Available online at www.sciencedirect.com
SCIENCE DIRECT®
Comput. Methods Appl. Mech. Engrg. 193 (2004) 2667–2698
Computer methods in applied mechanics and engineering
www.elsevier.com/locate/cma

Computational modeling of deformation bands in granular media. I. Geological and mathematical framework

Ronaldo I. Borja [a,*,1], Atilla Aydin [b,2]

[a] Department of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305-4020, USA
[b] Department of Geological and Environmental Sciences, Stanford University, Stanford, CA 94305-2115, USA

Received 9 May 2003; received in revised form 12 September 2003; accepted 14 September 2003

Abstract

Failure of granular media under natural and laboratory loading conditions involves a variety of micromechanical processes producing several geometrically, kinematically, and texturally distinct types of structures. This paper provides a geological framework for failure processes as well as a mathematical model to analyze these processes. Of particular interest is the formation of tabular deformation bands in granular rocks, which could exhibit distinct localized deformation features including simple shearing, pure compaction/dilation, and various possible combinations thereof. The analysis is carried out using classical bifurcation theory combined with non-linear continuum mechanics and theoretical/computational plasticity. For granular media, yielding and plastic flow are known to be influenced by all three stress invariants, and thus we formulate a family of three-invariant plasticity models with a compression cap to capture the entire spectrum of yielding of geomaterials. We then utilize a return mapping algorithm in principal stress directions to integrate the stresses over discrete load increments, allowing the solution to find the critical bifurcation point for a given loading path. The formulation covers both the infinitesimal and finite deformation regimes, and comparisons are made of the localization criteria in the two regimes. In the accompanying paper, we demonstrate with numerical examples the role that the constitutive model and finite deformation effects play on the prediction of the onset of deformation bands in geomaterials.

Keywords: Deformation bands; Granular media

1. Introduction

Failure in geomaterials such as concrete, soils, and rocks are often accompanied by the appearance of narrow tabular bands of intense deformation. The most common mode involves a shear offset combined

* Corresponding author. Fax: +1-650-723-7514.
E-mail address: borja@stanford.edu (R.I. Borja).
[1] Supported by NSF grant nos. CMS-97-00426 and CMS-03-24674; and DOE grant no. DE-FG02-03ER15454.
[2] Supported by NSF grant no. EAR-02-29862 and DOE grant no. DE-FG03-94ER14462.

0045-7825/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.cma.2003.09.019

with either compaction or dilation. However, basic deformation modes in rocks involving simple shearing with no significant volumetric increase or decrease, or pure compaction/dilation with no significant shear offset, also have been observed either in the laboratory or in the field [1–18]. It appears that bands with compaction and grain fracturing tend to undergo the largest volumetric deformation in natural settings.

Deformation bands in granular media are often interpreted as resulting from material instability influenced largely by existing defects or imperfections. Unfortunately, such imperfections are difficult if not impossible to quantify, and thus the occurrence of deformation bands are often analyzed as a bifurcation of the macroscopic inelastic constitutive behavior. The idea is derived from the works of Hadamard [19], Hill [20], Thomas [21] and Mandel [22] within the context of acceleration waves in solids, and involves investigation of the occurrence of alternate kinematical solutions (such as the emergence of tabular deformation bands) satisfying the governing field equations. Conditions for the onset of such bands in geomaterials have been presented in the context of elastoplasticity by Rudnicki and Rice [23–25], and in the context of hypoplasticity by a number of European schools [26–30].

Although the bifurcation theory for the analysis of deformation bands in geomaterials is fairly well understood, it appears that much application has focused in the past on the shear localization of pressure-sensitive dilatant materials [23–25]. Quite recently, the same theory has been used to model the occurrence of compaction bands [13,15,16,31], a certain type of geologic structure observed to form in porous rocks. Such type of structure could have important geologic implications since they represent fluid barriers and therefore the prediction of their occurrence is of scientific and engineering value. Finally, recent literature also indicates the occurrence in the field of so-called dilation bands [7], characterized by a predominantly opening mode. Such geologic structure provides a sharp contrast to planar opening-mode fractures or joints with two discrete surfaces, since dilation bands do not result in free surfaces. The latter mode completes the spectrum of observed localized deformation modes and provides strong motivations for the development of a unifying geologic and mathematical framework for characterizing these modes.

Our geologic framework for localized deformation in tabular bands is based on the relative contributions of shear and volumetric deformations. The three extreme modes are pure compaction, pure dilation, and simple shear; combination modes involve shearing with either compaction or dilation. Section 2 presents a formal classification of these failure modes as well as describes their geological characteristics. For simplicity we shall limit the scope of this paper to localized deformation modes in granular rocks.

Our point of departure for the mathematical characterization of localized deformation modes in tabular bands is the balance of traction across a surface where the displacement gradient field may be discontinuous. Theory of plasticity is used to characterize the inelastic constitutive response. The formulation results in a homogeneous system of equations, and conditions are sought for a non-trivial solution. This yields the critical band orientation at which the eigenvalue problem is first satisfied. The above procedure is fairly standard [32]; however, a by-product of the analysis that has not been fully exploited in the literature concerns the fact that the eigenvalue problem also produces a characteristic vector defining the direction of the relative velocity jump across the thickness of the band. The theory clearly defines not only the orientation of this characteristic vector but also its sense (i.e., direction). Together with the previously determined critical band orientation, the accompanying eigenvector predicts the nature of the resulting deformation bands.

We show that the mathematical framework is robust in that it covers the entire spectrum of localized deformation modes established in the geologic framework, including the extreme cases of pure compaction/dilation and the simple shear band localization modes. For pure compaction/dilation bands some theoretical analyses have been advanced fairly recently in the literature [13,16,31] indicating that the theoretical orientations of these bands coincide with those of the principal stress planes. In this paper we qualify this conclusion and restrict its validity to the case of coaxial plastic flow theory in which the principal directions of the stress tensor are assumed to coincide with those of the plastic strain increment tensor. Strictly, we show that compaction/dilation bands are theoretically parallel to a principal plane of the plastic strain

increment tensor, and not necessarily to a principal plane of the stress tensor. This distinction is significant particularly when dealing with non-coaxial plastic flow deformation.

Prediction of tabular bands as a bifurcation from a homogeneous deformation field is well understood. The results are known to be strongly dependent on the constitutive description of the homogeneous deformation. However, to date almost all of the modeling efforts have focused on two-invariant constitutive representations of the homogeneous deformation, which may not be adequate for cohesive-frictional materials such as granular rocks. In the first place, these materials exhibit lower yield and failure strengths in tension than in compression, suggesting some influence of the third stress invariant on the yield and plastic flow behavior. Indeed, evidence from numerous laboratory tests suggest the significant effect of the third stress invariant on the description of the mechanical responses of geomaterials [33–44].

Detection of the bifurcation point requires a robust numerical integration procedure for the elastoplastic constitutive relations. In computational plasticity the return mapping algorithm offers distinct advantages over the traditional explicit schemes, including simplicity in the implementation, compatibility with the structure of many existing finite element codes, and facility for extension to the finite deformation regime. Only very recently, the return mapping algorithm also has been applied to three-invariant plasticity models delivering optimal performance [45–48]. The next step then would be to use this powerful algorithm for the more challenging task of accurately capturing the inelastic loading history leading to different failure modes in granular materials.

Notations and symbols used in this paper are as follows: bold-face letters denote matrices and vectors; the symbol '$\cdot$' denotes an inner product of two vectors (e.g. $\boldsymbol{a}\cdot\boldsymbol{b}=a_ib_i$), or a single contraction of adjacent indices of two tensors (e.g. $\boldsymbol{c}\cdot\boldsymbol{d}=c_{ij}d_{jk}$); the symbol ':' denotes an inner product of two second-order tensors (e.g. $\boldsymbol{c}:\boldsymbol{d}=c_{ij}d_{ij}$), or a double contraction of adjacent indices of tensors of rank two and higher (e.g. $\boldsymbol{C}:\boldsymbol{\epsilon}^{\mathrm{e}}=C_{ijkl}\epsilon^{\mathrm{e}}_{kl}$); the symbol '$\otimes$' denotes a juxtaposition, e.g. $(\boldsymbol{a}\otimes\boldsymbol{b})_{ij}=a_ib_j$. For any symmetric second-order tensors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, we have $(\boldsymbol{\alpha}\otimes\boldsymbol{\beta})_{ijkl}=\alpha_{ij}\beta_{kl}$; $(\boldsymbol{\alpha}\oplus\boldsymbol{\beta})_{ijkl}=\alpha_{jl}\beta_{ik}$; and $(\boldsymbol{\alpha}\ominus\boldsymbol{\beta})_{ijkl}=\alpha_{il}\beta_{jk}$.

2. Classification of failure modes and their geological characteristics

We consider the entire spectrum of localized deformation in tabular bands resulting from distinct failure modes in granular rocks. The two top tiers in this spectrum are shear and volumetric deformation bands:

(1) Shear deformation bands
  (1.1) Pure shear bands
  (1.2) Compactive shear bands
  (1.3) Dilatant shear bands
(2) Volumetric deformation bands
  (2.1) Pure compaction bands
  (2.2) Pure dilation bands

Shear deformation bands, which will be referred to as "shear bands" for simplicity in this paper, are dominated by a component of velocity gradient parallel to the tabular band boundaries (Fig. 1a). This is recognizable by a clear evidence of macroscopic shear offset across the band in rocks (Fig. 2a). Similar to the classification of fractures in fracture mechanics, shear bands may have two basic modes, sliding (mode-II) and tearing (mode-III), based on the geometry and kinematics of their general propagation direction. However, this detail is not of further concern in this paper.

Although it is rare, a shear band without any volumetric deformation within the band (top flat line in Fig. 1b) may occur in nature [2]. This type of shear band will be referred to as pure, or simple shear band similar to the "simple shear" in mechanics. Most shear bands, however, undergo volumetric deformation in

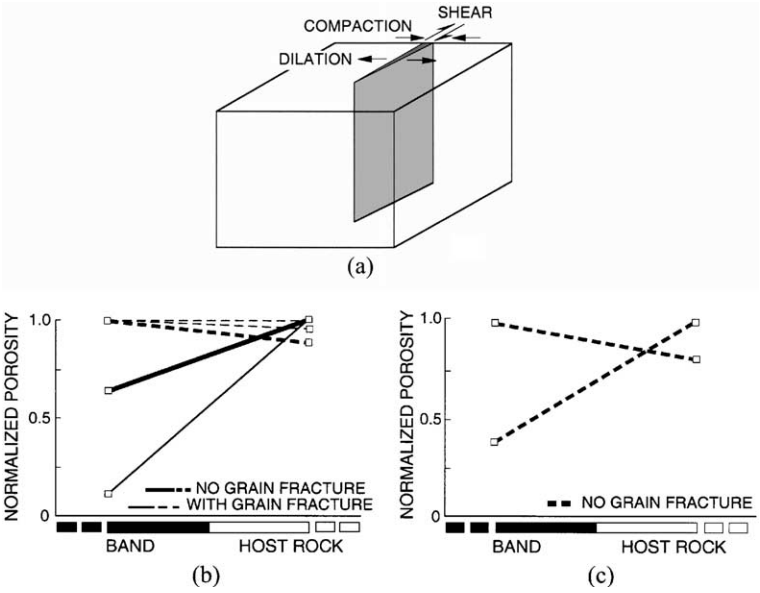

Fig. 1. Failure modes in granular rocks: (a) an idealized tabular band of localized deformation showing shear, compaction, and dilation modes; (b) based on a large amount of data in the literature, shear bands show three fundamental trends in terms of the component of shear and volumetric deformation: simple shear bands have no significant volume change (the flat line defined by 2 measurements); compactive and dilatant shear bands showing mixed-mode shear bands with compaction (18 samples) and dilation (6 samples), respectively; (c) porosity data showing pure compaction (2 samples) and pure dilation bands (4 samples from the same locality). Data from Refs. [1–4,7,15].

addition to shearing [1–4]. These mixed-mode bands, regardless of the relative magnitude of shear and volumetric components of deformation, will be referred to as "compactive shear bands" and "dilatant shear bands." Compactive and dilatant shear bands are associated with volume decrease and increase, respectively (Fig. 1b). It is, therefore, essential to characterize the volume change within tabular bands. Several methods were employed to determine pore or skeleton volume and porosity within bands and in relatively pristine rocks nearby. These methods include liquid/helium porosimetry or immersion samples into a liquid [4], petrographic image analysis [2] and X-ray computerized tomography—CT scanner [3], and point counting [7].

Volumetric deformation bands lack shearing across the bands and have two diametrically opposite modes: compaction bands and dilation bands. Compaction bands are tabular bands where the boundaries move toward each other (Fig. 1a) producing a volume decrease (Fig. 1c); dilation bands are those where the boundaries move away from each other (Fig. 1a) resulting in a volume increase (Fig. 1c). It is conceivable that a band may have a greater volumetric deformation component with respect to the shear component. However, it is not practical to distinguish between the volumetric deformation bands with shear components and the shear bands with volumetric deformation components (1.2 and 1.3 under the classification scheme) in most naturally formed deformation bands. Therefore, bands with shear and volumetric deformation components are not differentiated based on the relative magnitudes of these components except for extreme cases in which one of the components is or close to zero.

2.1. Shear bands

Tabular bands of shear in granular rocks were described in the geological literature [3,10]. The diagnostic character of shear bands is a macroscopic shear offset (slip) across them, which can be measured

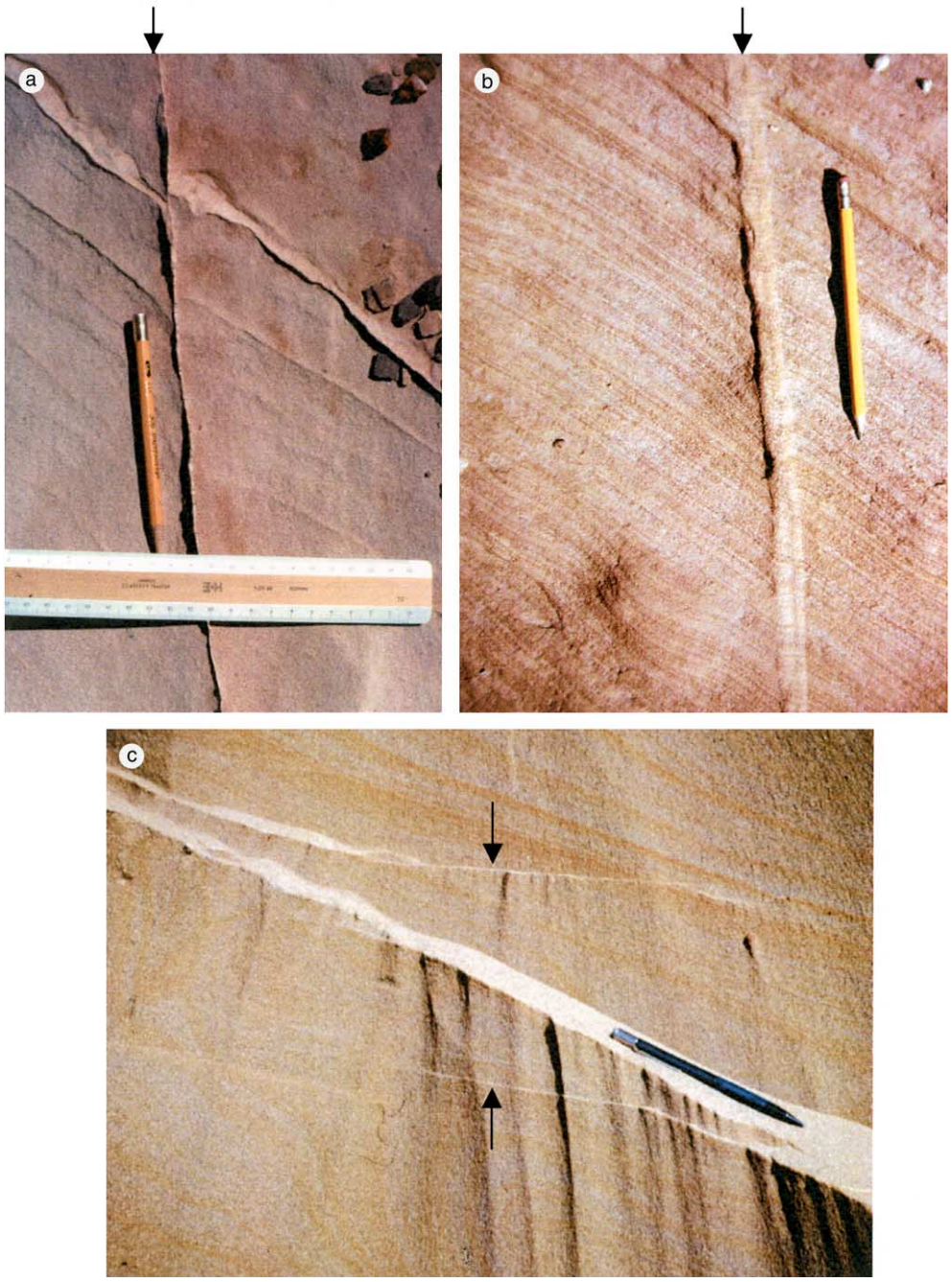

Fig. 2. (a) A shear band (marked by arrow) in the Entrada Sandstone, San Rafael Desert, Utah. The band is about 1–2 mm in thickness and displaces nearly horizontal beds by a few mm (left-hand side down). Ruler is about 20 cm long. (b) A compaction band in the Aztec Sandstone at Valley of Fire, Nevada. The band is about one cm thick, has no observable shear offset across it, and shows significant porosity decrease. Standard size pencil is for scale. (c) Dilation bands (nearly horizontal traces marked by arrows) in unconsolidated terrace sand near McKinleyville, Northern California. Pencil on the surface of the shear band is for scale. Porosity distribution across a thin section covering a horizontal band indicates about 7% porosity increase within the band.

from previously continuous reference markers (depositional beds) cut across by shear bands (Fig. 2a). Typically, the maximum slip across single shear bands is about a few millimeters to a few centimeters and occurs either at or near the midpoint along the trace of the bands [11]. The thickness of individual shear bands observed in rocks deformed under natural forces (Fig. 2a) is also limited to a few millimeters [2–4,10,11]. Similarly, shear bands produced in rock samples deformed under laboratory conditions have similar thicknesses [8,14]. Using these values for shear band thickness and slip, typical average shear strain across shear bands was calculated to be on the order of unity. The consistent values for shear bands thicknesses and limited slip across them were attributed to grain size and strain hardening, respectively [4]. The length dimension of single shear bands is also limited to about one to 100 m [11] or a few hundred meters [4] at most. Thus, it is necessary to form new shear bands adjacent to existing ones in order to widen and lengthen a shear band structure and to accommodate a larger magnitude of slip [5].

