^p = \dot{\lambda} \frac{\partial G}{\partial \sigma_i} \quad \text{if } F = \dot{F} = 0 \qquad (5b)$$

where F = a scalar-valued slip function with a negative value for nonsliding states of stress and zero for states of stress producing slip (positive values of F are undefined); G = a slip potential, with a gradient that gives the direction of slip; and $\dot{\lambda}$ = a non-negative slip multiplier that gives the magnitude of the slip. When $F = G$, the sliding rule is associated, and when $F \neq G$, the sliding rule is nonassociated. Eq. 5 is analogous to the flow rule used in the incremental theory of plasticity. Although plasticity and friction share many features, it has long been recognized that friction is strongly nonassociated (Drucker 1954) because the direction of sliding is pressure-independent.

The following assumes that hardening or softening behavior due to sliding-induced damage is strictly a function of the sliding work, W^p, where the rate of this quantity is $\dot{W}^p = \sigma_i \dot{g}_i^p$. Under this assumption, damage accumulates more rapidly for severe states of stress than for mild states, which is consistent with physical notions of wear. Some writers consider the effective slip, $\bar{g}^p$; $\dot{\bar{g}}^p = \sqrt{\dot{g}_i^p \dot{g}_i^p}$, as a measure of wear, but this may not be appropriate,

because it does not incorporate the effects of the severity of stress on wear. For example, consider two contact situations that are identical except that one has low-compressive stress while the other has high-compressive stress. Both situations undergo the same tangential-displacement history, and hence have almost identical effective slip. Clearly, it would be unreasonable to expect both situations to have the same damage due to wear.

The consistency equation is obtained by noting that if at a given instant in time, slip is imminent, $F = 0$, and, if at the next instant in time, the interface remains critical, then $\dot{F} = 0$, which can be written as

$$\frac{\partial F}{\partial \sigma_i}\dot{\sigma}_i + \frac{\partial F}{\partial W^p}\dot{W}^p = 0 \quad \ldots\ldots (6)$$

Combining Eqs. 3–6 and eliminating the slip multiplier leads to the constitutive law

$$\dot{\boldsymbol{\sigma}} = \mathbf{E}^{ep}\dot{\mathbf{g}} \quad \ldots\ldots (7)$$

where

$$\mathbf{E}^{ep} = \mathbf{E} \qquad \text{if } F < 0 \quad \text{or } \dot{F} < 0 \quad \ldots\ldots (8a)$$

$$\mathbf{E}^{ep} = \mathbf{E}\left(\mathbf{I} - \frac{\dfrac{\partial G}{\partial \boldsymbol{\sigma}}\dfrac{\partial F^T}{\partial \boldsymbol{\sigma}}\mathbf{E}}{\dfrac{\partial F^T}{\partial \boldsymbol{\sigma}}\mathbf{E}\dfrac{\partial G}{\partial \boldsymbol{\sigma}} - H}\right) \qquad \text{if } F = \dot{F} = 0 \quad \ldots\ldots (8b)$$

and the hardening or softening parameter is

$$H = \frac{\partial F}{\partial W^p}\boldsymbol{\sigma}^T\frac{\partial G}{\partial \boldsymbol{\sigma}} \quad \ldots\ldots (8c)$$

and T denotes transposition. Forms for F and G are problem-dependent but are generally derivable from, or related to, Coulomb's friction law in conjunction with an idealization for the contact-surface microstructure.

Eq. 7 is easy to evaluate and provides a clear relation between stresses and deformation that is valid for arbitrary sliding histories. For this reason, the constitutive model is ideal for implementation in analysis software. Unfortunately, during sliding the material matrix $\mathbf{E}^{ep}$ is asymmetric if $F \neq G$, which is usually the case.

Microstructural Considerations

Expressions for F and G can be obtained by considering the idealized contact-surface profile shown in Fig. 1(*d*). This model consists of sawtooth-asperity surfaces that degrade. If the friction on the active asperity surface is governed by Coulomb's law, then $|\sigma_1| \leq -\mu\sigma_2$, where μ is the coefficient of friction, and σ_1 and σ_2 are the tangential and normal stresses on the asperity surface. By transformation, this equation can be expressed in terms of the macroscopic stresses σ_t and σ_n. This leads to (Plesha 1987):

$$F = |\sigma_n \sin\alpha_k + \sigma_t \cos\alpha_k| + \mu(\sigma_n \cos\alpha_k - \sigma_t \sin\alpha_k) \quad \ldots\ldots (9a)$$

$$G = |\sigma_n \sin\alpha_k + \sigma_t \cos\alpha_k| \quad \ldots\ldots (9b)$$

where $k = L$ or R, depending on which asperity surface is active (i.e., whether $g_t \geq 0$ or $g_t < 0$, respectively).

For discontinuities, such as joints in rock and crack surfaces in concrete, there is extensive experimental data showing that asperity surfaces degrade. The asperity behavior in this model is characterized as follows: Under high-compressive stresses, high-tangential stresses are required to produce slip, and rapid asperity degradation can occur for small tangential displacements. Under low-compressive stresses, low-tangential stresses will produce slip; yet if the amount of slip is large, then asperity degradation can arise from wear. A simple tribological relation that replicates these characteristics is obtained by assuming that degradation is a function of sliding work

$$\alpha_k = (\alpha_k)_0 \exp(-cW^p) \quad (10)$$

where $(\alpha_k)_0$ = the initial asperity angle; and c = an experimentally determined positive constant that reflects how rapidly the surfaces degrade. Thus, if a particular degree of degradation is obtained by a state of high stress and low displacement, the same degree can be obtained at a lower stress level provided the displacement is large enough.

As discussed in Plesha (1987) and Zubelewicz et al. (1987), under conditions of cyclic shearing, this constitutive law may overpredict bulking (i.e., thickening of an interface due to the larger volume of space occupied by the rubbled asperity material relative to its initial, intact volume). A modification to the constitutive law reported herein, which removes this deficiency, is developed in Zubelewicz et al. (1987).

Corroboration of this constitutive law with rock-joint behavior, using Eqs. 7–10, has been carried out in Plesha (1987), Zubelewicz, et al. (1987), and Plesha and Haimson (1988) with extremely good agreement.

