International Journal of Solids and Structures 48 (2011) 3001–3010

Contents lists available at ScienceDirect

International Journal of Solids and Structures

journal homepage: www.elsevier.com/locate/ijsolstr

A C2 continuous approximation to the Mohr–Coulomb yield surface

A.J. Abbo *, A.V. Lyamin, S.W. Sloan, J.P. Hambleton

Centre for Geotechnical and Materials Modelling, University of Newcastle, Callaghan, NSW 2308, Australia

ARTICLE INFO

Article history:
Received 27 January 2011
Received in revised form 14 June 2011
Available online 1 July 2011

Keywords:
Mohr–Coulomb
Elastoplasticity
Consistent tangent
Finite element

ABSTRACT

In spite of the development of more sophisticated constitutive models for soil, the Mohr–Coulomb yield criterion remains a popular choice for geotechnical analysis due to its simplicity and ease of use by practising engineers. The implementation of the criterion in finite element programs, however, presents some numerical difficulties due to the gradient discontinuities which occur at both the edges and the tip of the hexagonal yield surface pyramid. Furthermore, some implicit techniques utilising consistent tangent stiffness formulations are unable to achieve full quadratic convergence as the yield criteria is not C2 continuous. This paper extends the previous work of Abbo and Sloan (1995) through the introduction of C2 continuous rounding of the Mohr–Coulomb yield surface in the octahedral plane. This approximation, when combined with the hyperbolic approximation in the meridional plane (Abbo and Sloan, 1995), describes a yield surface that is C2 continuous at all stress states. The new smooth yield surface can be made to approximate the Mohr–Coulomb yield function as closely as required by adjusting only two parameters, and is suitable for consistent tangent stiffness formulations.

1. Introduction

The Mohr–Coulomb yield criterion provides a relatively simple model for simulating the plastic behaviour of soil. Other more sophisticated constitutive models for predicting the behaviour of soil have been developed over the past three decades, however the complexity of these models, as well as the additional testing required to determine the various soil parameters involved, minimises their utility for practising geotechnical engineers. The Mohr–Coulomb yield function is also of importance to finite element researchers and practitioners as it forms the basis of many analytical solutions. These analytical solutions serve as crucial benchmarks for validating numerical algorithms and software.

In three-dimensional principal stress space, the Mohr–Coulomb yield criterion is a hexagonal pyramid whose central axis lies along the hydrostatic axis as shown in Fig. 1(a). The implementation of the Mohr–Coulomb yield surface in finite element programs is complicated by the presence of the vertices at the tip and along the sides of the Mohr–Coulomb pyramid. At these vertices, the yield function is not C1 continuous (i.e., first derivative is not continuous) let alone C2 continuous (i.e, second-derivative is not continuous). It is necessary to address these singularities because stress states lying at, or near, the vertices are often encountered in practice. One approach to overcoming the computational difficulties posed by the vertices is to consider the Mohr–Coulomb surface as six separate planar yield surfaces and implement the constitutive law as a multi-surface yield function using the formulation of Koiter (1953) (e.g. Ristinmaa and Tryding, 1993; Clausen et al., 2006). The approach used in this paper is to derive a smooth approximation to the yield surface that eliminates the sharp vertices by rounding the corners of the Mohr–Coulomb yield surface. The rounding is derived so that it closely approximates the true yield surface yet provides C2 continuity.

Mathematically the Mohr–Coulomb yield criterion can be described in terms of the principal stresses ($\sigma_1 \geqslant \sigma_2 \geqslant \sigma_3$) as

$$F = (\sigma_1 - \sigma_3) + (\sigma_1 + \sigma_3)\sin\phi - 2c\cos\phi = 0 \quad (1)$$

in which c and ϕ represent the cohesion and friction angle of the soil and tensile stresses are considered positive. A more convenient form of the criterion, which avoids explicit calculation of principal stresses, was proposed by Nayak and Zienkiewicz (1972). They expressed the criterion as a function of the three stress invariants ($\sigma_m, \bar{\sigma}, \theta$) (see Appendix A) as

$$F = \sigma_m \sin\phi + \bar{\sigma}K(\theta) - c\cos\phi = 0 \quad (2)$$

in which

$$K(\theta) = \cos\theta - \frac{1}{\sqrt{3}}\sin\phi\sin\theta \quad (3)$$

is a function controlling the shape of the surface in the octahedral plane (the plane orthogonal to the hydrostatic axis).

The gradient discontinuities at the tip and along the sides of the hexagonal pyramid can be considered separately by studying the meridional and octahedral sections of the yield surface. The

* Corresponding author. Tel.: +61 2 49215582; fax: +61 2 49216991.
E-mail address: Andrew.Abbo@newcastle.edu.au (A.J. Abbo).

0020-7683/$ - see front matter Crown Copyright © 2011 Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijsolstr.2011.06.021

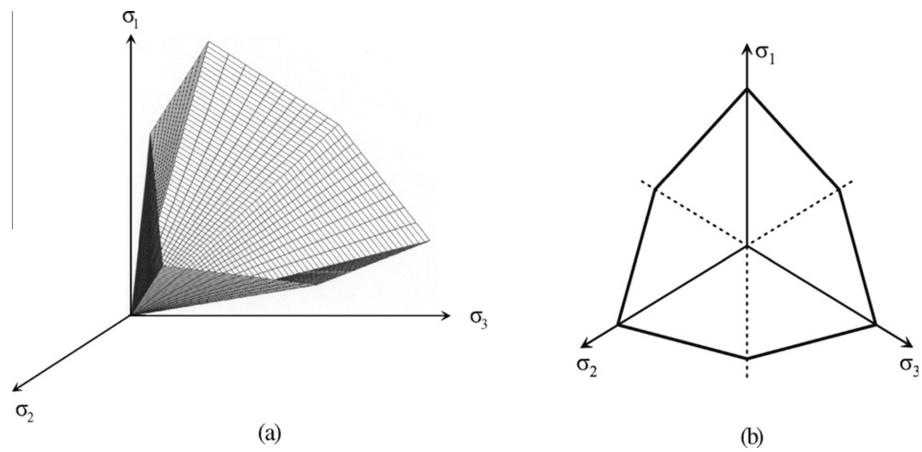

Fig. 1. Mohr–Coulomb failure criterion in (a) principal stress space and (b) in the octahedral plane.

meridional section, which is a cross section through the surface with a constant value of θ, defines a relationship between $\bar{\sigma}$ and σ_m. For the Mohr–Coulomb criterion, this relationship is linear and describes a straight line in $(\sigma_m, \bar{\sigma})$ space as shown in Fig. 2. This line intersects the σ_m-axis and it is this point of intersection that corresponds to the tip of the Mohr–Coulomb pyramid. A cross-section through the yield surface perpendicular to the hydrostatic axis , mathematically defined by a constant mean stress (i.e. σ_m = constant) is illustrated in Fig. 1(b). This cross section represents an octahedral section and is defined by a relationship between $\bar{\sigma}$ and θ. In this plane, the Mohr–Coulomb surface is represented as an irregular hexagon with sharp vertices (and hence gradient discontinuities) at the meridians corresponding to triaxial compression and extension ($\theta = \pm 30°$).

The form of the Mohr–Coulomb yield criterion is such that rounding of the tip in the meridional plane and rounding of the vertices in the octahedral plane can be accomplished independently. Various techniques to eliminate the sharp corners in the octahedral plane have been proposed, including those described by Zienkiewicz and Pande (1997), Owen and Hinton (1980) and Sloan and Booker (1986). The widely-used procedure of Sloan and Booker uses a trigonometric approximation to model the yield surface which is applied only in the vicinity of the corners. In doing so, it has the benefit over other rounding techniques of exactly modelling the Mohr–Coulomb yield surface away from the corners. The value of θ at which the yield surface moves from the true Mohr–Coulomb surface to the rounded approximation is defined by a transition angle θ_T. The value of the transition angle is typically set in the vicinity of 29°, but may be adjusted to model the Mohr–Coulomb yield surface as closely as desired. This method provides a convex rounded surface that is C1 continuous at all stress states. Furthermore, the trigonometric approximation also lies within the true Mohr–Coulomb criterion which ensures that the shear strength is modelled conservatively. The effect of this (small) reduction in strength is most noticeable under axisymmetric conditions in which the stresses are either in triaxial tension or compression and hence lie at the corners of the yield surface in the octahedral plane. However, any loss in strength for these cases can easily be predicted and used in the interpretation of the results. The rounding of the corners also influences the direction of plastic flow and the effect of this on elastoplastic calculations has recently been discussed by Taiebat and Carter (2008).