Shear bands are commonly associated with grain fracturing and grain size reduction described by the term "cataclasis" in the geological literature. Thin sections of shear bands when viewed under a petrographic microscope show evidence for grain fracturing and other micromechanical processes responsible for their formation. In most cases, grain fracturing can be detected at an initial stage or in the periphery of a shear band where the grains are damaged but not yet demolished. In more advanced stages of shear band development with a high intensity cataclasis, grain fracturing and grain crushing are reflected by grain size reduction and change of grain shape from a rounded form outside the band to an angular form within the band. Typical grain size distribution within shear bands is such that the range of grain size broadens indicating a poorer sorting due to an increasing number of smaller grains induced by the comminution and survival of a few original grains in the bands.

Although a majority of shear bands reported in the geological literature are associated with grain size reduction resulting from grain fracturing described earlier, the process of grain fracturing or the related comminution is not a prerequisite for shear band formation. It is possible that grain movement by sliding along grain contacts and pore collapse may localize shear and volumetric strains into a tabular band with finite thickness especially under low confining pressure and with no or little cement [2,9].

Following the classification scheme presented earlier and the common micromechanical properties discussed above, it is essential to know about the volume change within tabular deformation bands. The plots in Fig. 1b summarize some trends with respect to the nature of volume change and the micromechanics of grain fracturing in many naturally occurring shear bands distilled from the literature.

Simple shear bands have no volumetric deformation component by definition (Fig. 1b). Compactive and dilatant shear bands have compaction and dilation, respectively, in addition to the shearing components. It appears that those bands with compaction and grain fracturing have undergone the largest volumetric deformation in natural settings.

2.2. Compaction and dilation bands

Two end members of volumetric deformation bands (Fig. 1c) are: (a) compaction bands characterized by a volume decrease, and (b) dilation bands characterized by a volume increase with respect to corresponding undeformed parent rock. Fig. 2b shows an isolated compaction band in the Aztec Sandstone in Valley of Fire State Park in southeastern Nevada. These were the earliest examples for pure compaction localization reported in the literature [2,3,6,12]. Compaction bands have later been reported from other locations with independent evidence for the compactive character of the deformation [15]. The data indicate a porosity decrease from an average of about 20% to 25% for the undeformed rock to about one to 5% for the band [2,3], and it is distributed along the bands in form similar to that from an idelized anticrack in accordance to linear elastic fracture mechanics theory [17].

Deformation bands with porosity increase with respect to the undeformed state of the rock have been reported from many locations [2,3]. However, the shearing component along these bands or the lack thereof

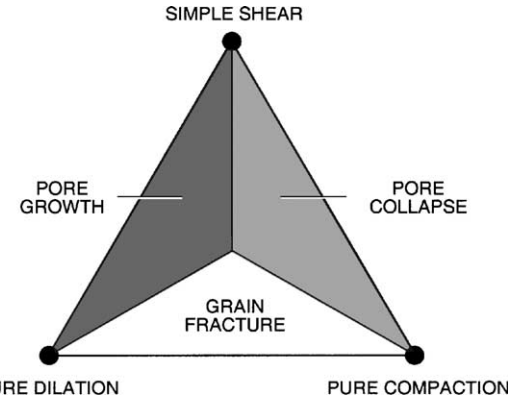

Fig. 3. Idealized diagram defining the failure mechanism and failure modes in porous rock. Note that shear/dilation/grain fracture and shear/compaction/grain fracture are permissible.

was unclear. The only unambiguous case in which shear offset has been ruled out and dilation has been independently affirmed is that reported by Du Bernard et al. [7]. Photograph in Fig. 2c shows an outcrop pattern of dilation bands (horizontal bands marked by arrows) from this study. The dilation bands linked to the segments of the associated shear band (inclined in the photograph marked by pencil). The linkage points of the dilation bands and the shear band segments appear to be marked by a sharp kink and the dilation bands occur at the dilational quadrants of a series of segmented shear bands based on the sense of slip across the shear band segments (right-hand side up). The graph in Fig. 1c representing porosity measurements from this locality shows about 7% porosity increase within the band.

Pure compaction bands have recently been produced in the laboratory [16,18], and both compaction and dilation bands have been analyzed theoretically by Issen and Rudnicki [13] and others as will be discussed later.

To summarize, we have presented a classification scheme that accounts for the entire spectrum of deformation field in the form of tabular bands. The end members are simple shear, pure compaction and pure dilation bands (Fig. 3). The mixed modes include shear with either volume decrease or volume increase. Field data indicate that although the three end members do occur in nature, mixed-mode localization structures, compactive shear bands and dilatant shear bands are the most common modes of localized failure. The volumetric deformation is the largest for compactive shear bands and compaction bands. The mathematical framework to analyze these failures modes, including the computer implementation of a proposed model, is next described in the following sections.

3. Formulation of deformation bands: infinitesimal case

In this section we revisit the general localization theory of deformation bands with the following main goals: (a) to demonstrate that the theory is complete in that it defines a necessary condition for the emergence of a deformation band, the likely orientation of this band, and the nature of the accompanying localized volumetric response; and (b) to show that the theory encompasses the extreme cases of pure compaction band and pure dilation band under certain constitutive hypotheses.

3.1. General analysis of deformation bands

The general kinematics of a deformation band in the infinitesimal regime is shown in Fig. 4. Here, $\Omega \subset R^{n_{sd}}$ represents the reference configuration of a body with smooth boundary $\partial\Omega$, and $\boldsymbol{x}$ represents the

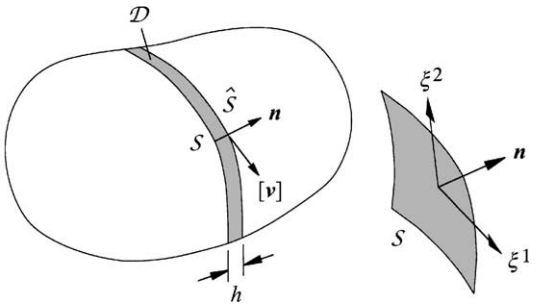

Fig. 4. Normal parameterization of shear band geometry: infinitesimal deformation case.

position vector of any particle in Ω. We consider a smooth material surface $\mathscr{S} \subset \Omega$ where some fields may be discontinuous. Adopting the notation of [49], we denote points in $\mathscr{S}$ by $\boldsymbol{y}$ so that

$$\mathscr{S} = \{\boldsymbol{y} = \widehat{\boldsymbol{y}}(\xi^1, \xi^2) | (\xi^1, \xi^2) \in \mathscr{B}\}, \tag{3.1}$$

where $\widehat{\boldsymbol{y}} : \mathscr{B} \to R^{n_{\mathrm{sd}}}$ is a smooth global parameterization. Thus, the unit normal to $\mathscr{S}$ is

$$\boldsymbol{n} = \widehat{\boldsymbol{n}}(\xi^1, \xi^2) = \widehat{\boldsymbol{y}}_{,1} \times \widehat{\boldsymbol{y}}_{,2} / \|\widehat{\boldsymbol{y}}_{,1} \times \widehat{\boldsymbol{y}}_{,2}\|. \tag{3.2}$$

The above parameterization for $\mathscr{S}$ provides a convenient normal parameterization in the closed tabular band domain $\overline{\mathscr{D}} = \mathscr{S} \times [0, h]$ so that any point $\widehat{\boldsymbol{x}}$ in the deformation band is defined by the mapping

$$\widehat{\boldsymbol{x}}(\xi^1, \xi^2, \eta) = \widehat{\boldsymbol{y}}(\xi^1, \xi^2) + \eta \widehat{\boldsymbol{n}}(\xi^1, \xi^2) \quad \text{for } 0 \leqslant \eta \leqslant h, \tag{3.3}$$

where $h > 0$ is the band thickness, herein assumed to be small but finite. A second smooth surface $\widehat{\mathscr{S}}$ can then be defined by the set relation

$$\widehat{\mathscr{S}} = \{\boldsymbol{y} = \widehat{\boldsymbol{x}}(\xi^1, \xi^2, h) | (\xi^1, \xi^2) \in \mathscr{B}\}, \tag{3.4}$$

so that $\mathscr{S}$ and $\widehat{\mathscr{S}}$ define opposite surfaces of discontinuity representing boundaries of the band domain.

We define the velocity field by the ramp-like relation

$$\boldsymbol{v} = \begin{cases} \bar{\boldsymbol{v}} & \text{if } \eta \leqslant 0; \\ \bar{\boldsymbol{v}} + \eta [\![\boldsymbol{v}]\!]/h & \text{if } 0 \leqslant \eta \leqslant h; \\ \bar{\boldsymbol{v}} + [\![\boldsymbol{v}]\!] & \text{if } \eta \geqslant h, \end{cases} \tag{3.5}$$

where $\bar{\boldsymbol{v}}$ is a continuous velocity field and $[\![\boldsymbol{v}]\!]$ represents the relative velocity of the opposite faces of the band. Assuming $[\![\boldsymbol{v}]\!]$ is uniform over $\hat{\mathscr{S}}$, the corresponding velocity gradient fields outside and inside the band take the form

$$\boldsymbol{l} = \begin{cases} \nabla \bar{\boldsymbol{v}} & \text{in } \Omega \setminus \overline{\mathscr{D}}; \\ \nabla \bar{\boldsymbol{v}} + ([\![\boldsymbol{v}]\!] \otimes \boldsymbol{n})/h & \text{in } \mathscr{D}, \end{cases} \tag{3.6}$$

where $\mathscr{D} = \mathscr{S} \times (0, h)$ is the open band domain and $\boldsymbol{n}$ is the unit normal vector to the band (since h is assumed small, $\boldsymbol{n}$ may be taken as normal to either $\mathscr{S}$ or $\hat{\mathscr{S}}$). We note that the orientation of $\boldsymbol{n}$ is perfectly symmetric in the sense that it may be directed either outward or inward to the band. The velocity gradient is thus discontinuous across the band, and upon evaluating just inside and just outside the surface of discontinuity we obtain the relations

$$\boldsymbol{l}^1 = \boldsymbol{l}^0 + \frac{1}{h} [\![\boldsymbol{v}]\!] \otimes \boldsymbol{n} \Longleftrightarrow \dot{\boldsymbol{\epsilon}}^1 = \dot{\boldsymbol{\epsilon}}^0 + \frac{1}{h} \mathrm{sym}([\![\boldsymbol{v}]\!] \otimes \boldsymbol{n}), \tag{3.7}$$

where $\dot{\boldsymbol{\epsilon}}^1 = \mathrm{sym}(\boldsymbol{l}^1)$ and $\dot{\boldsymbol{\epsilon}}^0 = \mathrm{sym}(\boldsymbol{l}^0)$. Throughout this paper we will use the superscript symbols "1" and "0" to refer to points on $\mathscr{S}$ interpreted to lie just inside and just outside this surface, respectively.

We assume an elastoplastic material with a yield function F and a plastic potential function Q, and denote their gradients with respect to the Cauchy stress tensor as

$$\boldsymbol{f} = \frac{\partial F}{\partial \boldsymbol{\sigma}}, \qquad \boldsymbol{q} = \frac{\partial Q}{\partial \boldsymbol{\sigma}}. \tag{3.8}$$

Further, we assume that at the moment of localization we have the inequalities

$$\boldsymbol{f} : \boldsymbol{c}^{\mathrm{e}} : \dot{\boldsymbol{\epsilon}}^1 > 0, \qquad \boldsymbol{f} : \boldsymbol{c}^{\mathrm{e}} : \dot{\boldsymbol{\epsilon}}^0 > 0, \tag{3.9}$$

where $\boldsymbol{c}^{\mathrm{e}}$ is the fourth-order tensor of elastic moduli. The above conditions imply that the material is yielding plastically on both sides of $\mathscr{S}$. The case with loading on one side and unloading on the other side of $\mathscr{S}$ has been investigated in [25], where it was demonstrated that this bifurcation mode is less critical than the case where plastic yielding occurs on both sides.

Just inside the surface $\mathscr{S}$ the rate constitutive equation takes the form

$$\dot{\boldsymbol{\sigma}}^1 = \boldsymbol{c}^{\mathrm{e}} : (\dot{\boldsymbol{\epsilon}}^1 - \dot{\lambda}\boldsymbol{q}), \tag{3.10}$$

where $\dot{\lambda}\boldsymbol{q} := \dot{\boldsymbol{\epsilon}}^{\mathrm{p}}$ is the plastic component of the strain rate $\dot{\boldsymbol{\epsilon}}$ (from the flow rule). We recall that the non-negative plastic multiplier $\dot{\lambda}$ satisfies the Kuhn–Tucker complementarity condition [50]

$$\dot{\lambda} \geqslant 0, \quad F \leqslant 0, \quad \dot{\lambda}F = 0. \tag{3.11}$$

Furthermore, with isotropy in the elastic response the elasticity tensor $\boldsymbol{c}^{\mathrm{e}}$ takes the form

$$\boldsymbol{c}^{\mathrm{e}} = K\boldsymbol{1} \otimes \boldsymbol{1} + 2\mu(\boldsymbol{I} - \tfrac{1}{3}\boldsymbol{1} \otimes \boldsymbol{1}), \tag{3.12}$$

where K and μ are the elastic bulk and shear moduli, and $\boldsymbol{I}$ is the symmetric fourth-order identity tensor with components $\boldsymbol{I}_{ijkl} := (\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})/2$.

The plastic multiplier $\dot{\lambda}$ may be determined from the consistency condition

$$\dot{\mathscr{F}} = \boldsymbol{f} : \dot{\boldsymbol{\sigma}}^1 - \dot{\lambda}H = \boldsymbol{f} : \boldsymbol{c}^{\mathrm{e}} : (\dot{\boldsymbol{\epsilon}}^1 - \dot{\lambda}\boldsymbol{q}) - \dot{\lambda}H = 0, \tag{3.13}$$

where H is the plastic modulus. Solving this last equation for $\dot{\lambda}$ gives

$$\dot{\lambda} = \dot{\bar{\lambda}} + \dot{\bar{\bar{\lambda}}}/h, \tag{3.14}$$

where

$$\dot{\bar{\lambda}} = \frac{1}{\chi}\boldsymbol{f} : \boldsymbol{c}^{\mathrm{e}} : \dot{\boldsymbol{\epsilon}}^0 > 0, \qquad \dot{\bar{\bar{\lambda}}} = \frac{1}{\chi}\boldsymbol{f} : \boldsymbol{c}^{\mathrm{e}} : \mathrm{sym}(\llbracket \boldsymbol{v} \rrbracket \otimes \boldsymbol{n}) > 0, \tag{3.15}$$

and

$$\chi = \overline{\chi} + H > 0, \quad \overline{\chi} = \boldsymbol{f} : \boldsymbol{c}^{\mathrm{e}} : \boldsymbol{q} > 0. \tag{3.16}$$

The restriction that χ be a non-negative function ensures that $\dot{\bar{\lambda}} > 0$, since the scalar product $\boldsymbol{f} : \boldsymbol{c}^{\mathrm{e}} : \dot{\boldsymbol{\epsilon}}^0$ is non-negative by assumption (3.9). The additional restriction that $\overline{\chi}$ be a non-negative function rules out an extreme non-associative plastic flow which could lead to 'objectionable' mechanical responses, see [25]. Finally, from the condition that $\dot{\bar{\bar{\lambda}}} > 0$ it follows from (3.15) that

$$\boldsymbol{f} : \boldsymbol{c}^{\mathrm{e}} : \mathrm{sym}(\llbracket \boldsymbol{v} \rrbracket \otimes \boldsymbol{n}) > 0. \tag{3.17}$$

This last inequality may be used as a criterion to determine whether the deformation band would exhibit a dilatant or compactive behavior at localization.

Substituting (3.14) into Eq. (3.10) leads to a rate constitutive relation just inside the surface of discontinuity $\mathscr{S}$ of the form

$$\dot{\boldsymbol{\sigma}}^1 = \boldsymbol{c}^{\mathrm{ep}} : \dot{\boldsymbol{\epsilon}}^1, \tag{3.18}$$

where $\boldsymbol{c}^{\mathrm{ep}}$ is the elastoplastic constitutive operator given by

$$\boldsymbol{c}^{\mathrm{ep}} = \boldsymbol{c}^{\mathrm{e}} - \frac{1}{\chi}\boldsymbol{c}^{\mathrm{e}} : \boldsymbol{q} \otimes \boldsymbol{f} : \boldsymbol{c}^{\mathrm{e}}. \tag{3.19}$$

From the hypothesis that yielding takes place on both sides of the band, a similar rate constitute equation may be written just outside the band as

$$\dot{\boldsymbol{\sigma}}^0 = \boldsymbol{c}^{\mathrm{ep}} : \dot{\boldsymbol{\epsilon}}^0. \tag{3.20}$$

Following standard arguments, a deformation band is then possible provided the traction rate vector across the surface of discontinuity is continuous,

$$\boldsymbol{n} \cdot \dot{\boldsymbol{\sigma}}^0 = \boldsymbol{n} \cdot \dot{\boldsymbol{\sigma}}^1. \tag{3.21}$$

Writing $[\![\boldsymbol{v}]\!] = \dot{\zeta}\boldsymbol{m}$, where $\dot{\zeta} > 0$ is the magnitude and $\boldsymbol{m}$ is the unit direction of $[\![\boldsymbol{v}]\!]$, the traction continuity condition becomes

$$(\dot{\zeta}/h)\boldsymbol{A} \cdot \boldsymbol{m} = \boldsymbol{0}, \quad \boldsymbol{A} = \boldsymbol{n} \cdot \boldsymbol{c}^{\mathrm{ep}} \cdot \boldsymbol{n}, \tag{3.22}$$

where $\boldsymbol{A}$ is the elastoplastic acoustic tensor.