Finite Element Implementation

An extensive review of the published writing on interface finite elements, with particular reference to rock-mechanics problems, is reported in Heuze and Barbour (1981). We present the development of a two-dimensional, isoparametric, quadratic displacement-interface finite element for contact-friction problems. Goodman et al. (1968) reported a similar element. In this paper, however, we emphasize a consistent isoparametric formulation that permits modeling of curved-contact surfaces and provides an element that is compatible with the quadratic-displacement-plane finite elements that are often used to model crack-tip stress singularities. Significantly, this finite element, in conjunction with the constitutive law, permits the treatment of contact-friction problems with the same methodologies and program architecture as those currently used for small-displacement material nonlinear problems.

The macroscopic contact surface shown in Fig. 1(*b*) is discretized into interface finite elements. One such element is shown in Fig. 2(*a*). The element can be curved and is assumed to have zero thickness in the undeformed state, but is shown separated in Fig. 2(*a*) for clarity. The element has six nodes, and each node has two degrees of freedom corresponding to horizontal and vertical displacements. Node pairs 1 and 6, 2 and 5, and 3 and 4 are assumed to be spatially coincident prior to deformation; nodes 2 and 5 need not be at the midpoint of the element's arc length, although accuracy is probably best when this is the case. Compatibility of displacements is

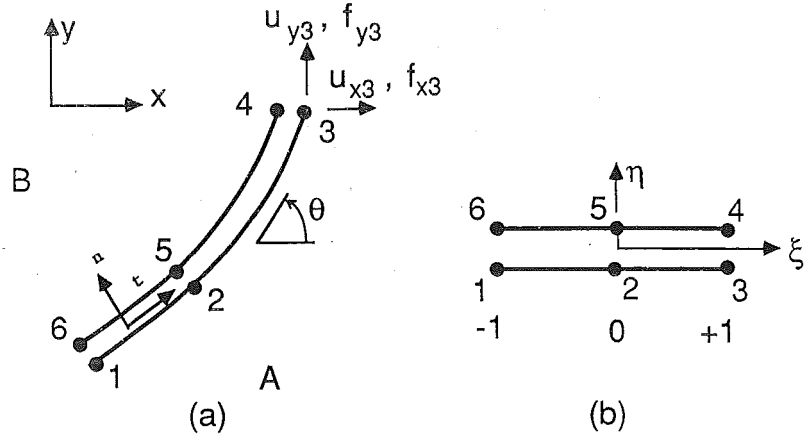

FIG. 2. (*a*) Geometry of Six-Node, Quadratic-Displacement-Isoparametric-Interface Finite Element in Global, or Physical, Coordinate System; and (*b*) Geometry of Reference Element

assured (excepting negative values of g_n as discussed earlier) when the interface element adjoins quadratic-edge-displacement continuum elements (e.g., the quadratic-displacement eight-node serendipity quadrilateral, nine-node Lagrange quadrilateral, and six-node triangle) such that end nodes of the interface element correspond to corner nodes of the continuum elements, and midpoint nodes of the interface element correspond to midpoint nodes of the continuum elements.

Geometric Mapping

The global, or physical, coordinates of the element are given parametrically in terms of the reference geometry shown in Fig. 2(*b*) and the reference coordinate ξ by the mapping

$$x = (N_1, N_2, N_3)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathbf{NX} \quad \ldots\ldots (11a)$$

$$y = (N_1, N_2, N_3)\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \mathbf{NY} \quad \ldots\ldots (11b)$$

where the shape functions $\mathbf{N}$ are given by

$$\mathbf{N} = \left[\frac{\xi(\xi - 1)}{2}, -(\xi + 1)(\xi - 1), \frac{\xi(\xi + 1)}{2}\right] \quad \ldots\ldots (12)$$

Since the two contact surfaces are assumed to be initially coincident, the spatial coordinates of only points one through three are necessary to define the shape of the element. The slope of the element, s, at any point is given by

$$s = \frac{dy}{dx} = \frac{\dfrac{dy}{d\xi}}{\dfrac{dx}{d\xi}} = \frac{\mathbf{N},_{\xi}\mathbf{Y}}{\mathbf{N},_{\xi}\mathbf{X}} \quad \ldots\ldots (13)$$

where $\mathbf{N},_{\xi} = d/d\xi N$. The tangent and normal unit vectors along the contact surface shown in Fig. 2(*a*) are given in terms of the slope by

$$\vec{\mathbf{t}} = \frac{(\vec{\mathbf{i}} + s\vec{\mathbf{j}})}{\sqrt{1+s^2}} \qquad \vec{\mathbf{n}} = \frac{(-s\vec{\mathbf{i}} + \vec{\mathbf{j}})}{\sqrt{1+s^2}} \quad \ldots\ldots (14)$$

where $\vec{\mathbf{i}}$ and $\vec{\mathbf{j}}$ = unit vectors in the global x and y directions, respectively. For future use, vector scalar products between $\vec{\mathbf{t}}$, $\vec{\mathbf{n}}$, and global vectors $\vec{\mathbf{i}}$, $\vec{\mathbf{j}}$ are

$$\vec{\mathbf{i}} \cdot \vec{\mathbf{t}} = \cos\theta = \frac{1}{\sqrt{1+s^2}} \qquad \vec{\mathbf{i}} \cdot \vec{\mathbf{n}} = -\sin\theta \quad \ldots\ldots (15a)$$

$$\vec{\mathbf{j}} \cdot \vec{\mathbf{t}} = \sin\theta = \frac{s}{\sqrt{1+s^2}} \qquad \vec{\mathbf{j}} \cdot \vec{\mathbf{n}} = \cos\theta \quad \ldots\ldots (15b)$$

The increment of arc length dt is related to the increment of reference length $d\xi$ by

$$dt = Jd\xi \quad \ldots\ldots (16)$$

where the Jacobian is given by

$$J = \left[\left(\frac{dx}{d\xi}\right)^2 + \left(\frac{dy}{d\xi}\right)^2\right]^{1/2} \quad \ldots\ldots (17)$$

Combining Eqs. 11*a*, 11*b*, and 17 yields

$$J = (\mathbf{X}^T\mathbf{N}^T_{,\xi}\mathbf{N},_{\xi}\mathbf{X} + \mathbf{Y}^T\mathbf{N}^T_{,\xi}\mathbf{N},_{\xi}\mathbf{Y})^{1/2} \quad \ldots\ldots (18)$$

which is easy to evaluate for specified ξ given $\mathbf{X}$ and $\mathbf{Y}$.