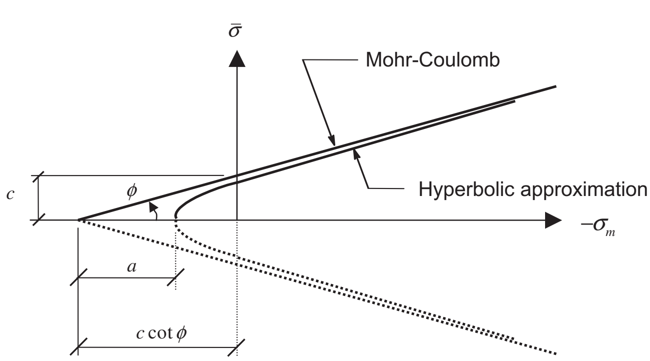

Fig. 2. Hyperbolic approximation to the Mohr–Coulomb yield criterion in the meridional plane.

Removal of the singularity at the apex or tip of the pyramid can be accomplished by adopting a suitable approximation to the Mohr–Coulomb surface in $(\sigma_m, \bar{\sigma})$ space. Zienkiewicz and Pande (1997) discuss various smooth approximations to the Mohr–Coulomb criterion, including the hyperbolic approximation shown in Fig. 2. A feature of the hyperbolic approximation is that it asymptotically approaches the Mohr–Coulomb yield surface as the mean stress increases and can be made to model the Mohr–Coulomb surface as closely as desired. The accuracy of the fit is controlled by adjusting a single parameter, which is the distance between the tip of the true yield surface and the apex of the hyperbolic surface. The hyperbolic surface is inside the Mohr–Coulomb surface at all stress states and therefore conservatively underpredicts strength in relation to the latter criterion. Hyperbolic yield criteria have been successfully utilized in a number of previous studies, an example from the field of rock mechanics being the work of Gens et al. (1990).

Abbo and Sloan (1995) combined a hyperbolic approximation with the octahedral rounding technique of Sloan and Booker (1986) to develop a smooth approximation to the Mohr–Coulomb yield surface that is continuous and differentiable for all values of the stresses (C1 continuous). In this paper we extend the technique of Sloan and Booker to derive a smooth approximation in the octahedral plane that has continuous second derivatives (C2 continuous). This, when combined with the hyperbolic approximation in the meridional plane, produces an approximation to the Mohr–Coulomb yield criterion that is C2 continuous at all stress states. The resulting surface can be used with a consistent tangent stiffness formulation to achieve full quadratic convergence of a global Newton Raphson iteration scheme.

2. Rounding the octahedral plane

The octahedral section of the yield surface is defined by a relationship between $\bar{\sigma}$ and θ (σ_m = constant), and for the

Mohr–Coulomb yield condition of Eq. (2), this relationship can be obtained as

$$\bar{\sigma} = \frac{c\cos\phi - \sigma_m \sin\phi}{K(\theta)} \tag{4}$$

Eq. (4) is a convenient form for plotting the yield surface using polar coordinates, as $\sqrt{2}\bar{\sigma}$ represents the distance between the yield surface and the origin in the octahedral plane (i.e. the radial coordinate). Smoothing of the Mohr–Coulomb surface to eliminate the vertices in the octahedral plane can be accomplished by redefining the form of the function $K(\theta)$. The exact form of this function can be selected to provide either C1 or C2 continuous smoothing of the yield surface.

2.1. C1 continuous smoothing

A C1 continuous smoothing was described by Sloan and Booker (1986) who adopted a trigonometric approximation for $K(\theta)$ in the vicinity of the vertices as shown in Fig. 3. In this scheme, when $|\theta|$ is greater than a user-specified transition angle θ_T, the function $K(\theta)$ is redefined as

$$K(\theta) = A - B\sin 3\theta \tag{5}$$

where A and B are coefficients that are obtained by enforcing C1 continuity of the original form of $K(\theta)$, as given by Eq. (3),with the trigonometric approximation at the transition angle θ_T. The transition angle θ_T specifies how accurately the rounded surface represents the true Mohr–Coulomb yield surface, with $\theta_T \to 30°$giving the most accurate approximation. In this paper, the form of the trigonometric approximation is varied slightly to that adopted in previous work by changing the sign of the second term. The form of the C1 continuous approximation to the function $K(\theta)$ adopted is defined as

$$K(\theta) = \begin{cases} A + B\sin 3\theta & |\theta| > \theta_T \\ \cos\theta - \frac{1}{\sqrt{3}}\sin\phi\sin\theta & |\theta| \leqslant \theta_T \end{cases} \tag{6}$$

in which the coefficients A and B are given by

$$A = \frac{1}{3}\cos\theta_T\left(3 + \tan\theta_T\tan 3\theta_T + \frac{1}{\sqrt{3}}\langle\theta\rangle(\tan 3\theta_T - 3\tan\theta_T)\sin\phi\right) \tag{7}$$

$$B = -\frac{1}{3\cos 3\theta_T}\left(\langle\theta\rangle\sin\theta_T + \frac{1}{\sqrt{3}}\sin\phi\cos\theta_T\right) \tag{8}$$

The function $\langle\theta\rangle$ is the sign function defined as

$$\langle\theta\rangle = \begin{cases} +1 & \text{for } \theta \geqslant 0° \\ -1 & \text{for } \theta < 0° \end{cases}$$

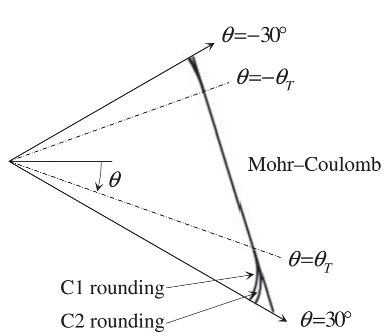

Fig. 3. Rounding of Mohr–Coulomb failure criterion in octahedral plane ($\phi = 30°, \theta_T = 20°$).

which is introduced to allow common expressions for the coefficients to be derived for both positive and negative ranges of θ via the relationship

$$\theta_t = \langle\theta\rangle\theta_T$$

The implementation of the C1 continuous approximation can benefit from the use of more convenient forms of the coefficients A and B which are presented in Appendix B. As shown in Appendix C, the C1 continuous surface is convex provided θ_T is greater than some value (computed from (C.6) in the Appendix). Choosing $\theta_T \geqslant 9.04°$, for example, ensures convexity for $\phi \leqslant 60°$.

The derivatives of the C1 continuous surface with respect to θ are

$$\frac{dK}{d\theta} = \begin{cases} 3B\cos 3\theta & |\theta| > \theta_T \\ -\sin\theta - \frac{1}{\sqrt{3}}\sin\phi\cos\theta & |\theta| \leqslant \theta_T \end{cases} \tag{9}$$

$$\frac{d^2K}{d\theta^2} = \begin{cases} -9B\sin 3\theta & |\theta| > \theta_T \\ -\cos\theta + \frac{1}{\sqrt{3}}\sin\phi\sin\theta & |\theta| \leqslant \theta_T \end{cases} \tag{10}$$

which are required later in order to calculate the gradients to the yield surface.

2.2. C2 continuous smoothing

A function that provides C2 continuous rounding of the vertices in the octahedral plane can be derived by adding an extra term to the function $K(\theta)$ proposed by Sloan and Booker (1986). A suitable function which potentially meets the requirement that the maximum extents of the yield function in the octahedral plane should occur at the vertices (with the condition $d\bar{\sigma}/d\theta = 0$ at $\theta = \pm\ 30°$) is given by

$$K(\theta) = A + B\sin 3\theta + C\sin^2 3\theta \tag{11}$$

where the coefficients A, B and C are functions of θ_T and ϕ.