For a non-trivial solution to exist, standard argument requires that

$$\det(\boldsymbol{A}) = 0. \tag{3.23}$$

The onset of a deformation band corresponds to the initial satisfaction of this determinant condition for some critical band orientation $\boldsymbol{n}$. Because of the homogeneous form of the localization condition it is not possible to solve for the strain rate $\dot{\zeta}/h$ even if h is known. However, we can always solve for the unit characteristic vector $\boldsymbol{m}$ of the tensor $\boldsymbol{A}$ once a critical band orientation $\boldsymbol{n}$ has been identified. Because the characteristic equation has a homogeneous form, two eigenvectors are possible, $\pm\boldsymbol{m}$. Following (3.17) the correct sign is then chosen such that

$$\boldsymbol{f} : \boldsymbol{c}^{\mathrm{e}} : \boldsymbol{\xi} > 0, \quad \boldsymbol{\xi} = \mathrm{sym}(\boldsymbol{m} \otimes \boldsymbol{n}). \tag{3.24}$$

Thus, the localization theory determines not only the critical band orientation normal vector $\boldsymbol{n}$ but also the characteristic tensor $\boldsymbol{\xi}$.

The 'trace' of the tensor $\boldsymbol{\xi}$, $\mathrm{tr}\,\boldsymbol{\xi} = \boldsymbol{n} \cdot \boldsymbol{m}$, determines the nature of the deformation band at localization. We define the following possible types of deformation band.

$$\begin{cases} \boldsymbol{m} \cdot \boldsymbol{n} = 1: & \text{pure dilation band;} \\ 0 < \boldsymbol{m} \cdot \boldsymbol{n} < 1: & \text{dilatant shear band;} \\ \boldsymbol{m} \cdot \boldsymbol{n} = 0: & \text{simple shear band;} \\ -1 < \boldsymbol{m} \cdot \boldsymbol{n} < 0: & \text{compactive shear band;} \\ \boldsymbol{m} \cdot \boldsymbol{n} = -1: & \text{pure compaction band.} \end{cases} \tag{3.25}$$

In a simple shear band the instantaneous velocity jump vector $[\![\boldsymbol{v}]\!]$ is tangent to the band. In a dilatant (compactive) shear band the angle between the unit vectors $\boldsymbol{n}$ and $\boldsymbol{m}$ is acute (obtuse), and thus the band exhibits some form of instantaneous expansion (contraction), see Figs. 1 and 5.

The nature of a deformation band obviously depends on the position of the stress point on the yield surface at the moment of localization. Assuming the metric tensor $\boldsymbol{c}^{\mathrm{e}}$ is given by (3.12), we can decompose $\boldsymbol{f}$ into volumetric and deviatoric parts as

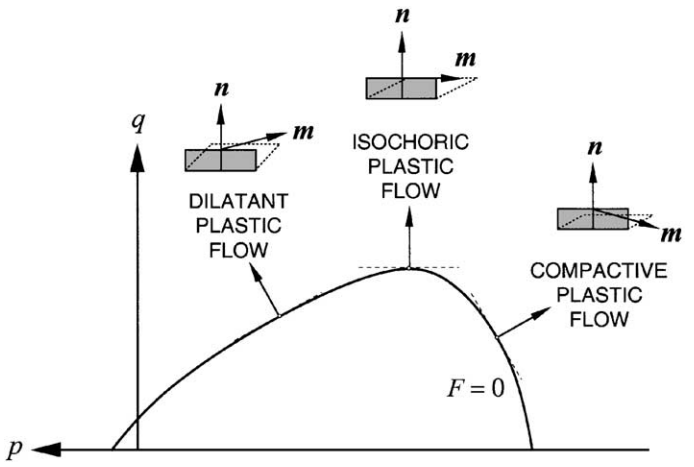

Fig. 5. Representation of dilatant, isochoric, and compactive plastic flow on a semi-meridian plane: p = mean normal stress; q = norm of second invariant of deviatoric stress; F = yield function.

$$\boldsymbol{f} = \bar{f}\mathbf{1} + \boldsymbol{f}', \quad \bar{f} = \operatorname{tr}\boldsymbol{f}/3, \quad \operatorname{tr}\boldsymbol{f}' = 0, \tag{3.26}$$

and write (3.24) as

$$\boldsymbol{f} : \boldsymbol{c}^{\mathrm{e}} : \boldsymbol{\xi} = 3K\bar{f}\operatorname{tr}\boldsymbol{\xi} + 2\mu f'_{mn} > 0, \quad f'_{mn} = \boldsymbol{m}\cdot\boldsymbol{f}'\cdot\boldsymbol{n}. \tag{3.27}$$

We note that in the regime of a pure shear band where $\operatorname{tr}\boldsymbol{\xi} \approx 0$, the quantity f'_{mn} must be non-negative or the inequality (3.27) will not be satisfied, while in the regime where $\boldsymbol{f}' \approx \mathbf{0}$, the relation $\operatorname{sign}(\bar{f}) = \operatorname{sign}(\operatorname{tr}\boldsymbol{\xi})$ must hold for the scalar product (3.27) to remain positive. The latter case implies that the shear band is either dilatant or compactive depending on whether the stress point at localization lies on the dilatant or compactive side of the yield surface. However, in the former case it is possible for a shear band to be compactive even if the stress point lies on the dilatant side of the yield surface if either $\bar{f}$ or $\operatorname{tr}\boldsymbol{\xi}$ is small enough for the scalar product (3.27) to remain positive. Graphical representations of dilatant, isochoric, and compactive plastic flows are shown in Fig. 5.

3.2. Pure compaction and pure dilation bands

We now investigate the possibility that the determinant condition (3.23) is satisfied for some critical band orientation $\boldsymbol{n}$ and that the characteristic vector $\boldsymbol{m}$ of the tensor $\boldsymbol{A}$ is parallel to $\boldsymbol{n}$. Because of the homogeneous form of the characteristic equation, it suffices to assume that $\boldsymbol{m} = \boldsymbol{n}$. The traction continuity condition then writes

$$(\dot{\zeta}/h)\boldsymbol{A}\cdot\boldsymbol{n} = \mathbf{0}. \tag{3.28}$$

Like the tensor $\boldsymbol{f}$, we can also decompose the tensor $\boldsymbol{q}$ into volumetric and deviatoric parts,

$$\boldsymbol{q} = \bar{q}\mathbf{1} + \boldsymbol{q}', \quad \bar{q} = \operatorname{tr}\boldsymbol{q}/3, \quad \operatorname{tr}\boldsymbol{q}' = 0. \tag{3.29}$$

The localization condition then specializes to the form

$$\boldsymbol{A}\cdot\boldsymbol{n} = \alpha\boldsymbol{n} - \beta\boldsymbol{n}\cdot\boldsymbol{q}' = \mathbf{0}, \tag{3.30}$$

where

$$\alpha = K + \frac{4\mu}{3} - \frac{3K\bar{q}}{\chi}(3K\bar{f} + 2\mu f'_{nn}), \tag{3.31a}$$

$$\beta = \frac{2\mu}{\chi}(3K\bar{f} + 2\mu f'_{nn}), \tag{3.31b}$$

$f'_{nn} = \boldsymbol{n} \cdot \boldsymbol{f}' \cdot \boldsymbol{n}$, and χ is given in (3.16).

Eq. (3.30) can be satisfied if and only if the vector $\boldsymbol{n}$ is parallel to the vector $\boldsymbol{n} \cdot \boldsymbol{q}'$. Consider now the following spectral representation of $\boldsymbol{q}'$

$$\boldsymbol{q}' = \sum_{A=1}^{3} q'_A \boldsymbol{n}^{(A)} \otimes \boldsymbol{n}^{(A)}, \tag{3.32}$$

where the q'_A's are the principal values and the $\boldsymbol{n}^{(A)}$'s are the corresponding principal directions. Substituting in (3.30) gives

$$\alpha \boldsymbol{n} - \beta \sum_{A=1}^{3} q'_A \cos\theta_A \boldsymbol{n}^{(A)} = \boldsymbol{0}, \tag{3.33}$$

where $\cos\theta_A = \boldsymbol{n} \cdot \boldsymbol{n}^{(A)}$ is the direction cosine of the angle between the unit band normal vector $\boldsymbol{n}$ and the principal direction $\boldsymbol{n}^{(A)}$. For this equation to make sense, $\boldsymbol{n} = \boldsymbol{n}^{(A)}$, which means that the unit normal to the band should coincide with one of the principal directions of $\boldsymbol{q}'$. Equivalently, $\boldsymbol{n}$ should coincide with one of the principal directions of the total tensor $\boldsymbol{q}$ itself, since the volumetric part $\bar{q}\boldsymbol{1}$ is 'neutral' with respect to the orientation of the principal axes.

In coaxial flow theory of plasticity on which the present discussion has focused so far, the principal axes of $\boldsymbol{q}$ coincide with those of the Cauchy stress tensor $\boldsymbol{\sigma}$. This means that $\boldsymbol{\sigma}$ is also amenable to the spectral representation

$$\boldsymbol{\sigma} = \sum_{A=1}^{3} \sigma_A \boldsymbol{n}^{(A)} \otimes \boldsymbol{n}^{(A)}, \tag{3.34}$$

where the σ_A's are the principal values, and the $\boldsymbol{n}^{(A)}$'s are the same principal directions as those of the tensor $\boldsymbol{q}$ itself (by definition of coaxiality). Consequently, the orientation of a compaction or dilation band coincides with a principal axis of the stress tensor, which is consistent with the conclusion of Issen and Rudnicki [13]. However, in non-coaxial flow theory this is not the case [51,52], and thus it must stated that, strictly speaking, the orientation of a compaction/dilation band coincides with the direction of a principal axis of the plastic flow direction $\boldsymbol{q}$ and not that of the stress tensor $\boldsymbol{\sigma}$.

In many cases coaxiality in the principal directions of $\boldsymbol{\sigma}$ and the second-order tensor $\boldsymbol{f}$ may also be demonstrated. Any isotropic yield function of stresses, for example, produces a stress gradient tensor $\boldsymbol{f}$ that has the same principal directions as those of $\boldsymbol{\sigma}$. Thus, we can also write the tensor $\boldsymbol{f}'$ in spectral form as

$$\boldsymbol{f}' = \sum_{A=1}^{3} f'_A \boldsymbol{n}^{(A)} \otimes \boldsymbol{n}^{(A)}. \tag{3.35}$$

Taking a band orientation $\boldsymbol{n}$ coinciding with a principal direction $\boldsymbol{n}^{(A)}$, the vector equation (3.33) reduces to

$$\left[K + \frac{4\mu}{3} - \frac{\tilde{q}_A \tilde{f}_A}{\chi}\right] \boldsymbol{n}^{(A)} = \boldsymbol{0} \quad (\text{no sum on } A), \tag{3.36}$$

where

$$\tilde{q}_A = 3K\bar{q} + 2\mu q'_A, \qquad \tilde{f}_A = 3K\bar{f} + 2\mu f'_A. \tag{3.37}$$

For non-trivial solution the above equation can be satisfied if and only if the scalar coefficient of $\boldsymbol{n}^{(A)}$ vanishes. Setting this quantity to zero gives

$$H = \left(K + \frac{4\mu}{3}\right)^{-1} \tilde{q}_A \tilde{f}_A - \overline{\chi} \quad \text{(no sum on } A\text{)}, \tag{3.38}$$

where

$$\overline{\chi} = \boldsymbol{f} : \boldsymbol{c}^{\mathrm{e}} : \boldsymbol{q} = 9K\bar{f}\bar{q} + 2\mu \sum_{A=1}^{3} f'_A q'_A > 0. \tag{3.39}$$

In a typical loading program where the load–displacement curve exhibits a degrading slope, the plastic modulus decreases with ongoing plastic deformation. Thus, the orientation for which the localization condition is first satisfied is that which yields the maximum value of H [24]. With reference to Eq. (3.38), this means that if a pure compaction/dilation band is to form the critical band orientation must coincide with one of the three principal directions of $\boldsymbol{q}$, and specifically at a value of A equal to either 1, 2 or 3 for which the quantity $\tilde{q}_A \tilde{f}_A$ is maximized. An ideal condition would be for $\tilde{q}_A$ and $\tilde{f}_A$ to carry the same sign in order for their product to be greater than zero and thus maximize its algebraic value. In fact, the associative flow rule gives $\boldsymbol{q} = \boldsymbol{f}$ and maximizes the quantity $\tilde{q}_A \tilde{f}_A$, thus favoring the development of either a compaction or dilation band.

For a deformation band with no shear offset the two possible characteristic tensors are

$$\boldsymbol{\xi} = \begin{cases} \boldsymbol{n} \otimes \boldsymbol{n} & \text{for pure dilation band;} \\ -\boldsymbol{n} \otimes \boldsymbol{n} & \text{for pure compaction band.} \end{cases} \tag{3.40}$$

Either tensor only has one non-zero eigenvalue corresponding to the stretching or compression mode, unlike the general shear band characteristic tensor $\mathrm{sym}(\boldsymbol{m} \otimes \boldsymbol{n})$ which has two non-zero eigenvalues corresponding to the stretching/compression and shearing modes. To determine the actual mode, i.e., compression or dilation, we check the sign

$$\boldsymbol{f} : \boldsymbol{c}^{\mathrm{e}} : \boldsymbol{\xi} = \pm(3\bar{f}K + 2\mu f'_A) \equiv \pm\tilde{f}_A > 0. \tag{3.41}$$

If $\tilde{f}_A > 0$, then $\boldsymbol{m} = \boldsymbol{n}$ and a pure dilation band would form; if $\tilde{f}_A < 0$, then $\boldsymbol{m} = -\boldsymbol{n}$ and we obtain a pure compaction band. Once again, the framework described above is complete in that it determines the critical band orientation $\boldsymbol{n}$ as well as the characteristic tensor $\boldsymbol{\xi}$.

Remark 1. If plastic flow is non-coaxial, then f'_{nn} must be used in lieu of f'_A and the rest of the formulation remains true. This is because relations (3.31a,b) are valid regardless of the nature of plastic flow, and f'_{nn} reduces to a principal value only for the case of coaxial theory.

Remark 2. If F and Q are both isotropic functions of stresses, then they are expressible in terms of the principal stresses, and thus we can write $f_A = \partial F/\partial\sigma_A$ and $q_A = \partial Q/\partial\sigma_A$. In this case, the invariants are $\bar{f} = (\sum_{A=1}^{3} f_A)/3$ and $\bar{q} = (\sum_{A=1}^{3} q_A)/3$, and the principal deviatoric values are $f'_A = f_A - \bar{f}$ and $q'_A = q_A - \bar{q}$.

3.3. Spectral representation of tangent constitutive operator

We next turn to the spectral representation of the elastoplastic constitutive operator $\boldsymbol{c}^{\mathrm{ep}}$. Here we restrict the discussion to coaxial flow theory and write the following tensors in spectral form as

$$\boldsymbol{\sigma} = \sum_{A=1}^{3} \sigma_A \boldsymbol{m}^{(A)}, \qquad \boldsymbol{\epsilon}^{\mathrm{e}} = \sum_{A=1}^{3} \epsilon_A^{\mathrm{e}} \boldsymbol{m}^{(A)}, \qquad \boldsymbol{f} = \sum_{A=1}^{3} f_A \boldsymbol{m}^{(A)}, \qquad \boldsymbol{q} = \sum_{A=1}^{3} q_A \boldsymbol{m}^{(A)}, \tag{3.42}$$

where σ_A, ϵ_A^{e}, $f_A = \partial F/\partial \sigma_A$, and $q_A = \partial Q/\partial \sigma_A$ are the spectral values of the respective tensors, $\boldsymbol{m}^{(A)} = \boldsymbol{n}^{(A)} \otimes \boldsymbol{n}^{(A)}$ are the spectral directions, and the $\boldsymbol{n}^{(A)}$'s are the (mutually orthogonal) unit eigenvectors. That the stress and elastic strain tensors have the same principal directions is a consequence of isotropy in the elastic response.