Kinematics

Eqs. 1 and 2 can be rewritten as

$$g_{\mathbf{t}} = (u_x^B - u_x^A)\vec{\mathbf{i}} \cdot \vec{\mathbf{t}} + (u_y^B - u_y^A)\vec{\mathbf{j}} \cdot \vec{\mathbf{t}} \quad \ldots\ldots (19a)$$

$$g_{\mathbf{n}} = (u_x^B - u_x^A)\vec{\mathbf{i}} \cdot \vec{\mathbf{n}} + (u_y^B - u_y^A)\vec{\mathbf{j}} \cdot \vec{\mathbf{n}} \quad \ldots\ldots (19b)$$

Combining Eqs. 15*a*, 15*b*, 19*a*, and 19*b* provides

$$\begin{pmatrix} g_{\mathbf{t}} \\ g_{\mathbf{n}} \end{pmatrix} = \mathbf{T}(u_x^A, u_x^B, u_y^A, u_y^B)^T \quad \ldots\ldots (20)$$

where

$$\mathbf{T} = \begin{pmatrix} -\cos\theta & \cos\theta & -\sin\theta & \sin\theta \\ \sin\theta & -\sin\theta & -\cos\theta & \cos\theta \end{pmatrix} \quad \ldots\ldots (21)$$

The finite-element-displacement interpolation is

$$u_x^A = (N_1, N_2, N_3)\begin{pmatrix} d_{x1} \\ d_{x2} \\ d_{x3} \end{pmatrix} \qquad u_x^B = (N_1, N_2, N_3)\begin{pmatrix} d_{x6} \\ d_{x5} \\ d_{x4} \end{pmatrix} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (22)$$

where the shape functions are the same as those used for the geometric mapping and are given by Eq. 12. Expressions for u_y^A and u_y^B are identical to Eq. 22 except that x is replaced by y. Collectively, the displacement interpolation is written as

$$(u_x^A, u_x^B, u_y^A, u_y^B)^T = \mathbf{MD} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (23)$$

where

$$\mathbf{M} = \left[\begin{array}{cc|cc|cc|cc|cc|cc} N_1 & 0 & N_2 & 0 & N_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & N_3 & 0 & N_2 & 0 & N_1 & 0 \\ 0 & N_1 & 0 & N_2 & 0 & N_3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & N_3 & 0 & N_2 & 0 & N_1 \end{array}\right] \quad \ldots\ldots\ldots \quad (24)$$

$$\mathbf{D} = (d_{x1}, d_{y1}, d_{x2}, d_{y2}, \ldots, d_{x6}, d_{y6})^T \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (25)$$

Combining Eqs. 20 and 23 gives

$$\begin{pmatrix} g_t \\ g_n \end{pmatrix} = \mathbf{g} = \mathbf{BD} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (26)$$

where $\mathbf{B} = \mathbf{TM}$ is analogous to the strain-displacement matrix used in the formulation of continuum finite elements except that it does not involve differential operators applied to shape functions.

Energy Balance

The work of the equivalent nodal forces must be equal to the work of the surface-stress distribution for an arbitrary displacement increment

$$d\mathbf{D}^T\mathbf{F}^j = \int_0^l d\mathbf{g}^T\boldsymbol{\sigma}^j dt \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (27)$$

where d = an increment; l = the element's arc length; j = pseudotime; and the nodal forces are

$$\mathbf{F} = (f_{x1}, f_{y1}, f_{x2}, f_{y2}, \ldots, f_{x6}, f_{y6})^T \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (28)$$

Combining Eqs. 26 and 27, using $\boldsymbol{\sigma}^j = \boldsymbol{\sigma}^{j-1} + d\boldsymbol{\sigma}$, and noting the arbitrariness of $d\mathbf{D}$ yields

$$\mathbf{F}^j = \int_0^l \mathbf{B}^T(\boldsymbol{\sigma}^{j-1} + d\boldsymbol{\sigma})dt \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (29)$$

Noting that $\mathbf{F}^{j-1} = \int_0^l \mathbf{B}^T\boldsymbol{\sigma}^{j-1}dt$ and combining with Eq. 16, the above equation becomes

$$\mathbf{F}^j = \mathbf{F}^{j-1} + \int_{-1}^{+1} \mathbf{B}^T d\boldsymbol{\sigma} J d\xi \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (30)$$

which is an expression commonly derived and conveniently used in transient finite element analysis of material-nonlinear problems with explicit time integration (Belytschko et al. 1976; Belytschko and Hughes 1983). With $d\boldsymbol{\sigma} = \mathbf{E}^{ep}\mathbf{B}d\mathbf{D}$ where $\mathbf{E}^{ep}$ is given by Eqs. 8*a–c*, Eq. 30 becomes

$$\mathbf{F}^j = \mathbf{F}^{j-1} + \mathbf{K}^{ep}d\mathbf{D} \quad \text{(31)}$$

where the tangent stiffness matrix is

$$\mathbf{K}^{ep} = \int_{-1}^{+1} \mathbf{B}^T\mathbf{E}^{ep}\mathbf{B}Jd\xi \quad \text{(32)}$$

Eq. 32 is convenient for use in quasi-static finite element programs and in transient finite element programs that use implicit integration.

As discussed earlier, the physics of frictional phenomena dictate that $\mathbf{E}^{ep}$, and hence $\mathbf{K}^{ep}$, usually are asymmetric. For explicit transient analysis, Eq. 30 is usually used with element-by-element computations in which a global-stiffness matrix is never formed or stored; hence, the lack of symmetry is of little consequence. The implications for quasi-static and implicit transient finite element analysis are more serious, although they are usually tolerable. Generally speaking, twice the amount of storage and about twice the amount of time for factorization are required for the tangent-stiffness matrix. To avoid the increased storage penalty, a symmetric equation-solution method for nonassociated plasticity such as proposed by Pande and Pietruszczak (1986) may be effective.