By adopting the C2 continuous trigonometric approximation given by Eq. (11), the function $K(\theta)$ is fully defined as

$$K(\theta) = \begin{cases} A + B\sin 3\theta + C\sin^2 3\theta & |\theta| > \theta_T \\ \cos\theta - \frac{1}{\sqrt{3}}\sin\phi\sin\theta & |\theta| \leqslant \theta_T \end{cases} \tag{12}$$

To obtain C2 continuity of the composite yield function it is necessary that both the first and second derivatives of $K(\theta)$ are continuous at the transition angle θ_T. Differentiation of Eq. (12) with respect to θ gives

$$\frac{dK}{d\theta} = \begin{cases} 3B\cos 3\theta + 3C\sin 6\theta & |\theta| > \theta_T \\ -\sin\theta - \frac{1}{\sqrt{3}}\sin\phi\cos\theta & |\theta| \leqslant \theta_T \end{cases} \tag{13}$$

and

$$\frac{d^2K}{d\theta^2} = \begin{cases} -9B\sin 3\theta + 18C\cos 6\theta & |\theta| > \theta_T \\ -\cos\theta + \frac{1}{\sqrt{3}}\sin\phi\sin\theta & |\theta| \leqslant \theta_T \end{cases} \tag{14}$$

Matching the first and second derivatives for the rounded surface to those for the Mohr–Coulomb surface at θ_T provides the two linear equations

$$3B\cos 3\theta_T + 3C\sin 6\theta_T = -\sin\theta_T - \frac{1}{\sqrt{3}}\sin\phi\cos\theta_T$$

$$18C\cos 6\theta_T - 9B\sin 3\theta_T = -\cos\theta_T + \frac{1}{\sqrt{3}}\sin\phi\sin\theta_T$$

which can be solved to give the following expressions for the coefficients B and C

$$C = \frac{1}{18\cos^3 3\theta_T}\left(\frac{d^2K}{d\theta^2}\cos 3\theta_T + \frac{dK}{d\theta}3\langle\theta\rangle\sin 3\theta_T\right) = \frac{-\cos 3\theta_T\left(\cos\theta_T - \frac{1}{\sqrt{3}}\sin\phi\langle\theta\rangle\sin\theta_T\right) - 3\langle\theta\rangle\sin 3\theta_T\left(\langle\theta\rangle\sin\theta_T + \frac{1}{\sqrt{3}}\sin\phi\cos\theta_T\right)}{18\cos^3 3\theta_T} \quad (15)$$

$$B = \frac{1}{3\cos 3\theta_T}\frac{dK}{d\theta} - 2C\langle\theta\rangle\sin 3\theta_T = \frac{\langle\theta\rangle\sin 6\theta_T\left(\cos\theta_T - \frac{1}{\sqrt{3}}\sin\phi\langle\theta\rangle\sin\theta_T\right) - 6\cos 6\theta_T\left(\langle\theta\rangle\sin\theta_T + \frac{1}{\sqrt{3}}\sin\phi\cos\theta_T\right)}{18\cos^3 3\theta_T} \quad (16)$$

Finally, imposing continuity of $K(\theta)$ at θ_T gives the relationship

$$A + B\sin 3\theta_T + C\sin^2 3\theta_T = \cos\theta_T - \frac{1}{\sqrt{3}}\sin\phi\sin\theta_T$$

which furnishes the coefficient A as

$$A = \frac{-1}{\sqrt{3}}\sin\phi\langle\theta\rangle\sin\theta_T - B\langle\theta\rangle\sin 3\theta_T - C\sin^2 3\theta_T + \cos\theta_T \quad (17)$$

Note that the original C1 continuous scheme of Sloan and Booker (1986) is a special case which can be recovered by setting $C = 0$ and enforcing only C1 continuity at the transition angle. Fig. 3 shows an example of the C2 continuous yield surface.

In general (11) describes a non-convex yield function but, by placing some restrictions on the choice of θ_T, the convexity of the rounded Mohr–Coulomb yield surface can be guaranteed for the portions of the curve that are used to smooth the vertices. In Appendix C it is shown that the yield surface is convex provided one chooses a sufficiently large value of θ_T, where the minimum admissible value of θ_T is computed from (C.18) in the Appendix. If attention is restricted to $\phi \leqslant 60°$, for example, the yield surface is convex for $\theta_T \geqslant 9.55°$. This restriction poses no problems in practice, since the transition angle is usually selected such that $25° \leqslant \theta_T \leqslant 29°$.

2.3. Accuracy of smooth approximations

Rounding the yield surface in the above manner leads to a small reduction in the shear strength in the vicinity of the vertices where $\theta \to \pm 30°$. Using Eq. (4), the reduction in the shear strength, as measured by the reduction in the radial polar co-ordinate $\sqrt{2}\bar{\sigma}$, can be expressed as

$$r(\theta) = 1 - \frac{K_{mc}(\theta)}{K(\theta)}$$

where $K_{mc}(\theta)$ denotes the form of $K(\theta)$ associated with the Mohr–Coulomb yield function given by Eq. (2). The maximum reduction occurs under triaxial compression with $\theta = 30°$, and is presented for a range of transition angles and friction angles in Table 1.

From Table 1 it can be seen that the C1 and C2 rounding both reduce the shear strength by similar amounts, with the latter giving a slightly better approximation to the strength from the Mohr–Coulomb yield surface. Of most significance is the maximum reduction in the shear strength for different transition angles. For $\theta_T = 25°$ and $\phi = 45°$ the shear strength reduction is at most 5.3%, while the maximum reduction for $\theta_T = 29.5°$ is an order of magnitude smaller at just 0.56%. It should be emphasised again that this reduction only occurs in the vicinity of the vertices where $|\theta| > \theta_T$, and that away from the vertices the Mohr–Coulomb yield surface is modelled exactly. Eq. (2) can be coupled with Eq. (6) or (12) to generate, respectively, a smooth approximation to the Mohr–Coulomb yield surface that is C1 or C2 continuous in the octahedral plane. The closeness of the fit to the parent yield surface is controlled by the parameter θ_T.

In practice, θ_T should not be too near 30° to avoid ill-conditioning of the approximation, with a typical value being in the range 25° to 29.5°. In choosing a suitable transition angle, consideration should be given to both the accuracy and efficiency of the analysis. For axisymmetric analyses, many of the plastic stress states lie near a vertex of the Mohr–Coulomb yield surface, for which the strength of the material is reduced by the proportions listed in Table 1. For plane strain and three-dimensional analysis, this clustering does not occur and the effect of the rounding on the strength is reduced. Indeed, in practical finite element analysis, the authors have observed that the reduction in the collapse load caused by the smoothing procedure is significantly less than the values quoted in Table 1.

The efficiency of a finite element analysis will be influenced by the choice of transition angle θ_T. For large values of the transition angle (i.e close to 30°) the curvature of the surface becomes more pronounced, which has a direct influence on the performance of algorithms used to integrate the stress strain relationships. For example, with the adaptive explicit substepping methods of Sloan et al. (2001), increasing the curvature of the yield surface will increase the number of substeps required for stress points in this zone. For schemes that do not employ substepping to integrate the constitutive laws, such as an implicit backward Euler method, increasing the curvature will increase the number of iterations required at the stress point level.

Table 1
Reduction in shear strength due to rounding in octahedral plane.

θ_t	C1 rounding		C2 rounding	
	ϕ = 0° (%)	ϕ = 45° (%)	ϕ = 0° (%)	ϕ = 45° (%)
25°	2.5	5.3	1.9	4.2
26°	2.0	4.3	1.5	3.4
27°	1.5	3.2	1.1	2.5
28°	1.0	2.2	0.76	1.7
29°	0.50	1.1	0.38	0.84
29.5°	0.25	0.6	0.19	0.42

3. Rounding the apex in the meridional plane

The Mohr–Coulomb yield surface is characterised by a sharp vertex that lies at its apex. To smooth this singularity, which can become a problem for loading in tension, Abbo and Sloan (1995) formulated a hyperbolic approximation to the Mohr–Coulomb function in the meridional plane, as shown in Fig. 2. This approximation, which asymptotes to the Mohr–Coulomb surface, can be written as

$$F = \sigma_m\sin\phi + \sqrt{\bar{\sigma}^2K^2(\theta) + a^2\sin^2\phi} - c\cos\phi = 0 \quad (18)$$

where the parameter a is the distance between the tip of the Mohr–Coulomb surface and the tip of the hyperbolic approximation. Eq. (18) can be used with Eq. (6) or (12) to generate, respectively, a smooth hyperbolic approximation to the Mohr–Coulomb yield surface that is C1 or C2 continuous everywhere. The closeness of the fit to the parent yield surface is controlled by the two parameters θ_T and a.

4. Yield surface gradients

The gradients of the yield surface and plastic potential play an essential role in elastoplastic finite element analysis. These quantities are used to calculate the elastoplastic stress-strain matrix which, in turn, is used to integrate the elastoplastic stresses and form the elastoplastic tangent stiffness matrix. As the gradients are usually computed many times in a single analysis, they need to be evaluated efficiently. Nayak and Zienkiewicz (1972) proposed a convenient method for computing the gradient **a** of an isotropic function. The gradient is expressed in the form

$$\mathbf{a} = \frac{\partial F}{\partial \boldsymbol{\sigma}} = C_1 \frac{\partial \sigma_m}{\partial \boldsymbol{\sigma}} + C_2 \frac{\partial \bar{\sigma}}{\partial \boldsymbol{\sigma}} + C_3 \frac{\partial J_3}{\partial \boldsymbol{\sigma}} \tag{19}$$

where

$$C_1 = \frac{\partial F}{\partial \sigma_m}, C_2 = \frac{\partial F}{\partial \bar{\sigma}} - \frac{\tan 3\theta}{\bar{\sigma}} \frac{\partial F}{\partial \theta}, C_3 = -\frac{\sqrt{3}}{2\bar{\sigma}^3 \cos 3\theta} \frac{\partial F}{\partial \theta} \tag{20}$$

and $\boldsymbol{\sigma}^T = \{\sigma_x, \sigma_y, \sigma_z, \tau_{xy}, \tau_{yz}, \tau_{xz}\}$ is the vector of stress components. This arrangement permits different yield criteria to be implemented by simply calculating the appropriate coefficients C_1, C_2 and C_3, since all of the other derivatives are independent of F. The coefficients for the various yield surfaces discussed in this paper are described below. Note that the coefficients C_1, C_2 and C_3 have a superscript added to denote which surface they refer to.