From the spectral forms for $\boldsymbol{\sigma}$ and $\boldsymbol{\epsilon}^{\mathrm{e}}$ we readily write the elastic constitutive operator $\boldsymbol{c}^{\mathrm{e}} = \partial \boldsymbol{\sigma}/\partial \boldsymbol{\epsilon}^{\mathrm{e}}$ also in spectral form as [53]

$$\boldsymbol{c}^{\mathrm{e}} = \sum_{A=1}^{3} \sum_{B=1}^{3} a_{AB}^{\mathrm{e}} \boldsymbol{m}^{(A)} \otimes \boldsymbol{m}^{(B)} + \frac{1}{2} \sum_{A=1}^{3} \sum_{B \neq A} \left(\frac{\sigma_B - \sigma_A}{\epsilon_B^{\mathrm{e}} - \epsilon_A^{\mathrm{e}}} \right) \left(\boldsymbol{m}^{(AB)} \otimes \boldsymbol{m}^{(AB)} + \boldsymbol{m}^{(AB)} \otimes \boldsymbol{m}^{(BA)} \right), \tag{3.43}$$

where

$$[a_{AB}^{\mathrm{e}}] = \begin{bmatrix} a & b & b \\ b & a & b \\ b & b & a \end{bmatrix}, \quad a = K + \frac{4\mu}{3}, \quad b = K - \frac{2\mu}{3}$$

is the elasticity matrix in principal axes, and $\boldsymbol{m}^{(AB)} = \boldsymbol{n}^{(A)} \otimes \boldsymbol{n}^{(B)}$. The matrix $[a_{AB}^{\mathrm{e}}]$ relates the principal Cauchy stress σ_A to the principal elastic strain $\varepsilon_B^{\mathrm{e}}$ for $A, B = 1, 2, 3$ in accordance with the generalized Hooke's law of linear elasticity. As for the expression for $\boldsymbol{c}^{\mathrm{e}}$, the first summations on the right-hand side represent the contributions of the material part, whereas the second summations reflect the spin of principal axes.

From the orthogonality of the principal axes, we see that $\boldsymbol{m}^{(AB)} : \boldsymbol{m}^{(C)} = (\boldsymbol{n}^{(A)} \cdot \boldsymbol{n}^{(C)})(\boldsymbol{n}^{(B)} \cdot \boldsymbol{n}^{(C)}) \equiv 0$ for any combinations of the eigendirections A, B, and C, provided that $A \neq B$. Thus the spin component of $\boldsymbol{c}^{\mathrm{e}}$ is orthogonal to the tensors $\boldsymbol{f}$ and $\boldsymbol{q}$, implying that its inner products with these tensors vanish. Consequently, $\boldsymbol{c}^{\mathrm{ep}}$ is also amenable to the spectral representation

$$\boldsymbol{c}^{\mathrm{ep}} = \sum_{A=1}^{3} \sum_{B=1}^{3} a_{AB}^{\mathrm{ep}} \boldsymbol{m}^{(A)} \otimes \boldsymbol{m}^{(B)} + \frac{1}{2} \sum_{A=1}^{3} \sum_{B \neq A} \left(\frac{\sigma_B - \sigma_A}{\epsilon_B^{\mathrm{e}} - \epsilon_A^{\mathrm{e}}} \right) \left(\boldsymbol{m}^{(AB)} \otimes \boldsymbol{m}^{(AB)} + \boldsymbol{m}^{(AB)} \otimes \boldsymbol{m}^{(BA)} \right), \tag{3.44}$$

where $[a_{AB}^{\mathrm{ep}}]$ is the matrix of elastoplastic moduli in principal axes with components

$$a_{AB}^{\mathrm{ep}} = a_{AB}^{\mathrm{e}} - \frac{1}{\chi} \tilde{q}_A \tilde{f}_B, \quad \chi = \sum_{A=1}^{3} \sum_{B=1}^{3} f_A a_{AB}^{\mathrm{e}} q_B + H, \quad \tilde{f}_B = \sum_{C=1}^{3} f_C a_{CB}^{\mathrm{e}}, \quad \tilde{q}_A = \sum_{D=1}^{3} a_{AD}^{\mathrm{e}} q_D. \tag{3.45}$$

Substituting the spectral form of $\boldsymbol{c}^{\mathrm{ep}}$ into the localization condition (3.28) for pure compaction/dilation band gives

$$\boldsymbol{A} \cdot \boldsymbol{n} = \sum_{A=1}^{3} \sum_{B=1}^{3} a_{AB}^{\mathrm{ep}} \cos\theta_A \cos^2\theta_B \boldsymbol{n}^{(A)} + \frac{1}{2} \sum_{A=1}^{3} \sum_{B \neq A} (\sigma_B - \sigma_A) \cos\theta_A \cos^2\theta_B \boldsymbol{n}^{(A)} = \boldsymbol{0}, \tag{3.46}$$

where $\cos\theta_A = \boldsymbol{n} \cdot \boldsymbol{n}^{(A)}$ is the direction cosine of the angle between the potential compaction/dilation band normal $\boldsymbol{n}$ and the principal direction $\boldsymbol{n}^{(A)}$. This vector equation can have a solution if and only if $\boldsymbol{n} = \pm \boldsymbol{n}^{(A)}$, i.e., if $\boldsymbol{n}$ is parallel to any of the principal axes. The second summations thus drop out (since $A \neq B$), and for non-trivial solution to exist we must have

$$a_{AA}^{\mathrm{ep}} \boldsymbol{n}^{(A)} = \boldsymbol{0} \Rightarrow a_{AA}^{\mathrm{ep}} = 0 \quad (\text{no sum on } A). \tag{3.47}$$

This is an alternative form of the localization condition (3.36). This result states that for localization to take place in the form of either pure compaction or pure dilation bands the initial vanishing of the determinant of the elastoplastic acoustic tensor must be due to the vanishing of a diagonal element of the elastoplastic constitutive matrix in principal axes. Furthermore, the unit normal to the compaction/dilation band is parallel to the principal axis corresponding to this particular vanishing diagonal element. If the initial vanishing of the determinant of the elastoplastic acoustic tensor is not due to the vanishing of any of the diagonal elements of the elastoplastic constitutive matrix in principal axes, then a pure compaction/dilation band is not possible and we expect to have a shear band.

4. Formulation of deformation bands: finite deformation case

The general analysis of deformation bands in the finite deformation regime is described in [32] (see also [54,55]) and only key points relevant to the assessment of the relative contributions of compaction/dilation and shearing are summarized herein. The kinematics of the problem changes slightly from the infinitesimal case in that we now deal with two configurations, reference and deformed.

4.1. General analysis of deformation bands

We then let $\phi : \mathscr{B} \to \mathscr{B}'$ be a C^1 configuration of $\mathscr{B}$ in $\mathscr{B}'$, where $\mathscr{B}$ and $\mathscr{B}'$ are the reference and deformed configurations of a body with smooth boundaries $\partial\mathscr{B}$ and $\partial\mathscr{B}'$, respectively, see Fig. 6. We assume an emerging deformation band defined by a pair of surfaces $\mathscr{S}_0$ and $\widehat{\mathscr{S}}_0$ and separated by the band thickness

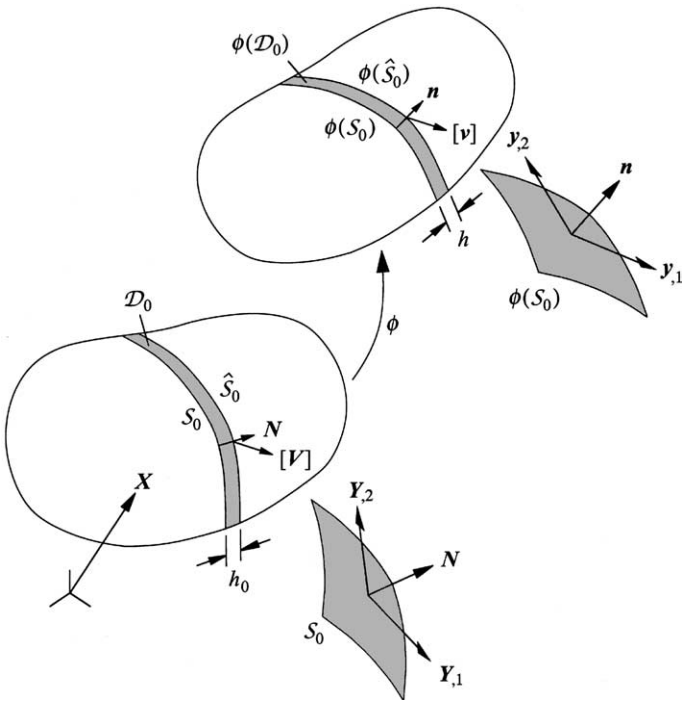

Fig. 6. Normal parameterization of shear band geometry: finite deformation case.

h_0, all reckoned with respect to the undeformed configuration. The mathematical representation of $\mathscr{S}_0 \subset \mathscr{B}$ is

$$\mathscr{S}_0 = \{ \boldsymbol{Y} = \widehat{\boldsymbol{Y}}(\xi^1, \xi^2) \mid (\xi^1, \xi^2) \in \mathscr{B} \}, \tag{4.1}$$

where $\widehat{\boldsymbol{Y}} : \mathscr{B} \to R^{n_{\mathrm{sd}}}$ is a smooth global parameterization, and ξ^1, ξ^2 are two tangential parameters to $\mathscr{S}_0$. Thus, the unit normal to $\mathscr{S}_0$ is

$$\boldsymbol{N} = \widehat{\boldsymbol{N}}(\xi^1, \xi^2) = \widehat{\boldsymbol{Y}}_{,1} \times \widehat{\boldsymbol{Y}}_{,2} / \| \widehat{\boldsymbol{Y}}_{,1} \times \widehat{\boldsymbol{Y}}_{,2} \|. \tag{4.2}$$

In the deformed configuration the surface $\mathscr{S} = \phi(\mathscr{S}_0)$ is given by

$$\mathscr{S} = \{ \boldsymbol{y} = \widehat{\boldsymbol{y}}(\zeta^1, \zeta^2) | (\zeta^1, \zeta^2) \in \phi(\mathscr{B}) \}, \tag{4.3}$$

where, again, $\widehat{\boldsymbol{y}} : \phi(\mathscr{B}) \to R^{n_{\mathrm{sd}}}$ is a smooth global parameterization. The unit normal to $\mathscr{S}$ is

$$\boldsymbol{n} = \widehat{\boldsymbol{n}}(\zeta^1, \zeta^2) = \widehat{\boldsymbol{y}}_{,1} \times \widehat{\boldsymbol{y}}_{,2} / \| \widehat{\boldsymbol{y}}_{,1} \times \widehat{\boldsymbol{y}}_{,2} \|. \tag{4.4}$$

The two unit normal vectors $\boldsymbol{N}$ and $\boldsymbol{n}$ are related by Nanson's formula, $\boldsymbol{n}\,\mathrm{d}a = J\boldsymbol{F}^{-\mathrm{t}} \cdot \boldsymbol{N}\mathrm{d}A$, where $\mathrm{d}a$ and $\mathrm{d}A$ are infinitesimal surface areas whose unit normals are $\boldsymbol{n}$ and $\boldsymbol{N}$, respectively, $\boldsymbol{F}$ is the deformation gradient, and $J = \det(\boldsymbol{F}) = \mathrm{d}v/\mathrm{d}V$ is the Jacobian [53]. If we denote the band thicknesses as h_0 and h in the reference and deformed configurations, respectively, then $\mathrm{d}v = h\,\mathrm{d}a$, $\mathrm{d}V = h_0\,\mathrm{d}A$, and we thus obtain the relation

$$\boldsymbol{N} \cdot \boldsymbol{F}^{-1} / h_0 = \boldsymbol{n}/h. \tag{4.5}$$

As usual, we herein assume the band thicknesses h_0 and h to be small.

We now investigate the emergence of a ramp-like velocity field across the band, and denote the relative velocity between the opposite band faces $\mathscr{S}_0$ and $\widehat{\mathscr{S}_0}$ by $[\![\boldsymbol{V}]\!]$. These two surfaces are the same material surfaces $\mathscr{S}$ and $\widehat{\mathscr{S}}$ in the deformed configuration (see Fig. 6), so the relative velocity $[\![\boldsymbol{v}]\!]$ in the deformed configuration is the same as $[\![\boldsymbol{V}]\!]$ itself, i.e., $[\![\boldsymbol{V}]\!] = [\![\boldsymbol{v}]\!] \circ \phi$. The jump discontinuity may be expressed through the rate of deformation gradient,

$$\dot{\boldsymbol{F}} = \begin{cases} \dot{\overline{\boldsymbol{F}}} & \text{in } \mathscr{B} \setminus \overline{\mathscr{D}}_0; \\ \dot{\overline{\boldsymbol{F}}} + ([\![\boldsymbol{V}]\!] \otimes \boldsymbol{N})/h_0 & \text{in } \mathscr{D}_0, \end{cases} \tag{4.6}$$

where $\dot{\boldsymbol{F}} = \mathrm{GRAD}\,\boldsymbol{V}$, $\dot{\overline{\boldsymbol{F}}} = \mathrm{GRAD}\,\overline{\boldsymbol{V}}$, $\overline{\boldsymbol{V}}$ is the continuous velocity field, and $\mathscr{D}_0 = \mathscr{S}_0 \times (0, h_0)$ is the (open) shear band domain. Alternately, we can define the jump discontinuity through the velocity gradient field,

$$\boldsymbol{l} = \begin{cases} \bar{\boldsymbol{l}} & \text{in } \phi(\mathscr{B}) \setminus \overline{\mathscr{D}}; \\ \bar{\boldsymbol{l}} + ([\![\boldsymbol{v}]\!] \otimes \boldsymbol{n})/h & \text{in } \mathscr{D}, \end{cases} \tag{4.7}$$

where $([\![\boldsymbol{v}]\!] \otimes \boldsymbol{n})/h = ([\![\boldsymbol{V}]\!] \otimes \boldsymbol{N} \cdot \boldsymbol{F}^{-1})/h_0$. Upon evaluating just inside and just outside the surface of discontinuity, we obtain the equivalent relations

$$\dot{\boldsymbol{F}}^1 = \dot{\boldsymbol{F}}^0 + \frac{1}{h_0}[\![\boldsymbol{V}]\!] \otimes \boldsymbol{N} \Longleftrightarrow \boldsymbol{l}^1 = \boldsymbol{l}^0 + \frac{1}{h}[\![\boldsymbol{v}]\!] \otimes \boldsymbol{n}, \tag{4.8}$$

where $\dot{\boldsymbol{F}}^1$ and $\dot{\boldsymbol{F}}^0$ ($\boldsymbol{l}^1$ and $\boldsymbol{l}^0$) are the rates of deformation gradients (spatial velocity gradients) just inside and just outside the surface of discontinuity, respectively.

Next we investigate the mode of deformation at bifurcation, again focusing on the relative shear and volumetric responses exhibited by the potential deformation band. Here we employ multiplicative plasticity theory and formulate the constitutive model in terms of the symmetric Kirchhoff stress tensor $\boldsymbol{\tau} = J\boldsymbol{\sigma}$. Accordingly, we assume that the yield and plastic potential functions, F and Q, respectively, are now expressed in terms of the stress tensor $\boldsymbol{\tau}$, and re-define the stress gradient tensors as $\boldsymbol{f} = \partial F/\partial \boldsymbol{\tau}$ and $\boldsymbol{q} = \partial G/\partial \boldsymbol{\tau}$. As in the infinitesimal theory, we assume the inequalities

$$\boldsymbol{f} : \boldsymbol{\alpha}^{\mathrm{e}} : \boldsymbol{l}^{1} > 0, \qquad \boldsymbol{f} : \boldsymbol{\alpha}^{\mathrm{e}} : \boldsymbol{l}^{0} > 0, \tag{4.9}$$

where $\boldsymbol{\alpha}^{\mathrm{e}}$ is the fourth-order tensor of hyperelastic moduli relating the Kirchhoff stress rate tensor $\dot{\boldsymbol{\tau}}$ to the elastic velocity gradient tensor $\boldsymbol{l}^{\mathrm{e}}$ [32]. The above inequalities suggest plastic yielding on both sides of the band at localization.

Just inside the surface $\mathscr{S}$ we write the rate constitutive equation as

$$\dot{\boldsymbol{\tau}} = \boldsymbol{\alpha}^{\mathrm{e}} : (\boldsymbol{l}^{1} - \dot{\lambda}\boldsymbol{q}), \tag{4.10}$$

where $\dot{\lambda}\boldsymbol{q} := \boldsymbol{l}^{\mathrm{p}}$ and $\boldsymbol{l}^{\mathrm{p}}$ is the plastic component of the velocity gradient $\boldsymbol{l}^{1}$. Here, $\boldsymbol{l}^{\mathrm{p}}$ is a symmetric tensor following a common assumption of zero plastic spin [56]. Issues pertaining to this assumption, as well as the development of plastic spin in the post-localization regime, are discussed in [56,57]. For continued yielding inside the band in the finite deformation regime, we must have

$$\boldsymbol{f} : \boldsymbol{\alpha}^{\mathrm{e}} : (\llbracket \boldsymbol{v} \rrbracket \otimes \boldsymbol{n}) = \frac{h}{h_0} \boldsymbol{f} : \boldsymbol{\alpha}^{\mathrm{e}} : (\llbracket \boldsymbol{V} \rrbracket \otimes \boldsymbol{N} \cdot \boldsymbol{F}^{-1}) > 0. \tag{4.11}$$

Note that the full value of the tensor $(\llbracket \boldsymbol{v} \rrbracket \otimes \boldsymbol{n})$ is now used in the finite deformation case, unlike in the infinitesimal deformation case where only the symmetric part of this tensor is relevant, cf. (3.17).