As with most isoparametric elements, Eqs. 30 and 32 must be evaluated using numerical integration. For an undistorted element (i.e., a straight element with midside nodes equidistant between ends) and uniform material matrix, the integrand of Eqs. 30 and 32 is a quartic polynomial that is exactly integrated using a three-point Gauss quadrature formula. Arguments can be made for using a rule with more integration points (but not necessarily a higher degree of precision) so that the material matrix $\mathbf{E}^{ep}$ can be sampled at more locations (Bathe 1982). However, using a similar finite element formulation, Kikuchi (1982) and Oden and Kikuchi (1982) advocate the use of reduced integration to avoid oscillating normal-stress distributions that result from full integration. A limited number of contact problems with the element reported herein have been solved by using two-, three-, and four-point Gauss rules, and Simpson's rule; in all cases except the first, smooth and accurate stress distributions were obtained. However, the question of what integration scheme gives the best results and optimal rate of convergence is still an open question.

The significant feature of this development is that small-displacement contact-friction problems, with their associated nonlinearities, can be treated as a material-nonlinear problem using existing finite-element-solution methodologies. In the examples that follow, standard techniques were used.

Numerical Results

Three examples illustrating the performance of the constitutive law and finite element for dilatant contact problems are considered. The first two examples are numerical simulations of laboratory direct-shear tests on initially fully mated rock joints. The third example considers the problem of

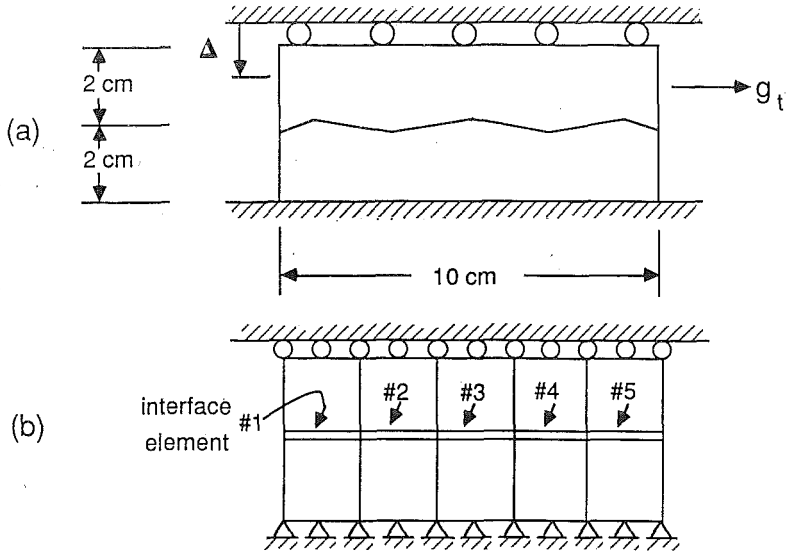

FIG. 3. Model and Mesh for the First Example: (*a*) Geometry of Rock Joint in Laboratory Direct-Shear Test with Constrained Dilation; and (*b*) Finite Element Model

determining the mixed-mode stress-intensity factors for a rough and tortuous crack subjected to a far-field shear stress. In all of the examples, standard incremental-solution procedures were used. In the first and second examples, incremental displacements were prescribed, and in the third example, incremental tractions were prescribed. A Gauss elimination equation solver for asymmetric matrices was used. For determination of the interface-element-stiffness matrices, $\mathbf{E}^{ep}$ was computed using the state of stress from the previous increment, and no iterations were performed within an increment to update the state of stress. Although this method for integration of the constitutive law is primitive, it was nevertheless satisfactory for the examples considered here. For general applications, more accurate and robust integration methods such as those reported in Ortiz and Popov (1985) are recommended. Interface elements were integrated using a three-point Gauss rule, and at integration points in which tension was detected ($\sigma_n > 0$), $\mathbf{E}^{ep}$ was taken as null. The solutions were relatively insensitive to the number of increments (except for the second example, where the onset of softening behavior was sensitive to the number of steps). The first two examples were repeated with 25, 50, and 100 incremental steps; results shown are for 100 steps. The third example used 45 incremental steps.

Example 1: Constrained Dilation Direct-Shear Test

Fig. 3(*a*) shows a model of a direct-shear test for an initially fully mated rock joint with asperity surfaces oriented at 10° with respect to the horizontal. The finite element discretization is shown in Fig. 3(*b*). The upper and lower plates, which are modeled using 10 eight-node quadrilateral (quadratic displacement) isoparametric elements, are separated by five interface ele-

ments. The bottom plate of rock is fixed, and the upper plate, after being compressed by an amount Δ sufficient to produce an initial compressive stress $\sigma_n = -5$ MPa, is permitted horizontal motion only. The physical parameters employed in the simulation are realistic for hard rock and are as follows: For the interface $E_{tt} = 5$ GN/m^3; $E_{nn} = 10$ GN/m^3; $\alpha = 10°$; $\mu = 0.5$; and $c = 0.1$ cm^2/joule. For the upper and lower plates of rock, Young's modulus $E = 10$ GPa; and Poisson's ratio $\nu = 0.3$. Since the intact rock material is much stiffer than the deformable joint, dilatancy during tangential sliding is almost completely constrained, i.e., $dg_n \cong 0$, and the phenomenon of dilatancy manifests itself through increasing compressive stresses in the presence of asperity degradation according to Eq. 10.

The loading consists of a prescribed displacement history of the top surface of the upper plate of rock. The average tangential stress and average normal stress at the interface resulting from this motion is shown in Fig. 4(*a*) and the average degradation of the asperity-surface angle is shown in Fig. 4(*b*). Although they are not presented here, the numerical results obtained in this simulation indicate that except near the ends of the interface, the average values of the stresses and displacements differ only slightly from the local values. Also shown in Figs. 4(*a*) and 4(*b*) are the results obtained by directly evaluating Eq. 7 with the constraint $g_n = -\Delta$. While corroborating quantitative experimental data are not available for this simulation, the results show good qualitative agreement with real joint behavior (Goodman 1976) where the normal and tangential stresses increase with increasing tangential displacement as a result of the boundary-condition restraining dilation. Furthermore, the rates of increase become smaller at large displacements because of the asperity degradation.

Example 2: Unconstrained Dilation Direct-Shear Test

In this example a simulation is performed of a laboratory direct-shear test at constant compressive stress reported by Bandis et al. (1981). The geometry of the specimen and model is the same as that in the first example, Figs. 3(*a*) and 3(*b*), except that the boundary condition along the top surface of the upper plate of rock consists of a uniform compressive load of $(\sigma_n)_0 = -90$ kPa. The bottom plate of rock is fixed and the shear-displacement history of the upper plate is prescribed. The physical parameters used simulate the soft model material reported in Bandis et al. (1981) and are as follows: $E_{tt} = 0.08$ GN/m^3; $E_{nn} = 1.0$ GN/m^3; $\alpha = 16°$; $\mu = 0.625$; and $c = 10$ cm^2/joule. For the interface, and for the intact rock, $E = 1$ GPa; and $\nu = 0.3$.