4.1. Rounded Mohr–Coulomb yield criterion

The coefficients C_1, C_2 and C_3 for the rounded Mohr–Coulomb yield criterion are obtained by differentiating equation (2) with respect to the three stress invariants. Upon substitution into (20) this gives the coefficients

$$C_1^{rmc} = \sin\phi, \quad C_2^{rmc} = K - \tan 3\theta \frac{dK}{d\theta}, \quad C_3^{rmc} = -\frac{\sqrt{3}}{2\bar{\sigma}^2 \cos 3\theta} \frac{dK}{d\theta} \tag{21}$$

Gradients for the rounded form of the Mohr–Coulomb yield surface are computed using Eq. (21) with the rounded $K(\theta)$ function given by Eq. (12). The gradients to the Mohr–Coulomb surface, with an unrounded octahedral cross-section, may also be evaluated using the above expressions except that Eq. (3) is used to define $K(\theta)$.

The constants given in Eq. (21) are not suitable for implementation in a computer program as $1/\cos 3\theta$ and $\tan 3\theta$ tend to infinity at $\theta = \pm 30°$. These terms can be eliminated for the rounded surface by substituting the expressions for K and $dK/d\theta$, as given by Eqs. (9) and (13), into Eq. (21). The constants may now be evaluated as

$$C_1^{rmc} = \sin\phi \tag{22}$$

$$C_2^{rmc} = \begin{cases} A - 2B\sin 3\theta - 5C\sin^2 3\theta & |\theta| > \theta_T \\ K - \frac{dK}{d\theta}\tan 3\theta & |\theta| \leqslant \theta_T \end{cases} \tag{23}$$

$$C_3^{rmc} = \begin{cases} -\frac{3\sqrt{3}}{2\bar{\sigma}^2}(B + 2C\sin 3\theta) & |\theta| > \theta_T \\ -\frac{\sqrt{3}}{2\bar{\sigma}^2\cos 3\theta}\frac{dK}{d\theta} & |\theta| \leqslant \theta_T \end{cases} \tag{24}$$

which avoids any computational problems. Further computational problems associated with small values of $\bar{\sigma}$ may also be avoided by expressing the gradients in the form

$$\mathbf{a} = \frac{\partial F}{\partial \boldsymbol{\sigma}} = C_1 \frac{\partial \sigma_m}{\partial \boldsymbol{\sigma}} + C_2 \frac{\partial \bar{\sigma}}{\partial \boldsymbol{\sigma}} + (\bar{\sigma}^2 C_3)\left(\frac{1}{\bar{\sigma}^2}\frac{\partial J_3}{\partial \boldsymbol{\sigma}}\right) \tag{25}$$

and computing the quantities $\bar{\sigma}^2 C_3$ and $(1/\bar{\sigma}^2)(\partial J_3/\partial \boldsymbol{\sigma})$. In this way the division of values by $\bar{\sigma}$ can either be avoided through cancellation or factored so that it divides a quantity of similar magnitude such as the components of the deviatoric stresses.

4.2. Hyperbolic yield criterion

The coefficients for the hyperbolic yield surface are obtained by differentiation of Eq. (18). These can be expressed very simply in terms of the above Mohr–Coulomb coefficients as

$$C_1^h = C_1^{rmc}, \quad C_2^h = \alpha C_2^{rmc}, \quad C_3^h = \alpha C_3^{rmc} \tag{26}$$

where

$$\alpha = \frac{\bar{\sigma}K}{\sqrt{\bar{\sigma}^2 K^2 + a^2 \sin^2\phi}}$$

A hyperbolic Mohr–Coulomb surface which is rounded in the octahedral plane is obtained by using Eq. (12) for $K(\theta)$, while an unrounded surface can be modelled by using Eq. (3). Use of the former ensures the yield surface is C2 continuous everywhere, even for a purely hydrostatic tensile stress state.

5. Gradient derivatives

In standard implicit stress integration methods, such as the backward Euler return algorithm discussed by Crisfield (1991), it is necessary to compute the derivatives of the gradient vector with respect to the stresses. Since implicit integration schemes are widely used in finite element codes in combination with a consistent tangent stiffness formulation, expressions for the gradient derivatives of the rounded hyperbolic surface are now derived.

Differentiation of Eq. (19) gives

$$\frac{\partial \mathbf{a}}{\partial \boldsymbol{\sigma}} = \frac{\partial C_2}{\partial \boldsymbol{\sigma}}\frac{\partial \bar{\sigma}}{\partial \boldsymbol{\sigma}} + C_2 \frac{\partial^2 \bar{\sigma}}{\partial \boldsymbol{\sigma}^2} + \frac{\partial C_3}{\partial \boldsymbol{\sigma}}\frac{\partial J_3}{\partial \boldsymbol{\sigma}} + C_3 \frac{\partial^2 J_3}{\partial \boldsymbol{\sigma}^2} \tag{27}$$

where the derivatives of the stress invariants $\partial\sigma/\partial\boldsymbol{\sigma}$, $\partial J_3/\partial\boldsymbol{\sigma}$, $\partial^2\bar{\sigma}/\partial\boldsymbol{\sigma}^2$ and $\partial^2 J_3/\partial\boldsymbol{\sigma}^2$ are all defined in Appendix A. The derivatives of the coefficients (C_2, C_3) with respect to the stresses are now evaluated for each of the smoothed yield functions.

5.1. Rounded Mohr–Coulomb yield criterion

For the rounded Mohr–Coulomb criterion the derivatives of the gradient coefficients are

$$\frac{\partial C_2^{rmc}}{\partial \boldsymbol{\sigma}} = \frac{\partial \theta}{\partial \boldsymbol{\sigma}}\left(\frac{dK}{d\theta} - \frac{d^2K}{d\theta^2}\tan 3\theta - 3\frac{dK}{d\theta}\sec^2 3\theta\right) \tag{28}$$

$$\frac{\partial C_3^{rmc}}{\partial \boldsymbol{\sigma}} = \frac{-\sqrt{3}}{2\bar{\sigma}^2\cos 3\theta}\left[\frac{\partial \theta}{\partial \boldsymbol{\sigma}}\left(\frac{d^2K}{d\theta^2} + 3\frac{dK}{d\theta}\tan 3\theta\right) - \frac{2}{\bar{\sigma}}\frac{dK}{d\theta}\frac{\partial \bar{\sigma}}{\partial \boldsymbol{\sigma}}\right] \tag{29}$$

where

$$\frac{\partial \theta}{\partial \boldsymbol{\sigma}} = \frac{-\sqrt{3}}{2\bar{\sigma}^3\cos 3\theta}\left(\frac{\partial J_3}{\partial \boldsymbol{\sigma}} - \frac{3J_3}{\bar{\sigma}}\frac{\partial \bar{\sigma}}{\partial \boldsymbol{\sigma}}\right)$$

The expressions given in Eqs. (28) and (29) are not suitable for implementation in a computer program as some of the trigonometric terms tend to infinity at $\theta = \pm 30°$. For the rounded surface it is possible to eliminate these terms by substituting the expressions for K and $dK/d\theta$, as given by Eqs. (9) and (13). The derivatives of the constants may now be evaluated as

$$\frac{\partial C_2^{rmc}}{\partial \boldsymbol{\sigma}} = \begin{cases} -6\cos 3\theta(B + 5C\sin 3\theta)\frac{\partial \theta}{\partial \boldsymbol{\sigma}} & |\theta| > \theta_T \\ \frac{\partial \theta}{\partial \boldsymbol{\sigma}}\left(\frac{dK}{d\theta} - \frac{d^2K}{d\theta^2}\tan 3\theta - 3\frac{dK}{d\theta}\sec^2 3\theta\right) & |\theta| \leqslant \theta_T \end{cases} \tag{30}$$

$$\frac{\partial C_3^{rmc}}{\partial \boldsymbol{\sigma}} = \begin{cases} \frac{3\sqrt{3}}{\bar{\sigma}^3}\left[-3C\bar{\sigma}\cos 3\theta\frac{\partial \theta}{\partial \boldsymbol{\sigma}} + (B + 2C\sin 3\theta)\frac{\partial \bar{\sigma}}{\partial \boldsymbol{\sigma}}\right] & |\theta| > \theta_T \\ \frac{-\sqrt{3}}{2\bar{\sigma}^2\cos 3\theta}\left[\frac{\partial \theta}{\partial \boldsymbol{\sigma}}\left(\frac{d^2K}{d\theta^2} + 3\frac{dK}{d\theta}\tan 3\theta\right) - \frac{2}{\bar{\sigma}}\frac{dK}{d\theta}\frac{\partial \bar{\sigma}}{\partial \boldsymbol{\sigma}}\right] & |\theta| \leqslant \theta_T \end{cases} \tag{31}$$

which, in conjunction with the grouping of terms to ensure division by numbers of similar order, avoids computational difficulties. Gradient derivatives for the rounded Mohr–Coulomb criterion with C2 continuity are obtained by using Eq. (12) for $K(\theta)$. This form can be used in a consistent tangent formulation with implicit integration, provided the mean normal stresses are such that the apex of the Mohr–Coulomb surface is avoided.