In the finite deformation regime the rate constitutive equations just inside and just outside the surface of discontinuity take the form

$$\dot{\boldsymbol{\tau}}^{1} = \boldsymbol{\alpha}^{\mathrm{ep}} : \boldsymbol{l}^{1}, \qquad \dot{\boldsymbol{\tau}}^{0} = \boldsymbol{\alpha}^{\mathrm{ep}} : \boldsymbol{l}^{0}, \tag{4.12}$$

where

$$\boldsymbol{\alpha}^{\mathrm{ep}} = \boldsymbol{\alpha}^{\mathrm{e}} - \frac{1}{\chi} \boldsymbol{\alpha}^{\mathrm{e}} : \boldsymbol{q} \otimes \boldsymbol{f} : \boldsymbol{\alpha}^{\mathrm{e}}, \quad \chi = \boldsymbol{f} : \boldsymbol{\alpha}^{\mathrm{e}} : \boldsymbol{q} + H, \tag{4.13}$$

and H is the usual plastic modulus. Introducing the non-symmetric first Piola–Kirchhoff stress tensor $\boldsymbol{P} = \boldsymbol{\tau} \cdot \boldsymbol{F}^{-\mathrm{t}}$, the associated rates are

$$\dot{\boldsymbol{P}}^{1} = \mathbf{A}^{\mathrm{ep}} : \dot{\boldsymbol{F}}^{1}, \qquad \dot{\boldsymbol{P}}^{0} = \mathbf{A}^{\mathrm{ep}} : \dot{\boldsymbol{F}}^{0}, \tag{4.14}$$

where $\mathbf{A}^{\mathrm{ep}}$ is the elastoplastic tangential moduli tensor with components

$$\mathsf{A}^{\mathrm{ep}}_{iAjB} = F^{-1}_{Ak} F^{-1}_{Bl} \mathsf{a}^{\mathrm{ep}}_{ikjl}, \quad \mathsf{a}^{\mathrm{ep}}_{ikjl} = \alpha^{\mathrm{ep}}_{ikjl} - \tau_{il}\delta_{jk}, \tag{4.15}$$

τ_{il} is a component of the Kirchhoff stress tensor on either side of the band, and δ_{jk} is the Kronecker delta.

For a deformation band to be possible the nominal traction rate must be continuous,

$$\dot{\boldsymbol{P}}^{1} \cdot \boldsymbol{N} = \dot{\boldsymbol{P}}^{0} \cdot \boldsymbol{N}. \tag{4.16}$$

Writing $\llbracket \boldsymbol{v} \rrbracket = \llbracket \boldsymbol{V} \rrbracket = \dot{\zeta}\boldsymbol{m} = \dot{\zeta}\boldsymbol{M}$, where $\dot{\zeta} > 0$ is the magnitude and $\boldsymbol{m} \equiv \boldsymbol{M}$ is the unit direction of the relative velocity vector, the localization condition takes the following alternative forms

$$\frac{\dot{\zeta}}{h_0} \boldsymbol{A} \cdot \boldsymbol{M} = \boldsymbol{0}, \quad A_{ij} = N_A \mathsf{A}^{\mathrm{ep}}_{iAjB} N_B, \tag{4.17}$$

or

$$\frac{\dot{\zeta} h_0}{h^2} \boldsymbol{a} \cdot \boldsymbol{m} = \boldsymbol{0}, \quad a_{ij} = n_k \mathsf{a}^{\mathrm{ep}}_{ikjl} n_l, \tag{4.18}$$

where $\boldsymbol{A}$ and $\boldsymbol{a}$ are, respectively, the Lagrangian and Eulerian acoustic tensors related through the band thickness via the relation $\boldsymbol{A} = (h_0/h)^2 \boldsymbol{a}$ (see [32]).

For a finite band thickness non-trivial solutions to the above equations exist if and only if

$$\det \boldsymbol{A} = \det \boldsymbol{a} = 0. \tag{4.19}$$

Setting $\det \boldsymbol{A} = 0$ identifies the critical unit band normal vector $\boldsymbol{N}$ reckoned with respect to the reference configuration, whereas setting $\det \boldsymbol{a} = 0$ gives the unit normal vector $\boldsymbol{n}$ to the same material band, but now reckoned with respect to the deformed configuration. Since $\boldsymbol{N}$ and $\boldsymbol{n}$ refer to the same material band, the vanishing of the two determinants occurs at the same time. Furthermore, at the bifurcation point the eigenvectors of the acoustic tensors $\boldsymbol{A}$ and $\boldsymbol{a}$ are the same (since $\boldsymbol{A}$ and $\boldsymbol{a}$ are the same tensor save for a scalar multiplier). The unit eigenvectors are precisely either $\pm\boldsymbol{m}$ or $\pm\boldsymbol{M}$, where the correct sign is chosen such that

$$\boldsymbol{f} : \boldsymbol{\alpha}^{\mathrm{e}} : (\boldsymbol{m} \otimes \boldsymbol{n}) > 0. \tag{4.20}$$

Like in the infinitesimal case the type of the resulting deformation band depends on the value of the scalar product $\boldsymbol{m} \cdot \boldsymbol{n}$ as defined in (3.25). Note that a pure shear band in the finite deformation case requires that $\boldsymbol{m}$ be perpendicular to $\boldsymbol{n}$, i.e., $\boldsymbol{m} \cdot \boldsymbol{n} = (h/h_0)\boldsymbol{M} \cdot \boldsymbol{F}^{-\mathrm{t}} \cdot \boldsymbol{N} = 0$. In other words, the orthogonality is defined in the current configuration and not in the reference configuration.

4.2. Pure compaction and pure dilation bands

Still focusing on the finite deformation case, we again investigate the possibility that the determinant condition (4.19) is satisfied for some critical band orientation $\boldsymbol{n}$ and that the eigenvector $\boldsymbol{m}$ of the tensor $\boldsymbol{a}$ (or $\boldsymbol{A}$) is parallel to $\boldsymbol{n}$. It suffices to take $\boldsymbol{m} = \boldsymbol{n}$ in (4.18) to get

$$\boldsymbol{a} \cdot \boldsymbol{n} = \boldsymbol{0}, \quad a_{ij} = n_k \mathsf{a}^{\mathrm{ep}}_{ikjl} n_l. \tag{4.21}$$

Equivalently, using the second of (4.15) we rewrite (4.21) as

$$\boldsymbol{a} \cdot \boldsymbol{n} = \boldsymbol{n} \cdot \boldsymbol{\alpha}^{\mathrm{ep}} : (\boldsymbol{n} \otimes \boldsymbol{n}) - \boldsymbol{\tau} \cdot \boldsymbol{n} = \boldsymbol{0}. \tag{4.22}$$

Once again, we consider spectral representations of the tangent operator $\boldsymbol{\alpha}^{\mathrm{ep}}$ and the Kirchhoff stress tensor $\boldsymbol{\tau}$ to simplify the above localization condition. Assuming isotropic elasticity, we first introduce the elastic left Cauchy-Green deformation tensor $\boldsymbol{b}^{\mathrm{e}} := \boldsymbol{F}^{\mathrm{e}} \cdot \boldsymbol{F}^{\mathrm{et}}$ arising from a multiplicative decomposition of $\boldsymbol{F}$ into elastic and plastic parts [58], and write

$$\boldsymbol{\tau} = \sum_{A=1}^{3} \tau_A \boldsymbol{n}^{(A)} \otimes \boldsymbol{n}^{(A)}, \qquad \boldsymbol{b}^{\mathrm{e}} = \sum_{A=1}^{3} b_A \boldsymbol{n}^{(A)} \otimes \boldsymbol{n}^{(A)}, \tag{4.23}$$

where τ_A $(= J\sigma_A)$ and b_A are the principal values of $\boldsymbol{\tau}$ and $\boldsymbol{b}^{\mathrm{e}}$, respectively; and the vectors $\boldsymbol{n}^{(A)}$ are the corresponding principal directions. Note that $\boldsymbol{\tau}$ and $\boldsymbol{b}^{\mathrm{e}}$ have the same principal directions due to the assumed isotropy in the elastic response.

Next, the tensor $\boldsymbol{\alpha}^{\mathrm{e}}$ is expressed in the spectral form [45,46]

$$\boldsymbol{\alpha}^{\mathrm{e}} = \sum_{A=1}^{3}\sum_{B=1}^{3} a^{\mathrm{e}}_{AB} \boldsymbol{m}^{(A)} \otimes \boldsymbol{m}^{(B)} + \frac{1}{2}\sum_{A=1}^{3}\sum_{B\neq A} \left(\frac{\tau_B - \tau_A}{b_B - b_A}\right)\left(b_B \boldsymbol{m}^{(AB)} \otimes \boldsymbol{m}^{(AB)} + b_A \boldsymbol{m}^{(AB)} \otimes \boldsymbol{m}^{(BA)}\right), \tag{4.24}$$

where

$$[a^{\mathrm{e}}_{AB}] = \begin{bmatrix} a & b & b \\ b & a & b \\ b & b & a \end{bmatrix}, \quad a = K + \frac{4\mu}{3}, \ \ b = K - \frac{2\mu}{3}$$

is the elasticity matrix in principal axes, $\boldsymbol{m}^{(A)} = \boldsymbol{n}^{(A)} \otimes \boldsymbol{n}^{(A)}$, and $\boldsymbol{m}^{(AB)} = \boldsymbol{n}^{(A)} \otimes \boldsymbol{n}^{(B)}$. The matrix $[a^{\mathrm{e}}_{AB}]$ relates the principal Kirchhoff stress rate $\dot{\tau}_A$ to the principal elastic logarithmic strain rate $\dot{\varepsilon}^{\mathrm{e}}_B$ for $A, B = 1, 2, 3$ in accordance with the standard hyperelastic Hooke's law in finite deformation elasticity [59]. As for the

expression for $\boldsymbol{\alpha}^{\mathrm{e}}$, the first summations on the right-hand side represent the contributions of the material part, while the second summations represent the effect of the spin of principal axes.

Assuming the yield and plastic potential functions F and Q are now expressed in terms of the invariants of the Kirchhoff stresses, their stress gradients take the form

$$\boldsymbol{f} = \sum_{A=1}^{3} f_A \boldsymbol{n}^{(A)} \otimes \boldsymbol{n}^{(A)}, \qquad \boldsymbol{q} = \sum_{A=1}^{3} q_A \boldsymbol{n}^{(A)} \otimes \boldsymbol{n}^{(A)}, \tag{4.25}$$

where $f_A = \partial F/\partial \tau_A$ and $q_A = \partial Q/\partial \tau_A$. Substituting (4.24) and (4.25) into (4.13) gives

$$\boldsymbol{\alpha}^{\mathrm{ep}} = \sum_{A=1}^{3}\sum_{B=1}^{3} a_{AB}^{\mathrm{ep}} \boldsymbol{m}^{(A)} \otimes \boldsymbol{m}^{(B)} + \frac{1}{2}\sum_{A=1}^{3}\sum_{B\neq A} \left(\frac{\tau_B - \tau_A}{b_B - b_A}\right)\left(b_B \boldsymbol{m}^{(AB)} \otimes \boldsymbol{m}^{(AB)} + b_A \boldsymbol{m}^{(AB)} \otimes \boldsymbol{m}^{(BA)}\right), \tag{4.26}$$

where

$$\begin{aligned} &a_{AB}^{\mathrm{ep}} = a_{AB}^{\mathrm{e}} - \frac{1}{\chi}\tilde{q}_A \tilde{f}_A, \quad \tilde{f}_A = \sum_{C=1}^{3} f_C a_{CA}^{\mathrm{e}}, \quad \tilde{q}_A = \sum_{D=1}^{3} a_{AD}^{\mathrm{e}} q_D, \\ &\chi = \overline{\chi} + H > 0, \quad \overline{\chi} = \sum_{A=1}^{3}\sum_{B=1}^{3} f_A a_{AB}^{\mathrm{e}} q_B > 0. \end{aligned} \tag{4.27}$$

Substituting the spectral forms of $\boldsymbol{\alpha}^{\mathrm{ep}}$ and $\boldsymbol{\tau}$ into the localization condition (4.22) gives

$$\boldsymbol{a}\cdot\boldsymbol{n} = \sum_{A=1}^{3}\sum_{B=1}^{3} a_{AB}^{\mathrm{ep}} \cos\theta_A \cos^2\theta_B \boldsymbol{n}^{(A)} + \frac{1}{2}\sum_{A=1}^{3}\sum_{B\neq A} (\tau_B - \tau_A)\cos\theta_A \cos^2\theta_B \boldsymbol{n}^{(A)} - \sum_{A=1}^{3} \tau_A \cos\theta_A \boldsymbol{n}^{(A)} = \boldsymbol{0}, \tag{4.28}$$

where $\cos\theta_A = \boldsymbol{n}\cdot\boldsymbol{n}^{(A)}$ is the direction cosine of the angle between the potential compaction/dilation band normal $\boldsymbol{n}$ and the principal direction $\boldsymbol{n}^{(A)}$. This vector equation can have a solution if and only if $\boldsymbol{n} = \pm\boldsymbol{n}^{(A)}$, i.e., if $\boldsymbol{n}$ is parallel to any of the principal axes. However, this causes the second summations on the right-hand side of (4.28) to drop out due to the orthogonality of the principal directions, and thus the condition for the onset of compaction/dilation band simplifies to

$$(a_{AA}^{\mathrm{ep}} - \tau_A)\boldsymbol{n}^{(A)} = \boldsymbol{0} \Rightarrow a_{AA}^{\mathrm{ep}} - \tau_A = 0 \quad (\text{no sum on } A). \tag{4.29}$$

This result states that, in the finite deformation regime, for localization to take place in the form of either pure compaction or pure dilation bands the initial vanishing of the determinant of the elastoplastic acoustic tensor must be due to the vanishing of the difference expression $a_{AA}^{\mathrm{ep}} - \tau_A$ for any specific principal axis A. Furthermore, the unit normal to the compaction/dilation band is parallel to the principal axis A on which the above difference expression vanishes. Once again, if the initial vanishing of the determinant of the elastoplastic acoustic tensor is not due to the vanishing of any of the above difference expressions for any principal direction A, then a pure compaction/dilation band is not possible and we expect to have a shear band.

For the case of pure compaction/dilation bands the critical plastic modulus has the form

$$H = \left(K + \frac{4\mu}{3} - J\sigma_A\right)^{-1} \tilde{q}_A \tilde{f}_A - \overline{\chi}. \tag{4.30}$$

To determine whether a pure compaction or pure dilation band will emerge, we again check the sign $\boldsymbol{f} : \boldsymbol{\alpha}^{\mathrm{e}} : \boldsymbol{\xi} > 0$, where $\boldsymbol{\xi} = \boldsymbol{n}\otimes\boldsymbol{n}$ for pure dilation bands and $\boldsymbol{\xi} = -\boldsymbol{n}\otimes\boldsymbol{n}$ for pure compaction bands. Note that the above localization condition for the finite deformation case differs from that for the infinitesimal deformation case only by an additional stress term $J\sigma_A$, cf. (3.38). Of course, in the present case the

parameters K and μ now take on the meaning of being the tangential hyperelastic bulk and shear moduli relating the Kirchhoff stress increments to the elastic logarithmic strain increments.

Remark 3. In [34] it was shown that the stress term in the tangent operator *enhances* the onset of shear strain localization in finite deformation plasticity since it destroys the symmetry of the tangent operator. For pure compaction bands a compressive normal stress induced by finite deformation plasticity (i.e., $\sigma_A < 0$) decreases the numerical value of the critical plastic modulus and thus *delays* the initiation of such bands. The contrast applies equally well with respect to how strain localization is influenced by the plastic flow rule: a non-associative flow rule favors the development of shear bands, whereas an associative flow rule favors the development of compaction/dilation bands [13].

5. Constitutive model for granular rocks

In this section we present a class of three-invariant elastoplastic constitutive models applicable to the analysis of deformation bands in granular rocks. The formulation applies to both infinitesimal and finite deformation plasticity. As a matter of notation, we shall use the Cauchy stresses of the infinitesimal theory in the presentation. However, an extension to the finite deformation regime is fairly straightforward—one simply needs to use the Kirchhoff stresses and the logarithmic strains in lieu of the Cauchy stresses and the infinitesimal strains [59].

5.1. Formulation of the constitutive model

Consider a convex elastic domain $\mathbb{E}$ defined by a smooth yield surface F in the Cauchy stress space $\boldsymbol{\sigma}$:

$$\mathbb{E} = \{(\boldsymbol{\sigma}, \kappa) \in S \times R^1 \mid F(\boldsymbol{\sigma}, \kappa) \leqslant 0\}, \tag{5.1}$$

where S is the space of symmetric rank-two tensors, and $\kappa < 0$ is a stress-like plastic internal variable characterizing the hardening/softening response of the material. The constitutive equation is expressed in terms of a free energy function $\Psi(\boldsymbol{\epsilon}^{\mathrm{e}}, v^{\mathrm{p}})$, where $\boldsymbol{\epsilon}^{\mathrm{e}}$ is the elastic component of the infinitesimal strain tensor and $v^{\mathrm{p}} = \mathrm{tr}(\boldsymbol{\epsilon}^{\mathrm{p}}) < 0$ denotes the compactive plastic volumetric strain conjugate to the plastic internal variable κ. The constitutive equations for $\boldsymbol{\sigma}$ and κ take the form

$$\boldsymbol{\sigma} = \frac{\partial \Psi(\boldsymbol{\epsilon}^{\mathrm{e}}, v^{\mathrm{p}})}{\partial \boldsymbol{\epsilon}^{\mathrm{e}}}; \qquad \kappa = -\frac{\partial \Psi(\boldsymbol{\epsilon}^{\mathrm{e}}, v^{\mathrm{p}})}{\partial v^{\mathrm{p}}}. \tag{5.2}$$

We next consider a plastic potential function $Q(\boldsymbol{\sigma}, \kappa)$, which can be distinct from $F(\boldsymbol{\sigma}, \kappa)$. The special case $\partial Q/\partial \boldsymbol{\sigma} = \partial F/\partial \boldsymbol{\sigma}$ implies associative flow rule, while the case $\partial Q/\partial \kappa = \partial F/\partial \kappa$ pertains to associative hardening. Assuming the total infinitesimal strain tensor is given by the sum $\boldsymbol{\epsilon} = \boldsymbol{\epsilon}^{\mathrm{e}} + \boldsymbol{\epsilon}^{\mathrm{p}}$, then the evolution equations for $\boldsymbol{\epsilon}^{\mathrm{e}}$ and v^{p} take the form

$$\dot{\boldsymbol{\epsilon}}^{\mathrm{e}} = \dot{\boldsymbol{\epsilon}} - \dot{\lambda}\frac{\partial Q(\boldsymbol{\sigma}, \kappa)}{\partial \boldsymbol{\sigma}}; \qquad \dot{v}^{\mathrm{p}} = \dot{\lambda}\frac{\partial Q(\boldsymbol{\sigma}, \kappa)}{\partial \kappa}, \tag{5.3}$$

where $\dot{\lambda}$ is a plastic consistency parameter satisfying the Kuhn–Tucker complementarity conditions

$$\dot{\lambda} \geqslant 0; \quad F(\boldsymbol{\sigma}, \kappa) \leqslant 0; \quad \dot{\lambda} F(\boldsymbol{\sigma}, \kappa) = 0. \tag{5.4}$$

Now, if the strain-like plastic internal variable v^{p} is to represent the plastic volumetric strain, then $\dot{v}^{\mathrm{p}} = \mathrm{tr}(\dot{\boldsymbol{\epsilon}}^{\mathrm{p}}) = \dot{\lambda}\,\mathrm{tr}(\partial Q/\partial \boldsymbol{\sigma})$ by the flow rule, and thus the plastic potential function Q must be constructed such that $\partial Q/\partial \kappa = \mathrm{tr}(\partial Q/\partial \boldsymbol{\sigma})$, which may not necessarily be equal to $\partial F/\partial \kappa$. Thus, we generally have a non-associative hardening even we have an associative flow rule.