The average stresses as functions of tangential displacement are shown in Fig. 5(*a*), and the degradation of the asperities at various locations along the interface is shown in Fig. 5(*b*). Because the upper plate is not restrained in the vertical direction, a pronounced bending effect exists which leads to a nonuniform distribution of stresses where the left side of the interface experiences less compression than the right side, and hence less deterioration of the asperities. Moreover, at large tangential displacements, the numerical results indicate that separation at the left side of the interface occurred. Also shown in these figures are the results obtained by directly using Eq. 7. Discrepancies arise because the finite element results are for a finite-size specimen and include variations of stress and damage along the interface due to bending, while the direct use of Eq. 7 assumes a uniform state of stress

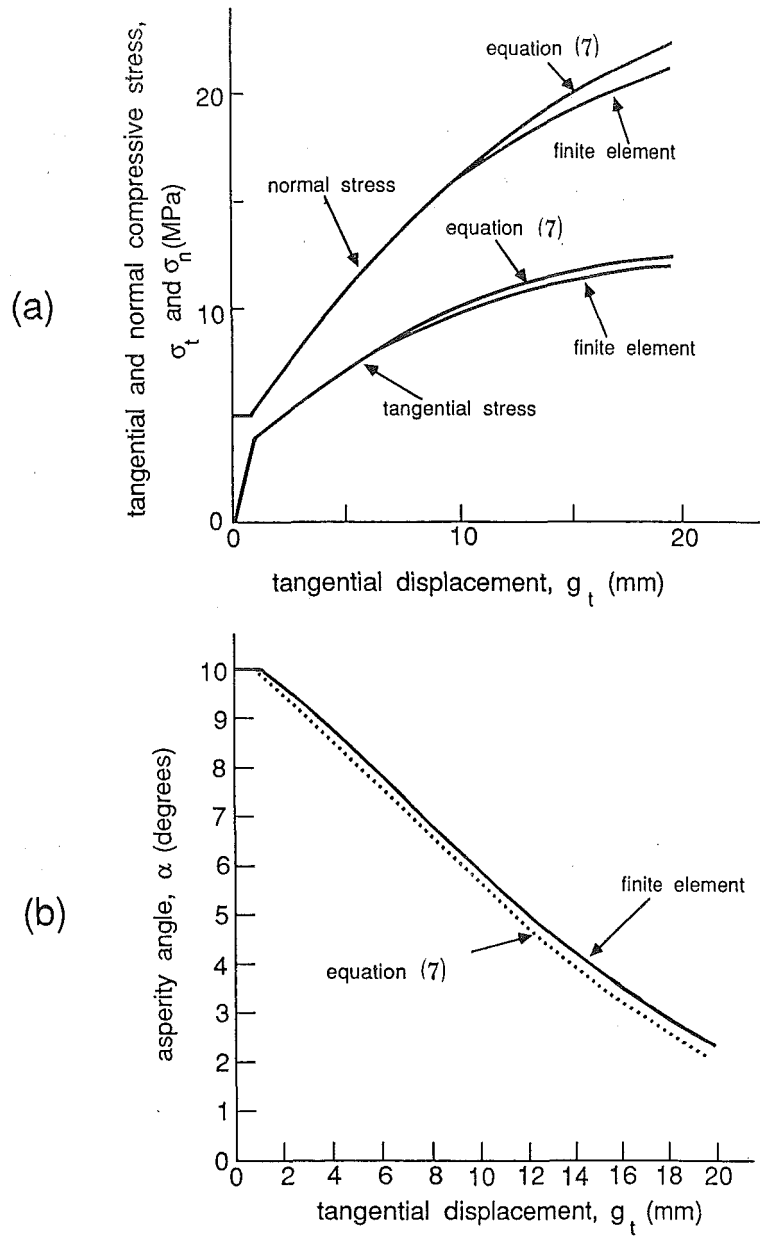

FIG. 4. Finite Element Results for First Example as Function of Tangential Displacement: (*a*) Tangential and Normal Stress; and (*b*) Asperity Degradation Due to Damage; Also Shown Are Results Obtained by Directly Evaluating Eq. 7 with Constraint $g_n = -\Delta$