5.2. Hyperbolic yield criterion

The derivatives of the coefficients for the hyperbolic yield surface can be expressed conveniently in terms of the Mohr–Coulomb coefficients and their derivatives according to

$$\frac{\partial C_2^h}{\partial \boldsymbol{\sigma}} = \alpha \frac{\partial C_2^{rmc}}{\partial \boldsymbol{\sigma}} + C_2^{rmc} \frac{\partial \alpha}{\partial \boldsymbol{\sigma}} \tag{32}$$

$$\frac{\partial C_3^h}{\partial \boldsymbol{\sigma}} = \alpha \frac{\partial C_3^{rmc}}{\partial \boldsymbol{\sigma}} + C_3^{rmc} \frac{\partial \alpha}{\partial \boldsymbol{\sigma}} \tag{33}$$

in which

$$\frac{\partial \alpha}{\partial \boldsymbol{\sigma}} = \frac{1-\alpha^2}{\sqrt{\bar{\sigma}^2 K^2 + a^2 \sin^2\theta}} \left(\frac{\partial \bar{\sigma}}{\partial \boldsymbol{\sigma}} K + \bar{\sigma} \frac{dK}{d\theta} \frac{\partial \theta}{\partial \boldsymbol{\sigma}} \right) \tag{34}$$

Thus the second derivatives for the hyperbolic surface with a rounded octahedral cross-section can be found from Eq. (27) by using (6) or (12) to define $K(\theta)$ in Eqs. (21), (28), (29), (32), (33) and (34).

6. Conclusions

A C2 continuous yield surface is derived that closely approximates the Mohr–Coulomb yield surface. The error in this approximation can be controlled by adjusting two simple parameters. As the new yield function is C2 continuous, it can be used with a consistent tangent solution scheme to provide quadratic convergence of the global iterations.

Acknowledgement

The research reported in this paper was made possible by the Australian Laureate Fellowship grant FL0992039 awarded to Professor Scott Sloan by the Australian Research Council.

Appendix A. Stress invariants

Nayak and Zienkiewicz (1972) proposed a form of the Mohr–Coulomb yield criterion that avoids the need to compute principal stresses. They expressed the criterion in the form

$$F = \sigma_m \sin\phi + \bar{\sigma} K(\theta) - c\cos\phi = 0 \tag{A.1}$$

where σ_m denotes the mean normal stress, $\bar{\sigma}$ is a measure of the deviatoric stress, and θ is the Lode angle. These three stress invariants are found from the Cartesian stresses $\boldsymbol{\sigma}^T = \{\sigma_x, \sigma_y, \sigma_z, \tau_{xy}, \tau_{yz}, \tau_{xz}\}$ using the following relationships

$$\sigma_m = \frac{1}{3}(\sigma_x + \sigma_y + \sigma_z)$$

$$\bar{\sigma} = \sqrt{\frac{1}{2}\left(s_x^2 + s_y^2 + s_z^2\right) + \tau_{xy}^2 + \tau_{yz}^2 + \tau_{xz}^2}$$

$$\theta = \frac{1}{3}\sin^{-1}\left(-\frac{3\sqrt{3}}{2}\frac{J_3}{\bar{\sigma}^3}\right) (|\theta| \leqslant 30^\circ)$$

where

$$J_3 = s_x s_y s_z + 2\tau_{xy}\tau_{yz}\tau_{zx} - s_x\tau_{yz}^2 - s_y\tau_{xz}^2 - s_z\tau_{xy}^2$$

and

$$s_x = \sigma_x - \sigma_m, s_y = \sigma_y - \sigma_m, s_z = \sigma_z - \sigma_m$$

are the deviatoric stresses.

Elastoplastic finite element calculations require the gradients to the yield function and the plastic potential. These are typically derived using the chain rule and require the derivatives of the three stress invariants with respect to the Cartesian stresses. The first derivatives of the stress invariants with respect to the Cartesian stresses are given by

$$\frac{\partial \sigma_m}{\partial \boldsymbol{\sigma}} = \frac{1}{3}\begin{Bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{Bmatrix}, \quad \frac{\partial \bar{\sigma}}{\partial \boldsymbol{\sigma}} = \frac{1}{2\bar{\sigma}}\begin{Bmatrix} s_x \\ s_y \\ s_z \\ 2\tau_{xy} \\ 2\tau_{yz} \\ 2\tau_{xz} \end{Bmatrix},$$

$$\frac{\partial J_3}{\partial \boldsymbol{\sigma}} = \begin{Bmatrix} s_y s_z - \tau_{yz}^2 \\ s_x s_z - \tau_{xz}^2 \\ s_x s_y - \tau_{xy}^2 \\ 2(\tau_{yz}\tau_{xz} - s_z\tau_{xy}) \\ 2(\tau_{xz}\tau_{xy} - s_x\tau_{yz}) \\ 2(\tau_{xy}\tau_{yz} - s_y\tau_{xz}) \end{Bmatrix} + \frac{\bar{\sigma}^2}{3}\begin{Bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{Bmatrix} \tag{A.2}$$

The second derivatives of the stress invariants, which are needed for consistent tangent formulations, are given by

$$\frac{\partial^2 \bar{\sigma}}{\partial \boldsymbol{\sigma}^2} = \begin{pmatrix} \frac{1}{3} - \frac{s_x s_x}{4\bar{\sigma}^2} & & & & & \\ -\frac{1}{6} - \frac{s_x s_y}{4\bar{\sigma}^2} & \frac{1}{3} - \frac{s_y s_y}{4\bar{\sigma}^2} & & & \text{symmetric} & \\ -\frac{1}{6} - \frac{s_x s_z}{4\bar{\sigma}^2} & -\frac{1}{6} - \frac{s_y s_z}{4\bar{\sigma}^2} & \frac{1}{3} - \frac{s_z s_z}{4\bar{\sigma}^2} & & & \\ -\frac{s_x \tau_{xy}}{2\bar{\sigma}^2} & -\frac{s_y \tau_{xy}}{2\bar{\sigma}^2} & -\frac{s_z \tau_{xy}}{2\bar{\sigma}^2} & 1 - \frac{\tau_{xy}\tau_{xy}}{\bar{\sigma}^2} & & \\ -\frac{s_x \tau_{yz}}{2\bar{\sigma}^2} & -\frac{s_y \tau_{yz}}{2\bar{\sigma}^2} & -\frac{s_z \tau_{yz}}{2\bar{\sigma}^2} & \frac{\tau_{yz}\tau_{xy}}{\bar{\sigma}^2} & 1 - \frac{\tau_{yz}\tau_{yz}}{\bar{\sigma}^2} & \\ -\frac{s_x \tau_{zx}}{2\bar{\sigma}^2} & -\frac{s_y \tau_{zx}}{2\bar{\sigma}^2} & -\frac{s_z \tau_{zx}}{2\bar{\sigma}^2} & \frac{\tau_{zx}\tau_{xy}}{\bar{\sigma}^2} & \frac{\tau_{zx}\tau_{yz}}{\bar{\sigma}^2} & 1 - \frac{\tau_{zx}\tau_{zx}}{\bar{\sigma}^2} \end{pmatrix}$$

and

$$\frac{\partial^2 J_3}{\partial \boldsymbol{\sigma}^2} = \frac{1}{3}\begin{bmatrix} s_x - s_y - s_z & & & & & \\ 2s_z & s_y - s_x - s_z & & & & \\ 2s_y & 2s_x & s_z - s_x - s_y & & & \\ 2\tau_{xy} & 2\tau_{xy} & -4\tau_{xy} & -6s_z & & \\ -4\tau_{yz} & 2\tau_{yz} & 2\tau_{yz} & 6\tau_{zx} & -6s_x & \\ 2\tau_{zx} & -4\tau_{zx} & 2\tau_{zx} & 6\tau_{yz} & 6\tau_{xy} & -6s_y \end{bmatrix}$$

Appendix B. Implementation

When implementing rounded yield surfaces in a finite element program it is convenient to adopt a constant value of the transition angle θ_T. This permits many of the terms in the coefficients A, B and C of the rounded form of $K(\theta)$ to be treated as constants and hard-coded to minimise computer arithmetic. In this Appendix, expressions for these coefficients are derived that are suitable for efficient implementation using a transition angle that is fixed within the software. The expressions may be simplified even further for the Tresca yield criterion for which $\sin\phi = 0$.