More specifically, we formulate yield and plastic potential functions appropriate for granular rocks in terms of translated principal stresses

$$\overline{\sigma}_1 = \sigma_1 - a, \qquad \overline{\sigma}_2 = \sigma_2 - a, \qquad \overline{\sigma}_3 = \sigma_3 - a, \tag{5.5}$$

where $a > 0$ is a stress offset along the hydrostatic axis accommodating the material's cohesion ($a = 0$ for cohesionless materials). The corresponding invariants of the translated principal stresses are

$$\bar{I}_1 = \overline{\sigma}_1 + \overline{\sigma}_2 + \overline{\sigma}_3, \qquad \bar{I}_2 = \overline{\sigma}_1\overline{\sigma}_2 + \overline{\sigma}_2\overline{\sigma}_3 + \overline{\sigma}_1\overline{\sigma}_3, \qquad \bar{I}_3 = \overline{\sigma}_1\overline{\sigma}_2\overline{\sigma}_3. \tag{5.6}$$

For future use we also define dimensionless invariant functions

$$f_1 = \frac{\bar{I}_1^2}{\bar{I}_2}, \qquad f_2 = \frac{\bar{I}_1\bar{I}_2}{\bar{I}_3}, \qquad f_3 = \frac{\bar{I}_1^3}{\bar{I}_3}. \tag{5.7}$$

Our goal is to formulate plasticity models in terms of the above stress invariant functions.

The family of yield surfaces of interest is of the form

$$F = -f^{\mu}\bar{I}_1 + \kappa = 0, \quad f = -c_0 + c_1 f_1 + c_2 f_2 + c_3 f_3 > 0, \tag{5.8}$$

where c_0, c_1, c_2 and c_3 are positive dimensionless coefficients denoting the relative contributions of the second and third stress invariants on the shape of the yield surface on the deviatoric plane; $\mu > 0$ is a material parameter characterizing the shape of the yield surface on meridian planes, and κ is the same stress-like plastic internal variable described earlier. Our region of interest lies in the negative octant of the principal stress space.

Several classes of yield functions may be recovered depending on the values of the material parameters. If $c_0 = 3c_1 + 9c_2 + 27c_3$, the yield functions can only intersect the hydrostatic axis at $\bar{I}_1 = 0$, and they all open up toward the negative hydrostatic axis to form cones. On the deviatoric plane the cross-sectional shapes of these cones depend on the parameters c_1, c_2, and c_3. If $c_2 = c_3 = 0$ the cross-section is circular; if any of the coefficients c_2 or c_3 is non-zero, then the stress invariant will destroy the circular cross-sectional shape of the yield function on the deviatoric plane. More specifically, if $\mu \to \infty$, then we can rewrite the yield function as $f(-\bar{I}_1)^{1/\mu} = (-\kappa)^{1/\mu}$, which approaches $f = 1$ as $\mu \to \infty$; and if $c_0 = 3c_1 + 9c_2 + 27c_3$ then we recover the following conical yield surfaces associated with the following plasticity models (see Figs. 7–9): (a) Drucker–Prager [60] if $c_1 \neq 0$ and $c_2 = c_3 = 0$; (b) Matsuoka–Nakai [42] if $c_2 \neq 0$ and $c_1 = c_3 = 0$; and (c) Lade–Duncan [37] if $c_3 \neq 0$ and $c_1 = c_2 = 0$. For a finite positive value of μ and for the same expression for c_0 the cones flatten out with increasing confining pressures, implying a decreasing effective friction angle as shown in Figs. 10 and 11 (see also [38]).

If $c_0 < 3c_1 + 9c_2 + 27c_3$, the yield function forms a cap on the compression side allowing the yield surface to close in toward the hydrostatic axis and intersect this axis at a second point given by the coordinate $\bar{I}_1 = \kappa/(3c_1 + 9c_2 + 27c_3 - c_0)^{\mu}$. If $c_0 = 0$ the yield surface resembles a teardrop on the meridian plane. If $c_1 > 0$ and $c_2 = c_3 = 0$, then we a have a two-invariant yield function and the yield criterion predicts the same yield stresses in tension and compression. If $c_3 > 0$ and $c_1 = c_2 = 0$, then the yield surface depends on all three stress invariants and its shape resembles an asymmetric teardrop on the meridian plane, with a higher yield stress predicted in compression than in tension. If $c_2 > 0$ and $c_1 = c_3 = 0$, the yield surface also depends on all three stress invariants but its curvature is sharper on the compression side and milder on the tension side, see Figs. 12–14. In this paper we shall consider only the case where c_0 is either zero (family of teardrop-shaped yield surfaces), or equal to $3c_1 + 9c_2 + 27c_3$ (family of conical shaped yield surfaces).

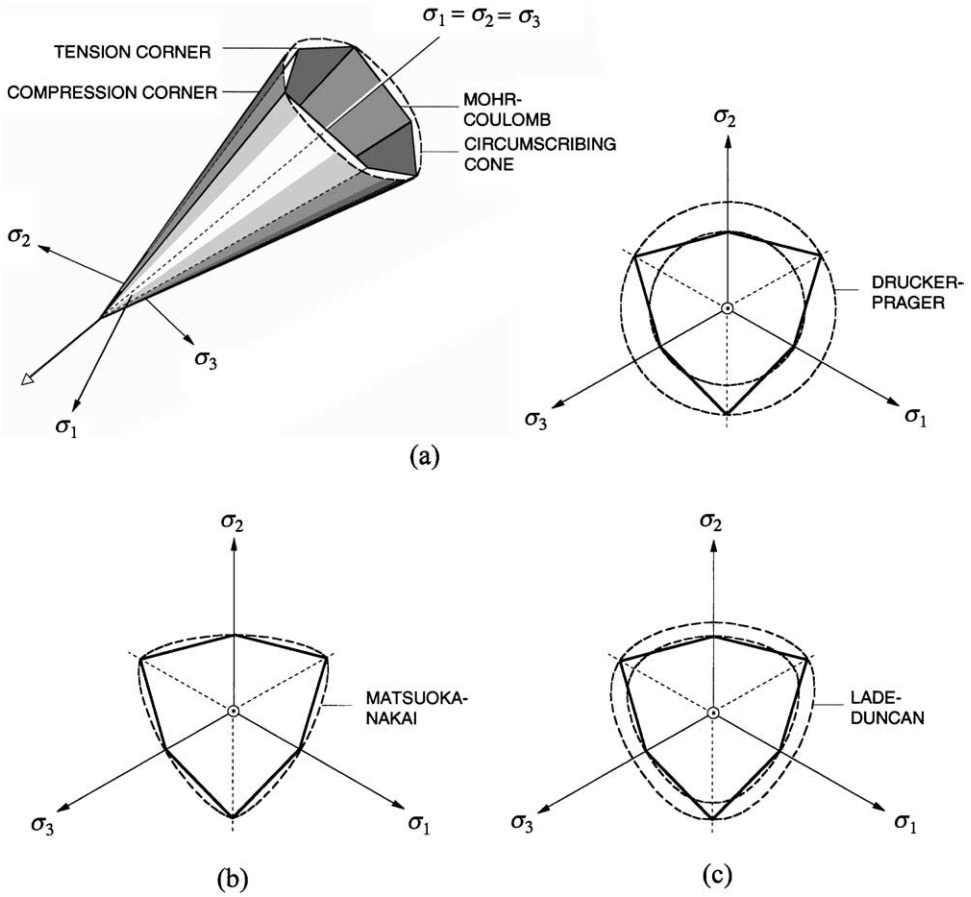

Fig. 7. Family of conical yield surfaces relative to the Mohr–Coulomb yield surface.

The family of plastic potential functions of interest is of the form

$$Q = -q^{v}\bar{I}_1 - Q_1(\kappa), \quad q = -\bar{c}_0 + c_1 f_1 + c_2 f_2 + c_3 f_3 > 0, \tag{5.9}$$

where c_1, c_2 and c_3 are the same positive dimensionless coefficients defined earlier, and $\bar{c}_0$ is an additional parameter that is now allowed to vary in the range $0 \leqslant \bar{c}_0 \leqslant 3c_1 + 9c_2 + 27c_3$. Along with a new exponent v, the parameters of the above family of plastic potential functions alter the meridional shape from that of a teardrop to that of an 'asymmetric cigar,' see Lade and Kim [36,39,40] who calibrated this plastic potential function to capture the plastic flow behavior of geomaterials. Fig. 15 shows such a plastic potential function in principal stress space.

The free energy function of interest is quadratic in the elastic strains and takes the form

$$\Psi = \Psi_0 + \tfrac{1}{2}\boldsymbol{\epsilon}^{e} : \boldsymbol{c}^{e} : \boldsymbol{\epsilon}^{e} + \Psi_1(v^{p}), \tag{5.10}$$

where Ψ_0 is a constant and $\boldsymbol{c}^{e}$ is the fourth-order elasticity tensor given in (3.12). The generalized Hooke's law is easily recovered from (5.2) as

$$\boldsymbol{\sigma} = \frac{\partial \Psi(\boldsymbol{\epsilon}^{e}, v^{p})}{\partial \boldsymbol{\epsilon}^{e}} = \boldsymbol{c}^{e} : \boldsymbol{\epsilon}^{e}. \tag{5.11}$$

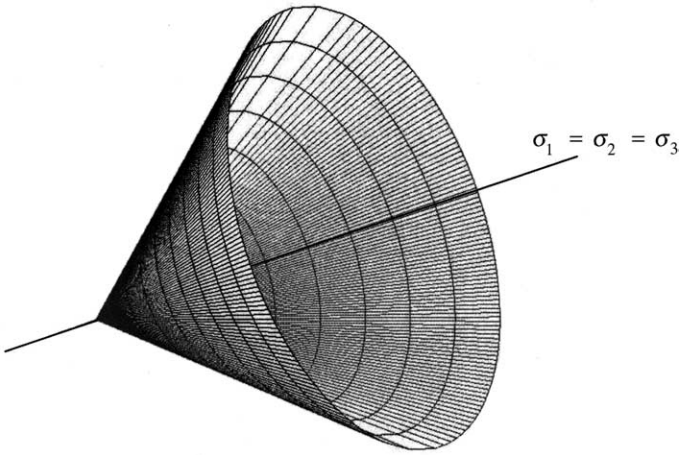

Fig. 8. Drucker–Prager yield surface.

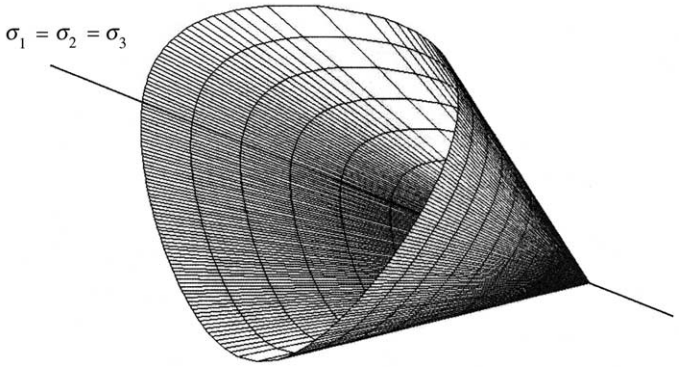

Fig. 9. Matsuoka–Nakai yield surface.

For a plastic process the consistency condition writes

$$\dot{F} = \frac{\partial F}{\partial \boldsymbol{\sigma}} : \dot{\boldsymbol{\sigma}} - \dot{\lambda} H = 0, \tag{5.12}$$

where H is the plastic modulus that takes the form

$$H = \frac{\partial F}{\partial \kappa} \Psi_1''(v^{\mathrm{p}}) \frac{\partial Q}{\partial \kappa}. \tag{5.13}$$

We recall that a hardening or softening response depends on the sign of the plastic modulus: hardening if the sign is positive, softening if the sign is negative, and perfect plasticity if the plastic modulus is equal to zero.

Now, let us consider the following hardening/softening law

$$\kappa = a_1 v^{\mathrm{p}} \exp(a_2 v^{\mathrm{p}}), \tag{5.14}$$

where a_1 and a_2 are positive scalar coefficients. The above equation postulates a dependence of κ on its strain-like conjugate plastic internal variable v^{p}. Differentiating (5.2) and (5.14) with time gives

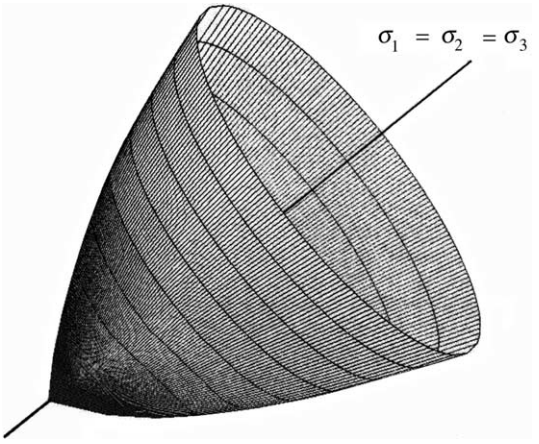

Fig. 10. Enhanced Drucker–Prager yield surface.

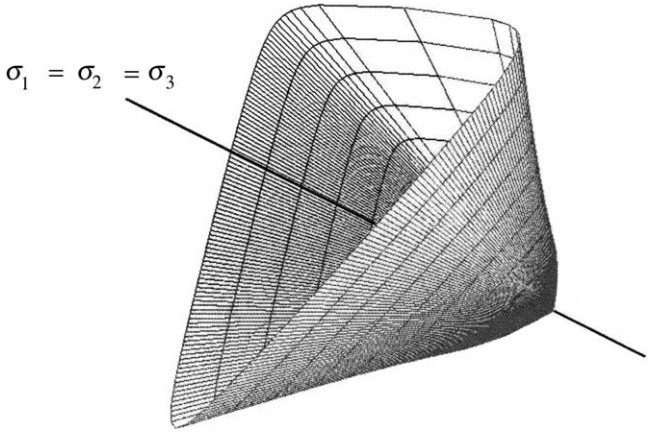

Fig. 11. Enhanced Matsuoka–Nakai yield surface. Note: friction angle is increased to exaggerate the triangular cross-sectional shape.

$$\Psi_1''(v^{\mathrm{p}}) = \kappa'(v^{\mathrm{p}}) = a_1(1 + a_2 v^{\mathrm{p}}) \exp(a_2 v^{\mathrm{p}}). \tag{5.15}$$

Thus, the plastic modulus H takes the more explicit form

$$H = -a_1(1 + a_2 v^{\mathrm{p}}) \exp(a_2 v^{\mathrm{p}}) \sum_{A=1}^{3} \frac{\partial Q}{\partial \sigma_A}. \tag{5.16}$$

If $\sum_{A=1}^{3} \partial Q/\partial \sigma_A < 0$ so that the plastic volumetric strain increment is compactive, then H is positive for small values of v^{p} and negative for negatively large values of v^{p}. The transition point at which the plastic modulus changes in sign is obtained by setting

$$1 + a_2 v^{\mathrm{p}} = 0 \Rightarrow v^{\mathrm{p}} = -1/a_2. \tag{5.17}$$

Thus, the parameter a_2 has the physical significance that the negative of its reciprocal is the critical value of the plastic volumetric strain v^{p} at which the plastic modulus H changes in sign.