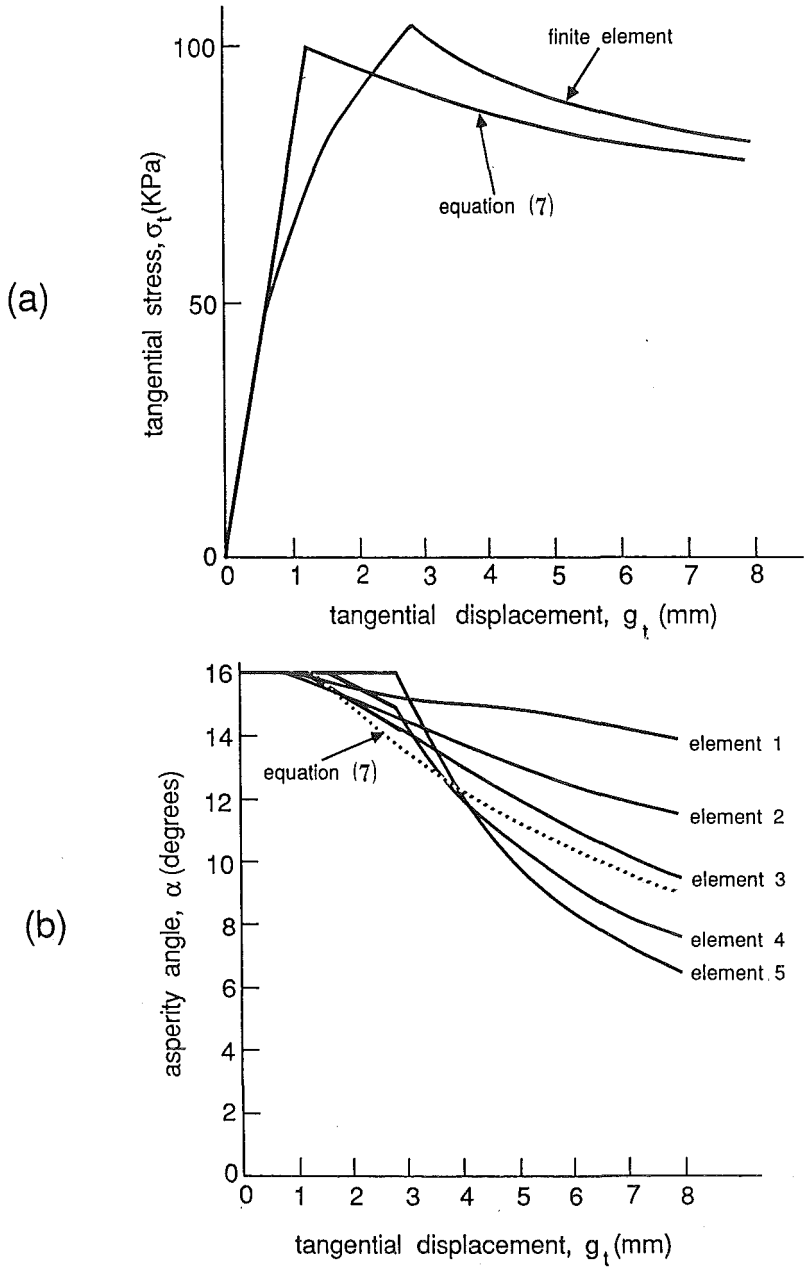

FIG. 5. Finite Element Results for Second Example as Function of Tangential Displacement: (*a*) Tangential Stress; and (*b*) Asperity Degradation Due to Damage; Also Shown Are Results Obtained by Directly Evaluating Eq. 7 with Constraint $\sigma_n = -90$ kPa

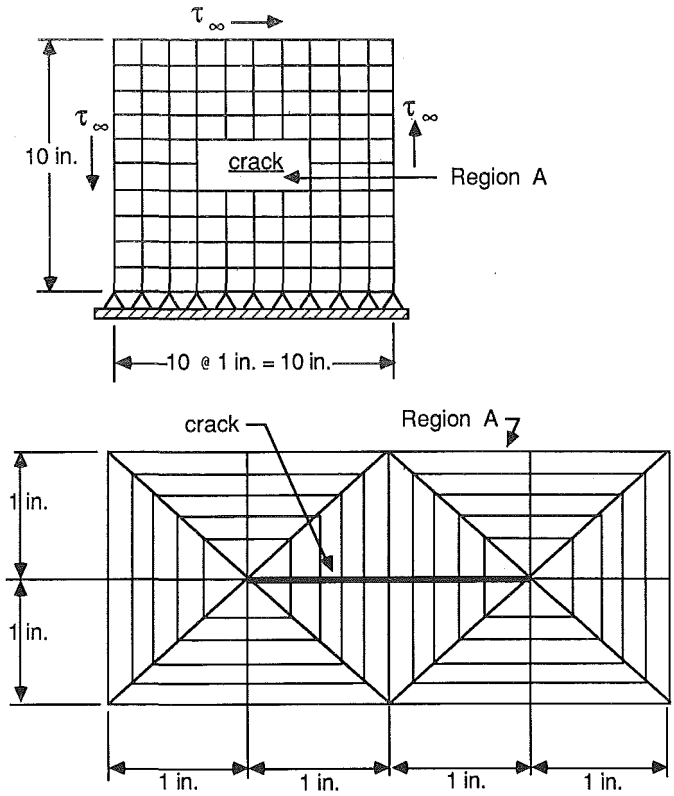

FIG. 6. Model and Mesh for Third Example

(i.e., no rotation of the upper plate of material) over the contact length.

The results of this simulation are in good quantitative agreement with those obtained from the experiments conducted by Bandis et al. (1981) and many others. Plesha (1987) provides a detailed comparison of the model and the experiment.

Example 3: Rough and Tortuous Crack in a Shear Field

This last example considers the plane-strain problem of the rough and tortuous crack shown in Fig. 6 subjected to far-field shear stresses. The finite element model of the elastic medium consists of six-node triangular (quadratic displacements) isoparametric elements surrounding the crack tip, and quadrilateral elements away from the crack tip. In this simulation the midside nodes of the triangular elements at the crack tip were moved to the quarter point to simulate the square root singularity in stresses at the crack tip. For the interface finite element closest to each crack tip, the midpoint nodes [i.e., generic nodes 2 and 5 in Fig. 2(*a*)] were also moved to the quarter point. This causes the Jacobian, Eq. 18, to vanish at the crack tip. However, this is of little consequence since the relative surface deformations, and hence the contact stresses, also vanish at the crack tip independently of the Jaco-

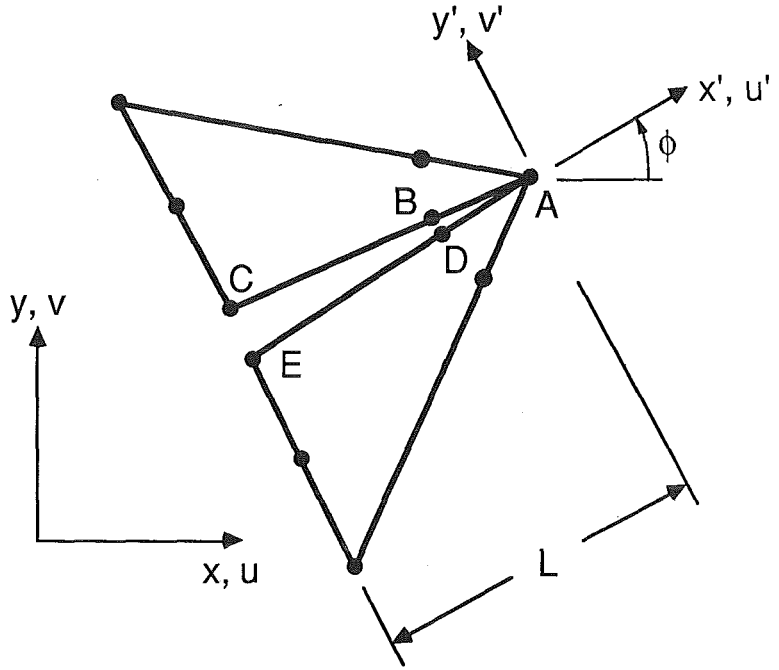

FIG. 7. Geometry of Displacements near Crack Tip for Computing Stress Intensity Factors (Shih et al. 1976)