B.1. Efficient implementation of C1 continuous rounding

For C1 continuous rounding in the octahedral plane, the coefficient A given by Eq. (7) can be expressed in the form

$$A = A_1 + A_2\langle\theta\rangle \sin\phi \tag{B.1}$$

in which

$$A_1 = \frac{1}{3}\cos\theta_T(3 + \tan\theta_T \tan 3\theta_T)$$
$$A_2 = \frac{1}{3}\cos\theta_T\left(\frac{1}{\sqrt{3}}(\tan 3\theta_T - 3\tan\theta_T)\right) \tag{B.2}$$

Similarly, the coefficient B in Eq. (8) can be decomposed into the form

$$B = B_1\langle\theta\rangle + B_2 \sin\phi \tag{B.3}$$

where

$$B_1 = \frac{\sin\theta_T}{3\cos 3\theta_T} \quad B_2 = \frac{\cos\theta_T}{3\sqrt{3}\cos 3\theta_T} \tag{B.4}$$

The expressions for A_1, A_2, B_1 and B_2 are functions of only the transition angle θ_T. Values of these parameters for a range of transition angles are given in Table B.1.

B.2. Efficient implementation of C2 continuous rounding

For the C2 continuous rounding in the octahedral plane, the coefficient C in Eq. (15) can be expressed in the form

$$C = C_1 + C_2\langle\theta\rangle \sin\phi \tag{B.5}$$

in which

$$C_1 = \frac{-\cos 3\theta_T \cos\theta_T - 3\sin 3\theta_T \sin\theta_T}{18\cos^3 3\theta_T}$$
$$C_2 = \frac{1}{\sqrt{3}}\left(\frac{\cos 3\theta_T \sin\theta_T - 3\sin 3\theta_T \cos\theta_T}{18\cos^3 3\theta_T}\right) \tag{B.6}$$

Similarly, the coefficient B in Eq. (16) can be decomposed into the form

$$B = B_1\langle\theta\rangle + B_2 \sin\phi \tag{B.7}$$

in which

$$B_1 = \frac{\cos\theta_T \sin 6\theta_T - 6\cos 6\theta_T \sin\theta_T}{18\cos^3 3\theta_T}$$
$$B_2 = \frac{-(\sin\theta_T \sin 6\theta_T + 6\cos 6\theta_T \cos\theta_T)}{18\sqrt{3}\cos^3 3\theta_T} \tag{B.8}$$

Substitution of equations (B.5) and (B.7) into Eq. (17) gives an expression for the coefficient A of the form

$$A = A_1 + A_2\langle\theta\rangle \sin\phi \tag{B.9}$$

Table B.1
Constants for C1 continuous rounding.

θ_t	A_1/A_2	B_1/B_2
25°	1.43205206204423 0.40694185837461	−0.54429052490231 −0.67390332449839
26°	1.58625207840266 0.56068026013645	−0.70281625348543 −0.83195415408635
27°	1.84646759264791 0.82053449275842	−0.96737101086344 −1.09614134894032
28°	2.37185544260506 1.34566308592589	−1.49710917042685 1.62561792415694
29°	3.95819258428804 2.93184419579307	−3.08780460604590 −3.21615679165482
29.5°	7.13865472324241 6.11226727092061	−6.27044775313959 −6.39876084142940

Table B.2
Constants for C2 continuous rounding.

θ_t	A_1/A_2	B_1/B_2	C_1/C_2
25°	−2.93057555085368 −3.93747122467738	8.48875837836269 8.32143144099294	−4.67585018301484 −4.65632790876395
26°	−7.12688371578337 −8.13395632105966	17.1127686084504 16.9458057150242	−9.10679781280996 −9.08746279997706
27°	−19.1707792133233 −20.1779910875781	41.5910878513868 41.4244083371757	−21.5444777026559 −21.5252868432642
28°	−69.4588436196005 −70.4661558583851	142.955616097339 142.789139113885	−72.6242056311263 −72.6051169464523
29°	−575.081604828925 −576.088977641021	1156.58107611761 1156.41472069709	−580.630173835517 −580.611146141268
29.5°	−4634.09083121302 −4635.09821920999	9279.37048135174 9279.20415632701	−4644.41198854414 −4644.39297606081

where

$$A_1 = \cos\theta_T - B_1 \sin 3\theta_T - C_1 \sin^2 3\theta_T$$
$$A_2 = \frac{-1}{\sqrt{3}}\sin\theta_T - B_2 \sin 3\theta_T - C_2 \sin^2 3\theta_T \tag{B.10}$$

The expressions for A_1, A_2, B_1, B_2, C_1 and C_2 are functions of only the transition angle (θ_t). Values of these parameters for a range of transition angles are given in Table B.2.

Appendix C. Proof of convexity

The rounded Mohr–Coulomb yield surface is in general non-convex, but in the following it will be demonstrated that the surface is convex provided that θ_T is sufficiently large.

As $\bar{\sigma}$ is inversely proportional to $K(\theta)$, convexity of the yield surface is ensured provided $1/K(\theta)$ is a convex function. As shown by van Eekelen (1980), a function $g(\theta)$ is convex in Cartesian space if the following relationship is satisfied

$$g(\theta)^2 + 2g'(\theta)^2 - g(\theta)g''(\theta) \geqslant 0 \tag{C.1}$$

Upon substitution of $g(\theta) = 1/K(\theta)$ into equation (C.1) and making use of the fact that $K(\theta) > 0$, the condition required for convexity reduces to

$$K''(\theta) + K(\theta) \geqslant 0 \tag{C.2}$$

For the Mohr–Coulomb yield surface, i.e., for $|\theta| \leqslant \theta_T$, it may readily be shown that $K''(\theta) = -K(\theta)$, which satisfies (C.2). It thus remains only to show that (C.2) is satisfied for $|\theta| \geqslant \theta_T$. This is demonstrated in the following sections for the C1 and C2 continuous surfaces.

C.1. Proof of convexity for C1 continuous rounding

For the C1 continuous surface defined by Eqs. (6)–(8), the convexity condition (C.2) can be written in the form

$$\sin\phi M_1(\theta) + N_1(\theta) \geqslant 0 \tag{C.3}$$

where

$$M_1(\theta) = \frac{8}{3\sqrt{3}}\cos\theta_T(\theta \sin^3\theta_T + \sin 3\theta)$$
$$N_1(\theta) = \frac{2}{3}\cos 2\theta_T + \frac{1}{3}\cos 4\theta_T + \frac{8}{3}\theta \sin\theta_T \sin 3\theta$$

For $\theta \geqslant 0$, it can be shown that $M_1(\theta) \geqslant 0$ and $N_1(\theta) \geqslant 0$ for all $\theta_T \in [0, 30°]$. Thus, for $\theta \geqslant 0$, the condition (C.3) is always satisfied and the yield surface is convex.

For $\theta \leqslant 0$, it can again be shown that $N_1(\theta) \geqslant 0$, whereas $M_1(\theta)$ can be positive or negative. In order to demonstrate convexity for $\theta \leqslant 0$, condition (C.3) is first rewritten as

$$\alpha_1 \sin 3\theta + \beta_1 \geqslant 0 \tag{C.4}$$

where $\alpha_1 = -8B\cos 3\theta_T$ and $\beta_1 = A\cos 3\theta_T$. For a particular value of θ_T and ϕ, the function $\alpha_1 \sin 3\theta + \beta_1$ in (C.4) can attain a minimum at an endpoint, $\theta = -\theta_T$ or $\theta = -30°$, or at some value on the interval $\theta \in (-30°, -\theta_T)$. Supposing that the minimum is at an endpoint, the sufficient convexity conditions are

$$-\alpha_1 \sin 3\theta_T + \beta_1 \geqslant 0, \quad -\alpha_1 + \beta_1 \geqslant 0 \tag{C.5}$$

If the minimum is not at an endpoint but rather on the interval $\theta \in (-30°, -\theta_T)$, the first derivative of $\alpha_1 \sin 3\theta + \beta_1$ must vanish at some point, viz.

$$\frac{d}{d\theta}(\alpha_1 \sin 3\theta + \beta_1) = 3\alpha_1 \cos 3\theta = 0$$

Since $\cos 3\theta = 0$ only at the endpoint $\theta = -30°$ and $\alpha_1 = 0$ would imply the derivative is zero everywhere, we conclude that no extremum can exist on the interval $\theta \in (-30°, -\theta_T)$ and that $\alpha_1 \sin 3\theta + \beta_1$ must be minimal at an endpoint. To prove that the yield function is convex, it thus suffices to check that the inequalities in (C.5) are satisfied.