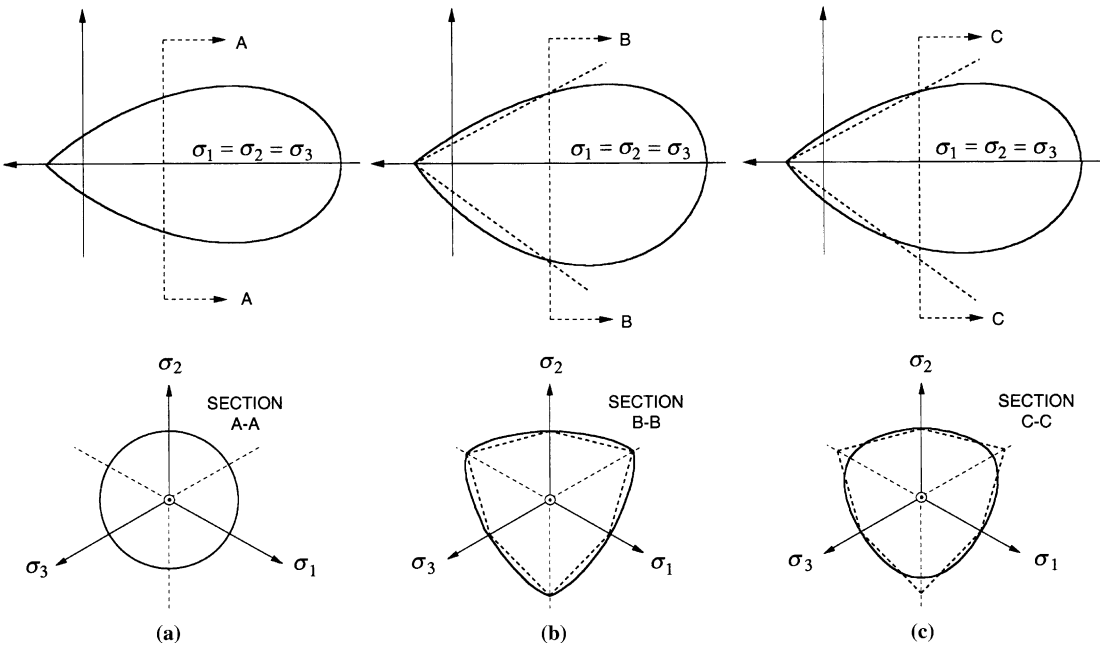

Fig. 12. Family of teardrop-shaped yield surfaces: (a) two-invariant; (b) three-invariant based on $I_1 I_2 / I_3$ invariant function; (c) three-invariant based on I_1^3 / I_3 invariant function.

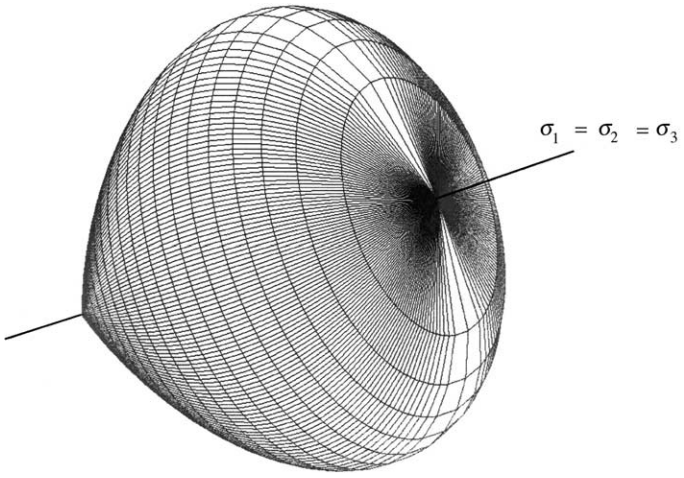

Fig. 13. Two-invariant teardrop-shaped yield surface.

To illustrate the physical significance of the internal variable κ, we consider the intersection of the yield surface with the hydrostatic axis at the point $\overline{\sigma}_1 = \overline{\sigma}_2 = \overline{\sigma}_3 < 0$. Here, we have $f_1 = 3$, $f_2 = 9$, and $f_3 = 27$. Thus, along the hydrostatic axis $\kappa = (3c_1 + 9c_2 + 27c_3)^{\mu} \bar{I}_1$, and κ is linearly proportional to $\bar{I}_1$. Equivalently, the value of $\bar{I}_1$ at the intersection point of the yield surface with the hydrostatic axis is $\bar{I}_1 = (3c_1 + 9c_2 + 27c_3)^{-\mu} \kappa$, which is the geometrical distance between the two extreme points on the yield

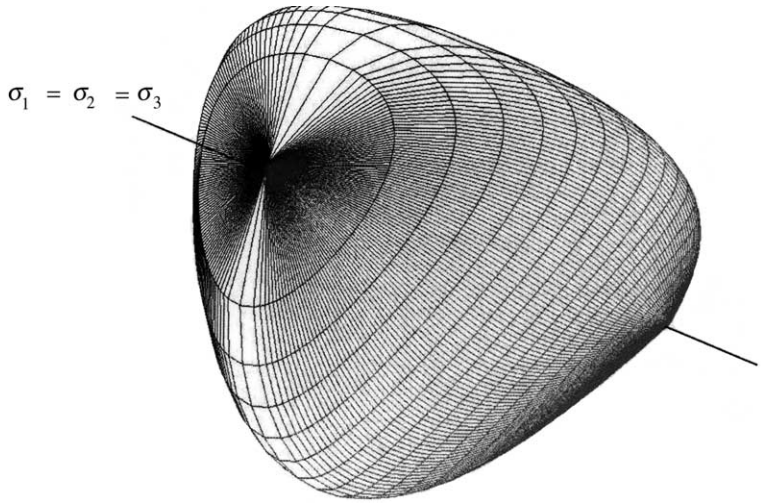

Fig. 14. Three-invariant teardrop-shaped yield surface.

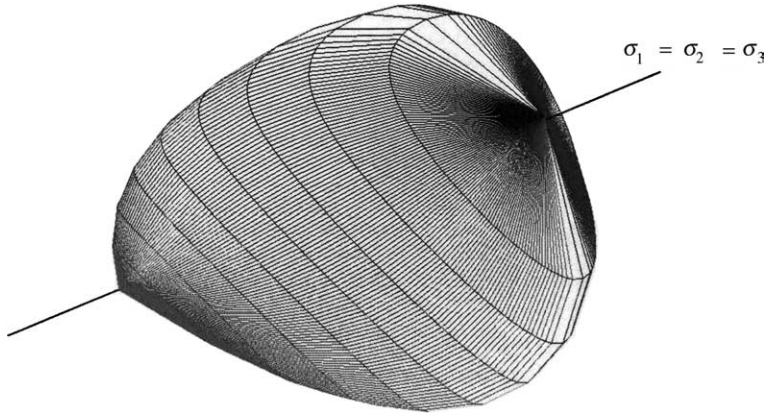

Fig. 15. Three-invariant teardrop-shaped plastic potential surface.

surface intersected by the hydrostatic axis. This makes κ and v^p legitimate conjugate plastic internal variables.

Remark 4. The dissipation inequality

$$\mathscr{D} = \boldsymbol{\sigma} : \dot{\boldsymbol{\epsilon}} - \frac{\mathrm{d}}{\mathrm{d}t}\Psi(\boldsymbol{\epsilon}^e, v^p) \geqslant 0$$

results in the constitutive equation $\boldsymbol{\sigma} = \partial\Psi/\partial\boldsymbol{\epsilon}^e$, along with the reduced dissipation inequality

$$\mathscr{D} = \boldsymbol{\sigma} : \dot{\boldsymbol{\epsilon}}^p + \kappa\dot{v}^p \geqslant 0,$$

where $\kappa = -\partial\Psi/\partial v^p$. Thus, the plastic dissipation must be non-negative, $\boldsymbol{\sigma} : \dot{\boldsymbol{\epsilon}}^p \geqslant 0$, which may not be satisfied if the stresses are translated excessively on the hydrostatic axis that the stress origin now lies

outside the yield function. Throughout this paper we assume that this scenario will not occur and that the yield surface is always large enough to enclose the stress space origin at all times.

5.2. Stress-point integration algorithm

For the family of three-invariant elastoplastic constitutive models described above the numerical integration is carried out using a recently published return mapping algorithm in principal stress axes [45,46,48]. This algorithm relies on a decomposition of the stress increment into material and spin parts, and has a structure that allows an explicit evaluation of the so-called algorithmic tangent operator. Essential aspects of the algorithm are briefly described below.

Integrating (5.3) using the implicit backward scheme, the evolution of the elastic strain tensor over a finite load increment takes the form

$$\boldsymbol{\epsilon}^{\mathrm{e}}_{n+1} = \boldsymbol{\epsilon}^{\mathrm{e\,tr}} - \Delta\lambda \left.\frac{\partial Q}{\partial \boldsymbol{\sigma}}\right|_{n+1}, \quad \boldsymbol{\epsilon}^{\mathrm{e\,tr}} = \boldsymbol{\epsilon}_n + \Delta\boldsymbol{\epsilon}, \tag{5.18}$$

where $\boldsymbol{\epsilon}_n$ is the converged elastic strain tensor of the previous load step, and $\Delta\boldsymbol{\epsilon}$ is the imposed strain increment. Since $\boldsymbol{\epsilon}^{\mathrm{e}}$ and $\partial Q/\partial\boldsymbol{\sigma}$ have the same spectral directions, it follows that the spectral directions of the trial elastic strain tensor $\boldsymbol{\epsilon}^{\mathrm{e\,tr}}$ are the same as those of $\boldsymbol{\epsilon}^{\mathrm{e}}$ and $\partial Q/\partial\boldsymbol{\sigma}$. Thus the return mapping equation may be carried out in the principal axes as

$$\sum_{B=1}^{3} a^{\mathrm{e}-1}_{AB}\sigma_B - \epsilon^{\mathrm{e\,tr}}_A + \Delta\lambda\frac{\partial Q}{\partial\sigma_A} = 0, \tag{5.19}$$

where $[a^{\mathrm{e}}_{AB}]$ is the matrix of elastic moduli having a structure given in (3.43). For a general 3D case, we thus have a set of three equations in the unknowns σ_1, σ_2, σ_3 and the discrete plastic multiplier $\Delta\lambda$.

A fourth equation is provided by the discrete hardening/softening law, herein written in residual form as

$$\kappa - \hat{\kappa} = 0, \quad \hat{\kappa} = a_1 v^{\mathrm{p}}\exp(a_2 v^{\mathrm{p}}), \quad v^{\mathrm{p}} = v^{\mathrm{p}}_n + \Delta\lambda\sum_{A=1}^{3}\frac{\partial Q}{\partial\sigma_A}, \tag{5.20}$$

where v^{p}_n is the converged cumulative plastic strain of the previous step. The above equation introduces an additional unknown parameter κ. The fifth and final equation is the discrete consistency condition

$$F(\sigma_1,\sigma_2,\sigma_3,\kappa) = 0. \tag{5.21}$$

Eqs. (5.19)–(5.21) constitute a set of five non-linear equations in five unknowns that can be solved iteratively by Newton's method. The 'driving force' in the algorithm is the prescribed elastic strain increment $\Delta\boldsymbol{\epsilon}$ describing the evolution of deformation.

For future use let us write the gradient with respect to a principal stress σ_A of the plastic potential function Q,

$$\frac{\partial Q}{\partial\sigma_A} = -q^{\nu} - \bar{I}_1\nu q^{\nu-1}\frac{\partial q}{\partial\sigma_A}, \tag{5.22}$$

where

$$\frac{\partial q}{\partial\sigma_A} = c_1\frac{\partial f_1}{\partial\sigma_A} + c_2\frac{\partial f_2}{\partial\sigma_A} + c_3\frac{\partial f_3}{\partial\sigma_A}. \tag{5.23}$$

Closed-form expressions are available for the derivatives of the invariant functions. As for the stress invariants themselves, we easily verify

$$\frac{\partial \bar{I}_1}{\partial \sigma_A} = 1, \qquad \frac{\partial \bar{I}_2}{\partial \sigma_A} = \bar{I}_1 - \overline{\sigma}_A, \qquad \frac{\partial \bar{I}_3}{\partial \sigma_A} = \bar{I}_3/\overline{\sigma}_A, \tag{5.24}$$

Simple calculus thus gives

$$\frac{\partial f_1}{\partial \sigma_A} = 2\frac{\bar{I}_1}{\bar{I}_2} - \frac{\bar{I}_1^2(\bar{I}_1 - \overline{\sigma}_A)}{\bar{I}_2^2}, \tag{5.25a}$$

$$\frac{\partial f_2}{\partial \sigma_A} = \frac{\bar{I}_1(\bar{I}_1 - \overline{\sigma}_A)}{\bar{I}_3} + \frac{\bar{I}_2}{\bar{I}_3} - \frac{\bar{I}_1\bar{I}_2}{\bar{I}_3\overline{\sigma}_A}, \tag{5.25b}$$

and

$$\frac{\partial f_3}{\partial \sigma_A} = 3\frac{\bar{I}_1^2}{\bar{I}_3} - \frac{\bar{I}_1^3}{\bar{I}_3\overline{\sigma}_A}. \tag{5.25c}$$

5.3. Algorithmic tangent operators

There are two algorithmic tangent operators of interest. The first is used in the *local* Newton iteration to calculate the five unknowns σ_1, σ_2, σ_3, κ, and $\Delta\lambda$. In the context of non-linear finite element analysis, this tangent operator, or local Jacobian, is used at the Gauss point level by the material subroutine to advance the solution to the next state for a given strain increment $\Delta\boldsymbol{\epsilon}$. The second tangent operator is the algorithmic moduli tensor $\boldsymbol{c} = \partial\boldsymbol{\sigma}/\partial\boldsymbol{\epsilon}$ supplied by the material subroutine to the calling element subroutine at the conclusion of the local iteration. This latter information is used by the finite element program to construct the overall tangent operator for *global* Newton iteration. There exists a closed-form relationship between these two tangent operators as demonstrated in this section. Furthermore, the algorithmic moduli tensor $\boldsymbol{c}$ may be used to approximate the material constitutive operator for analysis of the onset of deformation bands.

The first algorithmic tangent operator is determined considering (5.18)–(5.20) as a set of residual equations $\boldsymbol{r}(\boldsymbol{x})$ in the unknown variables $\boldsymbol{x} = \{\sigma_1, \sigma_2, \sigma_3, \kappa, \Delta\lambda\}$. The local Jacobian, $\boldsymbol{K} = \boldsymbol{r}'(\boldsymbol{x})$, has components

$$K_{AB} = \begin{bmatrix} \left(a_{AB}^{\mathrm{e}-1} + \Delta\lambda Q_{,\sigma_A\sigma_B}\right) & 0 & Q_{,\sigma_A} \\ -\hat{\kappa}_{,\sigma_B} & 1 & -\hat{\kappa}_{,\Delta\lambda} \\ F_{,\sigma_B} & F_{,\kappa} & 0 \end{bmatrix}_{5\times 5}. \tag{5.26}$$

The second algorithmic tangent operator has the form [45,46]

$$\boldsymbol{c} = \sum_{A=1}^{3}\sum_{B=1}^{3} a_{AB}\boldsymbol{m}^{(A)} \otimes \boldsymbol{m}^{(B)} + \frac{1}{2}\sum_{A=1}^{3}\sum_{B\neq A}\left(\frac{\sigma_B - \sigma_A}{\epsilon_B^{\mathrm{e\,tr}} - \epsilon_A^{\mathrm{e\,tr}}}\right)\left(\boldsymbol{m}^{(AB)} \otimes \boldsymbol{m}^{(AB)} + \boldsymbol{m}^{(AB)} \otimes \boldsymbol{m}^{(BA)}\right), \tag{5.27}$$

where

$$\boldsymbol{m}^{(A)} = \boldsymbol{n}^{(A)} \otimes \boldsymbol{n}^{(A)}, \qquad \boldsymbol{m}^{(AB)} = \boldsymbol{n}^{(A)} \otimes \boldsymbol{n}^{(B)}, \quad A \neq B. \tag{5.28}$$

The first term of (5.27) represents the linearization of the algorithm in the principal stress axes and reflects the features of the constitutive model as well as the return mapping algorithm in principal axes, whereas the second term reflects the spin of the spectral directions. The algorithmic matrix $[a_{AB}]$ is related to the local Jacobian $\boldsymbol{K}$ by the simple matrix relation

$$[a_{AB}] = [\boldsymbol{I}_{3\times 3}|\boldsymbol{0}] \cdot \boldsymbol{K}^{-1} \cdot \left[\frac{\boldsymbol{I}_{3\times 3}}{\boldsymbol{0}}\right], \tag{5.29}$$

where $\boldsymbol{I}_{3\times3}$ is a 3×3 identity matrix, i.e., $[a_{AB}]$ is the upper left-hand 3×3 submatrix of $\boldsymbol{K}^{-1}$, see [45,46] for further details.