bian. In fact, regardless of the criteria that determine the onset of sliding, which tends to ameliorate high-contact stresses, the interface element does not possess any traction singularities, regardless of the location of the midpoint nodes. Thus, the stress singularity at the crack tip is modeled exclusively by the quarter-point continuum elements. The stress-intensity factors K_{I} and K_{II} that characterize the stress and displacement fields in the vicinity of the crack tip are related to the displacements of element nodes for plane strain by

$$K_{\text{I}} = \sqrt{\frac{2\pi}{L}} \frac{E}{8(1-\nu^2)} [4(v'_B - v'_D) + v'_E - v'_C] \quad \ldots\ldots\ldots\ldots\ldots\ldots (33a)$$

$$K_{\text{II}} = \sqrt{\frac{2\pi}{L}} \frac{E}{8(1-\nu^2)} [4(u'_B - u'_D) + u'_E - u'_C] \quad \ldots\ldots\ldots\ldots\ldots\ldots (33b)$$

where L = the length of the singularity element along the direction of the crack surface; and the primes indicate that the global-coordinate nodal displacements have been transformed to the crack-tip coordinate system shown in Fig. 7 (Shih et al. 1976).

In the numerical calculations the physical parameters used for the intact material were $E = 30 \times 10^6$ psi; $\nu = 0.3$; and for the crack surfaces, $E_{tt} = E_{nn} = 1 \times 10^9$ psi/in.; $\mu = 0.2$; and $c = 0$. Fig. 8 shows the stress-intensity factors as functions of far-field shear stress for several values of asperity

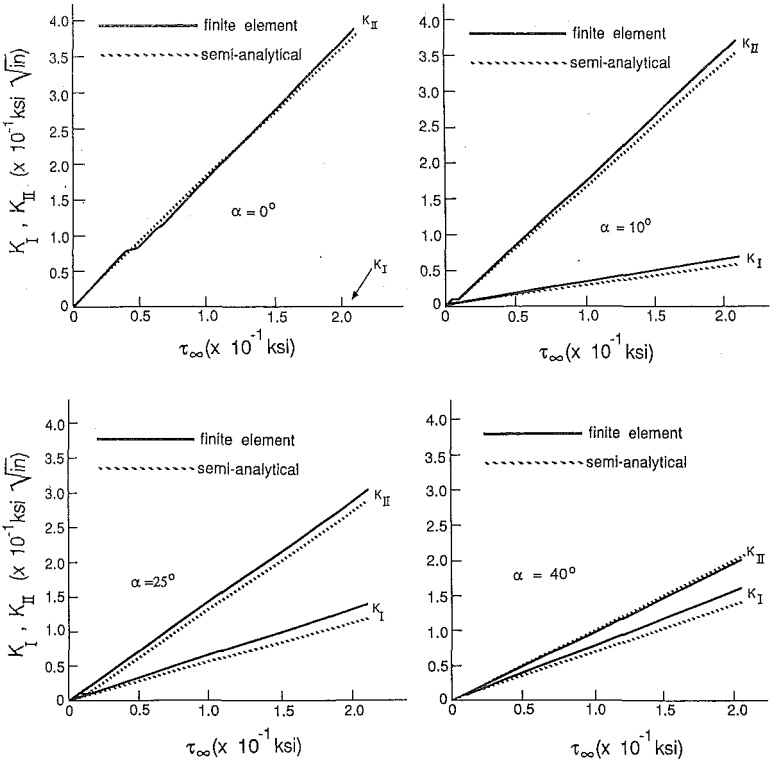

FIG. 8. Finite Element Results for Third Example; Stress Intensities versus Far-Field Shear Stress for Various Values of Asperity Orientation, α; Also Shown Are Results Obtained Semianalytically in Ballarini and Plesha (1987)

angle. For $\alpha = 0$, dilation is not present and hence the loading is strictly mode II ($K_I = 0$). As α increases, the resistance to tangential sliding increases thus reducing K_{II}, and because of dilatancy, K_I becomes nonzero.

Also shown in these figures are the results for the same problem but solved by a semianalytical method in Ballarini and Plesha (1987) using the constitutive law reported here in conjunction with nonlinear singular integral equations. The agreement is quite good.

Conclusions

Presentation of a comprehensive constitutive law and a suitable finite element implementation for dilatant contact-friction problems has been presented. Surface deformation in the normal and tangential directions are decomposed into elastic and plastic components. The resultant normal and tangential stresses are determined from the deformations through a general elastoplastic-type material matrix that accounts for several important physical features of the contact problem, namely surface roughness, relative mating of the contact surfaces, and an appropriate surface-friction law. Implemen-

tation in finite-element-analysis software conforms with the numerical procedures and program architecture that is commonly used for general material-nonlinear problems.

Acknowledgments

This research was partially supported by the U.S. Army Research Office (grant number DALL03-86-K-0134 to the University of Wisconsin) and by a Research Initiation Grant from the Ohio Board of Regents. The assistance of J. Hsu for investigating interface-element quadrature rules is gratefully acknowledged.

Appendix I. References

Ballarini, R., and Plesha, M. E. (1987). "The effects of crack surface friction and roughness on crack tip stress fields." *Int. J. Fracture,* 34(3), 195–207.

Bandis, S., Lumsden, A. C., and Barton, N. R. (1981). "Experimental studies of scale effects on the shear behavior of rock joints," *Int. J. Rock Mech., Mineral Sci. and Geomech. Abstracts,* 18(1), 1–21.

Bathe, K. J. (1982). *Finite element procedures in engineering analysis.* Prentice-Hall, Englewood Cliffs, N.J.

Bažant, Z. P., and Gambarova, P. (1980). "Rough cracks in reinforced concrete." *J. Struct. Div.,* ASCE, 106(4), 819–842.

Bažant, Z. P., and Gambarova, P. (1984). "Crack shear in concrete: Crack band microplane model." *J. Struct. Engrg.,* ASCE, 110(9), 2015–2035.