It can be shown that the first inequality in (C.5) is satisfied for all $\theta_T \in [0, 30°)$ and $\phi \in [0, 90°]$, however $-\alpha_1 + \beta_1$ can be positive or negative depending on θ_T and ϕ. By solving $-\alpha_1 + \beta_1 = 0$ with $\theta_T = \theta_{T,min}$, we find

$$\sin\phi = \frac{\sqrt{3}(2\cos 2\theta_{T,min} + \cos 4\theta_{T,min} + 8\sin\theta_{T,min})}{8\cos\theta_{T,min}(1 + \sin^3\theta_{T,min})} \tag{C.6}$$

It can be shown subsequently that the second inequality in (C.5) is satisfied only for $\theta_T \geqslant \theta_{T,min}$.

Combining results for $\theta \geqslant 0$ and $\theta \leqslant 0$, we conclude that the C1 continuous yield surface is convex provided $\theta_T \geqslant \theta_{T,min}$, where $\theta_{T,min}$ depends on ϕ according to (C.6).

C.2. Numerical convexity test for C2 continuous rounding

The convexity of the rounded forms of $K(\theta)$ is dependent upon both the transition angle θ_T and the friction angle ϕ. The convexity of the surface may be verified numerically by investigating values of the following function at discrete points

$$r(\theta, \theta_T, \phi) = K'' + K \tag{C.7}$$

where $r(\theta, \theta_T, \phi)$ must be positive for all values of θ on the intervals $\theta \in [-30°, -\theta_T]$ and $\theta \in [\theta_T, 30°]$ if the function $K(\theta)$ is to be convex. By considering the minimum values of $r(\theta, \theta_T, \phi)$ over each of the rounded intervals as defined by the functions

$$R^+(\theta_T, \phi) = \min\{r(\theta, \theta_T, \phi)\}\theta \in [\theta_T, 30°] \tag{C.8}$$
$$R^-(\theta_T, \phi) = \min\{r(\theta, \theta_T, \phi)\}\theta \in [-\theta_T, -30°] \tag{C.9}$$

the range of values of the transition angle θ_T and the friction angle ϕ at which the yield surface is convex may be illustrated as surface plots of $R^+(\theta_T, \phi)$ and $R^-(\theta_T, \phi)$. As the rounded surface is C2 continuous with the Mohr–Coulomb yield surface at θ_T the function $r(\theta, \theta_T, \phi)$ will evaluate to exactly zero when $\theta = \theta_T$. Hence the functions $R^+(\theta_T, \phi)$ and $R^-(\theta_T, \phi)$ will evaluate to exactly zero for values of θ_T and ϕ for which the yield surface is convex.

The functions $R^+(\theta_T, \phi)$ and $R^-(\theta_T, \phi)$ have been evaluated numerically with the minimum value of $r(\theta, \theta_T, \phi)$ determined by evaluating the functions at 1000 points in the intervals $\theta \in [-30°, -\theta_T]$ and $\theta \in [\theta_T, 30°]$. To illustrate the convexity of the yield surface, the functions $R^+(\theta_T, \phi)$ and $R^-(\theta_T, \phi)$ were evaluated over the ranges $\theta_T \in [0, 30°]$ and $\phi \in [0°, 50°]$ at intervals of 0.1°. The function $R^+(\theta_T, \phi)$ was found to be zero for all values of θ and ϕ on the specified range showing that the rounded surface is always convex on the interval $\theta \in [\theta_T, 30°]$. For negative values of $\theta \in [-30°, -\theta_T]$ the function $R^-(\theta_T, \phi)$ was found to be negative for small values of θ_T and ϕ . The function $R^-(\theta_T, \phi)$ is plotted in Figs. C.1 and C.2, which clearly show a region in which the yield function is non-convex. For friction angles up to 50° it can be seen

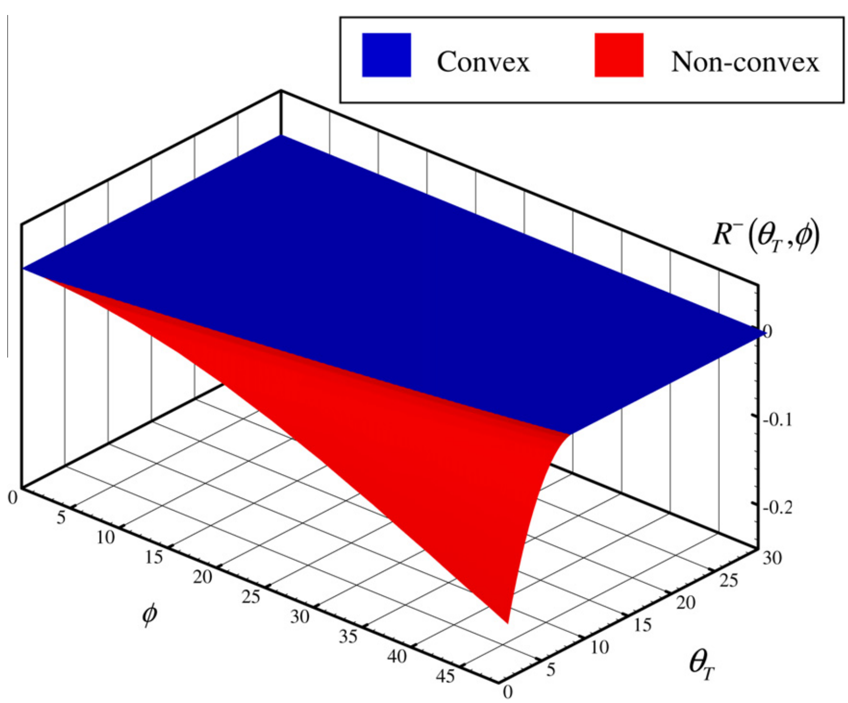

Fig. C.1. Plot of $R^-(\theta_T, \phi)$ showing convexity on interval $\theta \in [-\theta_T, -30°]$.

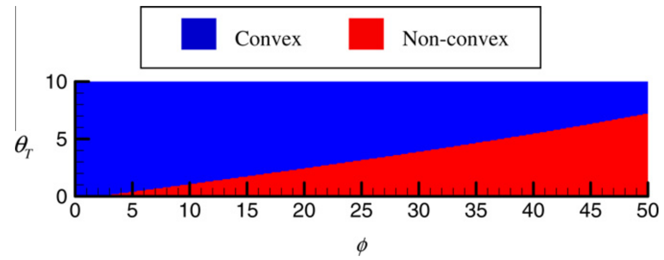

Fig. C.2. Plot of $R^{-}(\theta_T, \phi)$ showing convexity on interval $\theta \in [-\theta_T, -30^\circ]$ for small values of θ_T.

that choosing values of $\theta_T > 10^\circ$ will ensure that the surface is convex.

C.3. Proof of convexity for C2 continuous rounding

As in the proof of convexity for the C1 continuous surface, the convexity condition for the C2 continuous surface can be expressed in the form

$$\sin\phi M_2(\theta) + N_2(\theta) \geqslant 0 \tag{C.10}$$

which is obtained by substituting the expression for $K(\theta)$ from (12) (with coefficients from (15)–(17)) into (C.2). The functions $M_2(\theta)$ and $N_2(\theta)$, while straightforward to obtain, are not written explicitly due to their length. It can be shown that $N_2(\theta) \geqslant 0$ for $|\theta| \geqslant \theta_T$ and $\theta_T \in [0, 30^\circ]$, whereas $M_2(\theta)$ may be positive or negative. As an immediate consequence of (C.10) and $N_2(\theta) \geqslant 0$, the following is observed: if convexity can be demonstrated for some value of friction angle ϕ^*, then the yield surface is convex for all $\phi \in [0, \phi^*]$.