The proposed return mapping algorithm provides closed-form expressions for all the required derivatives, including the second derivatives of the function Q. The latter write

$$\frac{\partial^2 Q}{\partial\sigma_A\,\partial\sigma_B} = -vq^{v-1}\left(\frac{\partial q}{\partial\sigma_A}+\frac{\partial q}{\partial\sigma_B}\right) - \bar{I}_1 v(v-1)q^{v-2}\frac{\partial q}{\partial\sigma_A}\frac{\partial q}{\partial\sigma_B} - \bar{I}_1 vq^{v-1}\frac{\partial^2 q}{\partial\sigma_A\,\partial\sigma_B}, \tag{5.30}$$

where

$$\frac{\partial^2 q}{\partial\sigma_A\,\partial\sigma_B} = c_1\frac{\partial^2 f_1}{\partial\sigma_A\,\partial\sigma_B} + c_2\frac{\partial^2 f_2}{\partial\sigma_A\,\partial\sigma_B} + c_3\frac{\partial^2 f_3}{\partial\sigma_A\,\partial\sigma_B}. \tag{5.31}$$

The second derivatives of the invariant functions are given by

$$\frac{\partial^2 f_1}{\partial\sigma_A\,\partial\sigma_B} = \left(\frac{2}{\bar{I}_2} - 5\frac{\bar{I}_1^2}{\bar{I}_2^2} + 2\frac{\bar{I}_1^4}{\bar{I}_2^3}\right) + \left(\frac{\bar{I}_1^2}{\bar{I}_2^2}\right)\delta_{AB} + 2\left(\frac{\bar{I}_1}{\bar{I}_2^2} - \frac{\bar{I}_1^3}{\bar{I}_2^3}\right)(\overline{\sigma}_A+\overline{\sigma}_B) + 2\left(\frac{\bar{I}_1^2}{\bar{I}_2^3}\right)\overline{\sigma}_A\overline{\sigma}_B; \tag{5.32a}$$

$$\begin{aligned}\frac{\partial^2 f_2}{\partial\sigma_A\,\partial\sigma_B} &= 3\left(\frac{\bar{I}_1}{\bar{I}_3}\right) - \left(\frac{\bar{I}_1}{\bar{I}_3}\right)\delta_{AB} - \frac{1}{\bar{I}_3}(\overline{\sigma}_A+\overline{\sigma}_B) + \frac{\bar{I}_1}{\bar{I}_3}\left(\frac{\overline{\sigma}_A}{\overline{\sigma}_B}+\frac{\overline{\sigma}_B}{\overline{\sigma}_A}\right) - \frac{\bar{I}_1^2+\bar{I}_2}{\bar{I}_3}\left(\frac{1}{\overline{\sigma}_A}+\frac{1}{\overline{\sigma}_B}\right)\\ &\quad + 2\left(\frac{\bar{I}_1\bar{I}_2}{\bar{I}_3}\right)\frac{1}{\overline{\sigma}_A\overline{\sigma}_B} - \left(\frac{\bar{I}_1\bar{I}_2}{\bar{I}_3^2}\right)\Sigma_{AB};\end{aligned} \tag{5.32b}$$

and

$$\frac{\partial^2 f_3}{\partial\sigma_A\,\partial\sigma_B} = 6\left(\frac{\bar{I}_1}{\bar{I}_3}\right) - 3\left(\frac{\bar{I}_1^2}{\bar{I}_3}\right)\left(\frac{1}{\overline{\sigma}_A}+\frac{1}{\overline{\sigma}_B}\right) + 2\left(\frac{\bar{I}_1^3}{\bar{I}_3}\right)\frac{1}{\overline{\sigma}_A\overline{\sigma}_B} - \left(\frac{\bar{I}_1^3}{\bar{I}_3^2}\right)\Sigma_{AB}, \tag{5.32c}$$

where $\Sigma_{AB} = \partial^2\bar{I}_3/\partial\sigma_A\,\partial\sigma_B = (\bar{I}_3/\overline{\sigma}_A\overline{\sigma}_B)(1-\delta_{AB})$ is a symmetric matrix with components

$$[\Sigma_{AB}] = \begin{bmatrix} 0 & \overline{\sigma}_3 & \overline{\sigma}_2 \\ \overline{\sigma}_3 & 0 & \overline{\sigma}_1 \\ \overline{\sigma}_2 & \overline{\sigma}_1 & 0 \end{bmatrix}.$$

The remaining derivatives take the form

$$\frac{\partial\hat{\kappa}}{\partial\sigma_B} = \Delta\lambda\overline{\kappa}\sum_{A=1}^{3}\frac{\partial^2 Q}{\partial\sigma_A\,\partial\sigma_B}, \qquad \frac{\partial\hat{\kappa}}{\partial\,\Delta\lambda} = \overline{\kappa}\sum_{A=1}^{3}\frac{\partial Q}{\partial\sigma_A},$$

where

$$\overline{\kappa} = a_1(1+a_2 v^{\mathrm{p}})\exp(a_2 v^{\mathrm{p}}), \quad v^{\mathrm{p}} = v_n^{\mathrm{p}} + \Delta\lambda\sum_{A=1}^{3}\frac{\partial Q}{\partial\sigma_A}. \tag{5.33}$$

Finally,

$$\frac{\partial F}{\partial\sigma_A} = -f^{\mu} - \bar{I}_1\mu f^{\mu-1}\frac{\partial f}{\partial\sigma_A}, \quad \frac{\partial f}{\partial\sigma_A} = \frac{\partial q}{\partial\sigma_A}, \quad \frac{\partial F}{\partial\kappa} = 1. \tag{5.34}$$

Note that if we had worked in the general six-dimensional stress space, these derivatives would not have been so easy to get.

Remark 5. Comparing (3.44) and (5.27), we observe that if $a_{AB} \approx a_{AB}^{\mathrm{ep}}$ and $\epsilon_A^{\mathrm{e\,tr}} \approx \epsilon_A^{\mathrm{e}}$, then $\boldsymbol{c} \approx \boldsymbol{c}^{\mathrm{ep}}$. Note that $\epsilon_A^{\mathrm{e\,tr}} \approx \epsilon_A^{\mathrm{e}}$ if the *incremental* plastic strain $\Delta\boldsymbol{\epsilon}^{\mathrm{p}}$ is small compared to the *cumulative* elastic strain $\boldsymbol{\epsilon}^{\mathrm{e}}$. As shown in [61], under certain circumstances the algorithmic tangent operator may be used equally well to

approximate the elastoplastic constitutive operator for localization analysis. This could have some implementational advantages since the former tensor is already available from the Newton–Raphson iteration, and thus, need not be re-calculated.

6. Summary

We have presented a geological and mathematical framework for characterizing the appearance of various types of deformation bands in granular materials. The three end members are simple shear, pure compaction, and pure dilation bands. Field data indicate that although these three members do occur in nature, mixed-mode localization structures such as compactive shear bands and dilatant shear bands are the most common modes of localized failure. All of these modes are captured mathematically by classical bifurcation theory for planar bands in conjunction with non-linear continuum mechanics and theoretical/computational plasticity.

The plasticity models must be sufficiently robust to handle mixed-mode (volumetric and deviatoric) yielding and plastic flow. To this end, we have presented a family of three-invariant elastoplastic constitutive models capable of representing mixed-mode yielding and plastic flow in geomaterials. The models include some of the well known conical shaped yield surfaces for dilatant frictional materials, as well as a family of well calibrated teardrop-shaped yield surfaces capable of representing compactive volumetric yielding and plastic flow occurring in geomaterials at high confining pressures. To advance the solution at finite load increments, we have used a recently developed return mapping algorithm in principal axes applicable to any three-invariant plasticity models. Tangent operators are all expressed in spectral form suitable for efficient numerical implementation.

The resulting mathematical formulation is complete in the sense that the end result clearly defines the orientation and type of deformation bands expected to form at localization. In a companion paper we present some results of numerical simulations aimed at capturing the three extreme failure modes, namely, pure compaction, simple shear, and pure dilation bands, as well as the combination modes described in the geologic framework.

Acknowledgements

We are grateful to graduate student José Andrade for rendering the three-dimensional solid figures presented in this paper, as well as to the two anonymous reviewers for their constructive reviews. The first author acknowledges the support of National Science Foundation grant nos. CMS-97-00426 and CMS-03-24674, and US Department of Energy grant no. DE-FG02-03ER15454; the second author acknowledges the support of US Department of Energy grant no. DE-FG03-94ER14462 and National Science Foundation grant no. EAR-02-29862.

References

[1] M.A. Antonellini, A. Aydin, Effect of faulting on fluid flow in porous sandstones: petrophysical properties, Am. Assoc. Petrol. Geol. Bull. 78 (1994) 355–377.
[2] M.A. Antonellini, A. Aydin, D.D. Pollard, Microstructure of deformation bands in porous sandstones at Arches National Park, Utah, J. Struct. Geol. 16 (1994) 941–959.
[3] M.A. Antonellini, A. Aydin, D.D. Pollard, P. D'Onfro, Petrophysical study of faults in sandstones using petrographic image analysis and X-ray computerized tomography, Pure Appl. Geophys. 143 (1994) 181–201.
[4] A. Aydin, Small faults formed as deformation bands in sandstone, Pure Appl. Geophys. 116 (1978) 913–930.

[5] A. Aydin, A.M. Johnson, Analysis of faulting in porous sandstones, J. Struct. Geol. 5 (1983) 19–31.
[6] M. Cakir, A. Aydin, Tectonics and fracture characteristics of the northern Lake Mead, SE Nevada, in: Proceedings of the Stanford Rock Fracture Project Field Workshop, 1994, 34 p.
[7] X. Du Bernard, P. Eichhubl, A. Aydin, Dilation bands: a new form of localized failure in granular media, Geophys. Res. Lett. 29 (4) (2002) 2176, doi:10.1029/2002GL015966.
[8] D.E. Dunn, L.J. Lafountain, R.E. Jackson, Porosity dependence and mechanism of brittle faulting in sandstone, J. Geophys. Res. 78 (1973) 2303–2317.
[9] S. Cashman, K. Cashman, Cataclasis and deformation-band formation in unconsolidated marine terrace sand, Humboldt County, California, Geology 2 (2000) 111–114.
[10] T. Engelder, Cataclasis and the generation of fault gouge, Geol. Soc. Am. Bull. 85 (1974) 1515–1522.
[11] H. Fossen, J. Hesthammer, Geometric analysis and scaling relations of deformation bands in porous sandstone, J. Struct. Geol. 19 (1997) 1479–1493.
[12] R.E. Hill, Analysis of deformation bands in the Aztec Sandstone, Valley of Fire, Nevada, M.S. Thesis, Geosciences Department, University of Nevada, Las Vegas, 1989, 68 p.
[13] K.A. Issen, J.W. Rudnicki, Conditions for compaction bands in porous rock, J. Geophys. Res. 105 (2000) 21529–21536.
[14] K. Mair, I. Main, S. Elphick, Sequential growth of deformation bands in the laboratory, J. Struct. Geol. 22 (2000) 25–42.
[15] P.N. Mollema, M.A. Antonellini, Compaction bands: a structural analog for anti-mode I cracks in aeolian sandstone, Tectonophysics 267 (1996) 209–228.
[16] W.A. Olsson, Theoretical and experimental investigation of compaction bands in porous rock, J. Geophys. Res. 104 (1999) 7219–7228.
[17] K. Sternlof, D.D. Pollard, Numerical modelling of compactive deformation bands as granular anti-cracks, Eos. Trans. AGU 83(47) F1347, 2002.
[18] T.-F. Wong, P. Baud, E. Klein, Localized failure modes in a compactant porous rock, Geophys. Res. Lett. 28 (2001) 2521–2524.
[19] J. Hadamard, Lecons sur la Propagation des Ondes, Herman et fil, Paris, 1903.
[20] R. Hill, A general theory of uniqueness and stability in elastic–plastic solids, J. Mech. Phys. Solids 6 (1958) 236–249.
[21] T.Y. Thomas, Plastic Flow and Fracture of Solids, Academic Press, New York, 1961.
[22] J. Mandel, Conditions de stabilité et postulat de Drucker, in: Proceedings IUTAM Symposium on Rheology and Soil Mechanics, Springer-Verlag, Berlin, 1966, pp. 58–68.
[23] J.R. Rice, The localization of plastic deformation, in: W.T. Koiter (Ed.), Theoretical and Applied Mechanics, North-Holland Publishing Co., The Netherlands, 1976, pp. 207–220.
[24] J.W. Rudnicki, J.R. Rice, Conditions for the localization of deformation in pressure-sensitive dilatant materials, J. Mech. Phys. Solids 23 (1975) 371–394.
[25] J.R. Rice, J.W. Rudnicki, A note on some features of the theory of localization of deformation, Int. J. Solids Struct. 16 (1980) 597–605.
[26] R. Chambon, S. Crochepeyre, J. Desrues, Localization criteria for non-linear constitutive equations of geomaterials, Mech. Cohes.-Frict. Mater. 5 (2000) 61–82.
[27] J. Desrues, R. Chambon, Shear band analysis for granular materials: the question of incremental non-linearity, Ing. Arch. 59 (1989) 187–196.
[28] I. Herle, D. Kolymbas, Pressure- and density-dependent bifurcation of soils, in: H.B. Mühlhaus, A.V. Dyskin, E. Pasternak (Eds.), Bifurcation and Localisation Theory in Geomechanics, A.A. Balkema Lisse, 2001, pp. 53–58.
[29] D. Kolymbas, Bifurcation analysis for sand samples with a non-linear constitutive equation, Ing. Arch. 50 (1981) 131–140.
[30] C. Tamagnini, G. Viggiani, R. Chambon, Some remarks on shear band analysis in hypoplasticity, in: H.B. Mühlhaus, A.V. Dyskin, E. Pasternak (Eds.), Bifurcation and Localisation Theory in Geomechanics, A.A. Balkema Lisse, 2001, pp. 85–93.
[31] J.W. Rudnicki, Conditions for compaction and shear bands in a transversely isotropic material, Int. J. Solids Struct. 39 (2002) 3741–3756.
[32] R.I. Borja, Bifurcation of elastoplastic solids to shear band mode at finite strain, Comput. Methods Appl. Mech. Engrg. 191 (2002) 5287–5314.
[33] J.H. Argyris, G. Faust, J. Szimmat, E.P. Warnke, K.J. Willam, Recent developments in the finite element analysis of prestressed concrete reactor vessel, Nucl. Engrg. Des. 28 (1974) 42–75.
[34] L.F. Boswell, Z. Chen, A general failure criterion for plain concrete, Int. J. Solids Struct. 23 (1987) 621–630.
[35] J. Jiang, S. Pietruszczak, Convexity of yield loci for pressure sensitive materials, Comput. Geotech. 5 (1988) 51–63.
[36] M.K. Kim, P.V. Lade, Single hardening constitutive model for frictional materials I. Plastic potential function, Comput. Geotech. 5 (1988) 307–324.
[37] P.V. Lade, J.M. Duncan, Elastoplastic stress–strain theory for cohesionless soil, J. Geotech. Engrg. Div., ASCE 101 (1975) 1037–1053.
[38] P.V. Lade, Elasto-plastic stress–strain theory for cohesionless soil with curved yield surfaces, Int. J. Solids Struct. 13 (1977) 1019–1035.

[39] P.V. Lade, M.K. Kim, Single hardening constitutive model for frictional materials II. Yield criterion and plastic work contours, Comput. Geotech. 6 (1988) 13–29.
[40] P.V. Lade, M.K. Kim, Single hardening constitutive model for frictional materials III. Comparisons with experimental data, Comput. Geotech. 6 (1988) 31–47.
[41] J. Maso, J. Lerau, Mechanical behavior of Darney sandstone in biaxial compression, Int. J. Rock Mech. Min. Sci. Geomech. Abstr. 17 (1980) 109–115.
[42] H. Matsuoka, T. Nakai, Stress-deformation and strength characteristics of soil under three different principal stresses, Proc. JSCE 232 (1974) 59–70.
[43] G. Schickert, H. Winkler, Results of tests concerning strength and strain of concrete subjected to multiaxial compressive stresses, Deutscher Ausschus Stahlbeton 277 (1977), Berlin.
[44] K.J. Willam, E.P. Warnke, Constitutive model for the triaxial behaviour of concrete, in: ISMES Seminar on Concrete Structures Subjected to Triaxial Stresses, Bergamo, Italy, 1975, pp. 1–30.
[45] R.I. Borja, Plasticity Modeling and Computation, Lecture Notes, Stanford University, California, 2001.
[46] R.I. Borja, K.M. Sama, P.F. Sanz, On the numerical integration of three-invariant elastoplastic constitutive models, Comput. Methods Appl. Mech. Engrg. 192 (2003) 1227–1258.
[47] R.I. Borja, Algorithm for a class of three-invariant elastoplastic constitutive models suitable for the analysis of deformation bands in geomaterials, in: E. Oñate, D.R.J. Owen (Eds.), Seventh International Conference on Computational Plasticity (COMPLAS 2003), CIMNE, Barcelona, 2003, in CD-ROM.
[48] C. Tamagnini, R. Castellanza, R. Nova, A generalized backward Euler algorithm for the numerical integration of an isotropic hardening elastoplastic model for mechanical and chemical degradation of bonded geomaterials, Int. J. Numer. Anal. Meth. Geomech. 26 (2002) 963–1004.
[49] J.C. Simo, J. Oliver, A new approach to the analysis and simulation of strain softening in solids, in: Z.P. Bazant, Z. Bittnar, M. Jirásek, J. Mazars (Eds.), Fracture and Damage in Quasibrittle Structures, E&FN Spon, 2–6 Boundary Row, London, 1994, pp. 25–39 (Chapter 3).
[50] J.C. Simo, T.J.R. Hughes, Computational Inelasticity, Springer, New York, 1998.
[51] E. Papamichos, I. Vardoulakis, Shear band formation in sand according to non-coaxial plasticity model, Géotechnique 45 (1991) 649–661.
[52] J. Sulem, I. Vardoulakis, E. Papamichos, A. Oulahna, J. Tronvoll, Elasto-plastic modelling of Red Wildmoor sandstone, Mech. Cohes.-Frict. Mater. 4 (1999) 215–245.
[53] R.W. Ogden, Nonlinear Elastic Deformations, Chichester, Ellis Horwood, 1984.
[54] D. Bigoni, Bifurcation and instability of non-associative elastoplastic solids, in: CISM Lecture Notes on the Course: Material Instabilities in Elastic and Plastic Solids, H. Petryk (Coordinator), Udine, September 13–17, 1999.
[55] R. Larsson, P. Steinmann, K. Runesson, Finite element embedded localization band for finite strain plasticity based on a regularized strong discontinuity, Mech. Cohes.-Frict. Mater. 4 (1998) 171–194.
[56] Y.F. Dafalias, Plastic spin: necessity or redundancy?, Int. J. Plast. 14 (1998) 909–931.
[57] R.I. Borja, Finite element simulation of strain localization with large deformation: capturing strong discontinuity using a Petrov–Galerkin multiscale formulation, Comput. Methods Appl. Mech. Engrg. 191 (2002) 2949–2978.
[58] E.H. Lee, Elastic–plastic deformation at finite strains, J. Appl. Mech. (1969) 1–6.
[59] J.C. Simo, Algorithms for static and dynamic multiplicative plasticity that preserve the classical return mapping schemes of the infinitesimal theory, Comput. Methods Appl. Mech. Engrg. 99 (1992) 61–112.
[60] D.C. Drucker, W. Prager, Soil mechanics and plastic analysis or limit design, Quart. Appl. Math. 10 (1952) 157–165.
[61] R.I. Borja, Computational modeling of deformation bands in granular media. II. Numerical simulations, Comput. Methods Appl. Mech. Engrg. 193 (2004) 2699–2718.