Belytschko, T., Chiapetta, R. L., and Bartel, H. D. (1976). "Efficient large scale non-linear transient analysis by finite elements." *Int. J. Numerical Methods in Engrg.,* 10(3), 579–596.

Belytschko, T. (1983). "An overview of semidiscretization and time integration procedures." *Computational methods for transient analysis,* T. Belytschko and T. J. R. Hughes, eds., North-Holland Publishers, Amsterdam, The Netherlands, 1–65.

Cheng, J. H., and Kikuchi, N. (1985). "An incremental constitutive relation of unilateral contact friction for large deformation analysis." *J. Appl. Mech.,* 52(3), 639–648.

Curnier, A. (1984). "A theory of friction." *Int. J. Solids and Struct.,* 20(7), 637–647.

Divakar, M. P., Fafitis, A., and Shah, S. P. (1987). "A constitutive model for shear transfer in cracked concrete." *J. Struct. Engrg.,* ASCE, 113(5), 1046–1062.

Drucker, D. C. (1954). "Coulomb friction, plasticity and limit loads." *J. Appl. Mech.,* 21, 71–74.

Fredriksson, B. (1976). "Finite element solution of surface nonlinearities in structural mechanics with special emphasis to contact and fracture mechanics problems." *Computers and struct.,* 6(4/5), 281–290.

Ghaboussi, J., Wilson, E. L., and Isenberg, J. (1973). "Finite element for rock joints and interfaces." *J. Soil Mech. and Found. Engrg. Div.,* ASCE, 99(10), 833–848.

Goodman, R. E., Taylor, R. L., and Brekke, T. L. (1968). "A model for the mechanics of jointed rock." *J. Soil Mech. and Found. Engrg. Div.,* ASCE, 14(3), 637–659.

Goodman, R. E. (1976). *Methods of geological engineering in discontinuous rocks.* West Publishing Co., St. Paul, Minn.

Goodman, R. E. (1980). *Introduction to rock mechanics.* John Wiley and Sons, Inc., New York, N.Y.

Kikuchi, N. (1982). "A smoothing technique for reduced integration penalty methods in contact problems." *Int. J. Numerical Methods in Engrg.,* 18(3), 343–350.

Kikuchi, N., and Song, Y. J. (1981). "Penalty/finite element approximations of a class of unilateral problems in linear elasticity." *Quar. Appl. Math.,* 39(1), 1–22.

Michalowski, R., and Mroz, Z. (1978). "Associated and non-associated sliding rules

in contact friction problems." *Archives of Mech.*, 30(3), 259–276.
Oden, J. T., and Campos, L. (1981). "Some new results on finite element methods for contact problems with friction." *New concepts in finite element analysis*, 44, T. J. R. Hughes et al., eds., Am. Soc. of Mech. Engrs., New York, N.Y., 1–9.
Oden, J. T., and Kikuchi, N. (1982). "Finite element methods for constrained problems in elasticity." *Int. J. Numerical Methods in Engrg.*, 18(5), 701–725.
Ortiz, M., and Popov, E. P. (1985). "Accuracy and stability of integration algorithms for elastoplastic constitutive relations." *Int. J. for Numerical Methods in Engrg.*, 21(9), 1561–1576.
Pande, G. N., and Pietruszczak, S. (1986). "Symmetric tangential stiffness formulation for non-associated plasticity." *Computers and Geotech.*, 2(2), 89–99.
Plesha, M. E. (1987). "Constitutive models for rock discontinuities with dilatancy and surface degradation."*Int. J. Numerical and Analytical Methods in Geomech.*, 11(4), 345–362.
Plesha, M. E., and Haimson, B. C. (1988). "An advanced model for rock joint behavior: Analytical, experimental and implementational considerations." *Proc., 29th U.S. Symp. on Rock Mechs.*, P. A. Cundall et al. eds., Univ. of Minnesota, 119–126.
Riggs, H. R., and Powell, G. H. (1986). "Rough crack model for analysis of concrete." *J. Engrg. Mech.*, ASCE, 112(5), 448–464.
Seguchi, Y., et al. (1974). "Sliding rule of friction in plastic forming of metal." *Computational Methods in Nonlinear Mech.*, Univ. of Texas–Austin, Austin, Tex., 683–692.
Shih, C. F., de Lorenzi, H. G., and German, M. D. (1976). "Crack extension modeling with singular quadratic isoparametric elements." *Int. J. of Fracture*, 12(4), 647–651.
Zubelewicz, A., et al. (1987). "A constitutive model for the cyclic behavior of dilatant rock joints." *Proc. 2nd Int. Conf. on Constitutive Laws for Engrg. Mater.*, C. S. Desai et al., eds., Univ. of Arizona, 1137–1144.

APPENDIX II. NOTATION

The following symbols are used in this paper:

$\mathbf{B}$ = matrix relating interface displacements to nodal displacements;
c = surface degradation parameter;
$\mathbf{D}$ = vector of nodal displacements;
E = Young's modulus;
$E_{ij}, \mathbf{E}$ = interface elastic stiffnesses;
$E^{ep}_{ij}, \mathbf{E}^{ep}$ = elastic-plastic interface stiffnesses;
F = slip function;
$\mathbf{F}$ = vector of nodal forces;
G = slip potential;
$g_t, g_n, \mathbf{g}$ = relative displacements;
H = hardening/softening parameter;
$\mathbf{I}$ = identity matrix;
J = Jacobian;
$\mathbf{K}^{ep}$ = elastic-plastic tangent stiffness matrix;
$K_{\mathrm{I}}, K_{\mathrm{II}}$ = mode I and mode II stress intensity factors;
l = physical length of interface finite element;
$\mathbf{M}, \mathbf{N}$ = matrices of shape functions;
s = slope of finite element;
$\mathbf{T}$ = transformation matrix;
$\vec{\mathbf{t}}, \vec{\mathbf{n}}$ = vectors in tangent and normal directions of finite element;

$u_x, u_y, \mathbf{u}$ = displacements;
W^p = sliding work;
$\mathbf{X}, \mathbf{Y}$ = vectors containing finite element nodal coordinates;
α_k = orientations of asperity surfaces;
θ = orientation of finite element;
$\dot{\lambda}$ = slip multiplier;
μ = friction coefficient;
ν = Poisson's ratio;
ξ = reference coordinate for finite element; and
$\sigma_t, \sigma_n, \boldsymbol{\sigma}$ = surface stresses.