For the remainder of the proof the following alternative form of the convexity condition, obtained from manipulating (C.10), is used

$$\alpha_2 \sin 3\theta + \beta_2 \sin^2 3\theta + \gamma_2 \geqslant 0 \tag{C.11}$$

where

$$\alpha_2 = \frac{4}{\sqrt{3}} \sin\phi(7\cos 5\theta_T + 5\cos 7\theta_T) + 4\theta(-7\sin 5\theta_T + 5\sin 7\theta_T)$$
$$\beta_2 = \frac{35}{3}(6\cos 2\theta_T - 3\cos 4\theta_T + 8\sqrt{3}\theta \sin\phi \cos^3\theta_T \sin\theta_T)$$
$$\gamma_2 = \frac{1}{6}\sin 3\theta_T[-8\sqrt{3}\theta\sin\phi\cos^3\theta_T(36 - 29\cos 2\theta_T + 5\cos 4\theta_T) - 3(105\sin\theta_T + 14\sin 5\theta_T + 5\sin 7\theta_T)$$

The function $\alpha_2 \sin 3\theta + \beta_2 \sin^2 3\theta + \gamma_2$ in (C.11) may be minimal being at an endpoint ($\theta = \pm\theta_T; \theta = \pm 30^\circ$) or some intermediate point ($\theta \in (\theta_T, 30^\circ); \theta \in (-30^\circ, -\theta_T)$). Supposing that the minimum is at an endpoint, the sufficient convexity conditions are

$$\alpha_2 \sin 3\theta_T + \beta_2 \sin^2 3\theta_T + \gamma_2 \geqslant 0, \quad \alpha_2 + \beta_2 + \gamma_2 \geqslant 0 \text{ for } \theta \geqslant 0 \tag{C.12}$$

$$-\alpha_2 \sin 3\theta_T + \beta_2 \sin^2 3\theta_T + \gamma_2 \geqslant 0, \quad -\alpha_2 + \beta_2 + \gamma_2 \geqslant 0 \text{ for } \theta \leqslant 0 \tag{C.13}$$

If the minimum occurs on the interval, it follows that

$$\frac{d}{d\theta}(\alpha_2 \sin 3\theta + \beta_2 \sin^2 3\theta + \gamma_2) = 3\cos 3\theta(\alpha_2 + 2\beta_2 \sin 3\theta) = 0$$

Since $\cos 3\theta > 0$, this requires $\alpha_2 + 2\beta_2 \sin 3\theta = 0$ or

$$\sin 3\theta = -\frac{\alpha_2}{2\beta_2} \tag{C.14}$$

In order for a minimum to exist on the interval $\theta \in (\theta_T, 30^\circ)$ or $\theta \in (-30^\circ, -\theta_T)$, as opposed to some point outside the interval, the following is also required

$$\sin 3\theta_T \leqslant -\frac{\alpha_2}{2\beta_2} \leqslant 1 \text{ for } \theta \geqslant 0 \tag{C.15}$$

$$-1 \leqslant -\frac{\alpha_2}{2\beta_2} \leqslant -\sin 3\theta_T \text{ for } \theta \leqslant 0 \tag{C.16}$$

Upon combining (C.11) and (C.14), the convexity condition corresponding to the function in (C.11) attaining a minimum on the interval $\theta \in (\theta_T, 30^\circ)$ or $\theta \in (-30^\circ, -\theta_T)$ is

$$-\frac{\alpha_2^2}{4\beta_2} + \gamma_2 \geqslant 0 \tag{C.17}$$

To prove that the yield function is convex, it thus suffices to check that inequalities (C.12), (C.13) and (C.17) are satisfied, where (C.17) need not be satisfied if (C.15) and (C.16) are not satisfied.

It is straightforward to show $\alpha_2 \sin 3\theta_T + \beta_2 \sin^2 3\theta_T + \gamma_2 = 0$ with $\theta \geqslant 0$ and $-\alpha_2 \sin 3\theta_T + \beta_2 \sin^2 3\theta_T + \gamma_2 = 0$ with $\theta \leqslant 0$, implying that the first conditions in (C.12) and (C.13) are satisfied for arbitrary values of θ_T and ϕ. Assuming $\theta \geqslant 0$ and $\phi = 90^\circ$, it also can be shown that the second inequality in (C.12) is satisfied for all $\theta_T \in [0, 30^\circ]$, where $\alpha_2 + \beta_2 + \gamma_2$ has only one root in the interval at $\theta_T = 30^\circ$. Likewise, with $\theta \leqslant 0$ and $\phi = 90^\circ$, the second inequality in (C.13) is satisfied for all $\theta_T \in [0, 30^\circ]$, where $-\alpha_2 + \beta_2 + \gamma_2$ has one root at $\theta_T = 30^\circ$. This implies that conditions (C.12) and (C.13) are satisfied for all $\theta_T \in [0, 30^\circ]$ and $\phi \in [0, 90^\circ]$.

For $\theta \geqslant 0$, it can be shown for $\phi = 90^\circ$ that $-\frac{\alpha_2}{2\beta_2} < \sin 3\theta_T$ for all $\theta_T \in (0, 30^\circ)$. Therefore, according to (C.15), no extremum exists on the interval $\theta \in (-30^\circ, -\theta_T)$, and the minimum must be at an endpoint. By previous results, the yield surface is therefore convex for all $\theta_T \in [0, 30^\circ]$ and $\phi \in [0, 90^\circ]$ with $\theta \geqslant 0$.

For $\theta \leqslant 0$, there exists an interval $\theta_T \in [0, \theta_{T,min}]$ over which (C.16) is satisfied but (C.17) is not. The minimum admissible value, $\theta_{T,min}$, is computed from

$$\sin\phi = \frac{\sqrt{3}}{16} \frac{(35\sin\theta_{T,min} + 14\sin 5\theta_{T,min} - 5\sin 7\theta_{T,min})}{\cos^5\theta_{T,min}(11 - 10\cos 2\theta_{T,min})} \tag{C.18}$$

which may be obtained by solving either $-\frac{\alpha_2}{2\beta_2} = -\sin 3\theta_T$ or $-\frac{\alpha_2^2}{4\beta_2} + \gamma_2 = 0$ (the functions appearing in (C.16) and (C.17), respectively). It is therefore concluded that there exists at least one point (corresponding to the minimum) on the interval $\theta \in (-30^\circ, -\theta_T)$ at which the function on the left-hand side of (C.11) becomes negative. For $\theta_T \geqslant \theta_{T,min}$, the minimum must occur at an endpoint since (C.16) is not satisfied, in which case (C.11) is satisfied as previously demonstrated.

Combining results for $\theta \geqslant 0$ and $\theta \leqslant 0$, we conclude that the yield surface is convex for $|\theta| \geqslant \theta_T$ provided $\theta_T \geqslant \theta_{T,min}$, where $\theta_{T,min}$ is given by (C.18).

References

Abbo, A.J., Sloan, S.W., 1995. A smooth hyperbolic approximation to the Mohr–Coulomb yield criterion. Comput. Struct. 54, 427–441.
Clausen, J., Damkilde, L., Andersen, L., 2006. Efficient return algorithms for associated plasticity with multiple yield surfaces. Int. J. Numer. Methods Eng. 66, 1036–1059.
Crisfield, M.A., 1991. Non-linear Finite Element Analysis of Solids and Structures, vol. 1. Wiley, Chichester.
Gens, A., Carol, I., Alonso, E.E., 1990. A constitutive model for rock joints: formulation and numerical implementation. Comput. Geotech. 9, 3–20.
Koiter, W.T., 1953. Stress-strain relations, uniqueness and variational theorems for elastic-plastic materials with a singular yield surface. Q. Appl. Math. 11, 350–354.
Nayak, G.C., Zienkiewicz, O.C., 1972. Convenient form of stress invariants for plasticity. J. Struct. Div. ASCE. 98, 949–954.
Owen, D.R.J., Hinton, E., 1980. Finite Elements in Plasticity: Theory and Practice. Pineridge Press, Swansea.
Ristinmaa, M., Tryding, J., 1993. Exact integration of constitutive equation in elasto-plasticity. Int. J. Numer. Methods Eng. 36, 2525–2544.

Sloan, S.W., Booker, J.R., 1986. Removal of singularities in Tresca and Mohr–Coulomb yield functions. Commun. Appl. Numer. Methods 2, 173–179.
Sloan, S.W., Abbo, A.J., Sheng, D.C., 2001. Refined explicit integration of elastoplastic models with automatic error control. Eng. Comput. 18, 121–154, Erratum: 2002, Eng. Comput., 19, 594–594.
Taiebat, H.A., Carter, J.P., 2008. Flow rule effects in the Tresca model. Comput. Geotech. 35, 500–503.
van Eekelen, H.A.M., 1980. Isotropic yield surfaces in three dimensions for use in soil mechanics. Int. J. Numer. Anal. Methods Geomech. 4, 89–101.
Zienkiewicz, O.C., Pande, G.N., 1997. Some useful forms of isotropic yield surfaces for soil and rock mechanics. In: Gudehus, G. (Ed.), Finite Elements in Geomechanics. Wiley, Chichester, pp. 179–190.