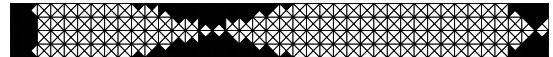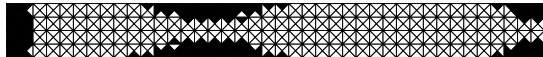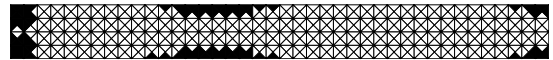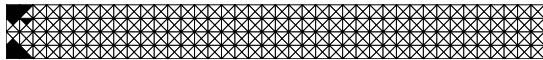# BASIC COMPUTATIONAL PLASTICITY

KRISTIAN KRABBENHØFT

Department of Civil Engineering
Technical University of Denmark

June 2002

Contents

1 Introduction

In structural mechanics the most common type of analysis is the linear static analysis where the displacements are assumed infinitesimal and Hooke's law is assumed valid. However, at high load levels a number of nonlinearities may be encountered. One of the most common of these is material nonlinearity where the stress-strain relation is nonlinear. Other common nonlinearities include those arising from significant changes in the geometry during loading. Although both material and geometric nonlinearities arise as a result of the structure being subjected to forces of a significant magnitude they can and should be separated and described independently of each other. Thus, it is easy to imagine situations where the displacements are so large that the equilibrium equations must be formulated in a coordinate system which follows the structure, while at the same time the material remains linear elastic following Hooke's law. This could be the case for slender structures such as masts and antennae. On the other hand, the situation where the displacements are sufficiently small to allow for the equilibrium equations to be formulated in the original coordinate system, but where the stress-strain relation is nonlinear are just as common. As examples, the simple tension and compression tests of many engineering materials such as concrete, steel and soil produce stress-strain curves which are nonlinear while the displacements are still 'small'. The same could be observed for many real structures composed of these materials.

The most challenging problems are of course those involving a combination of the above mentioned nonlinearities and possibly others. Car crash simulation is such a problem. Here the displacements are large while at the same time the stress levels are so high that the stress-strain relation is nonlinear. Additionally, inertial and contact forces must be taken into consideration. Another related problem is so-called push-over analysis of off shore platforms. These types of problems almost always involve some sort of material nonlinearity, and thus, material nonlinearity can be viewed both as one of the most common types of nonlinearities arising in real structures as well as a de facto prerequisite for more advanced problems such as the above mentioned.

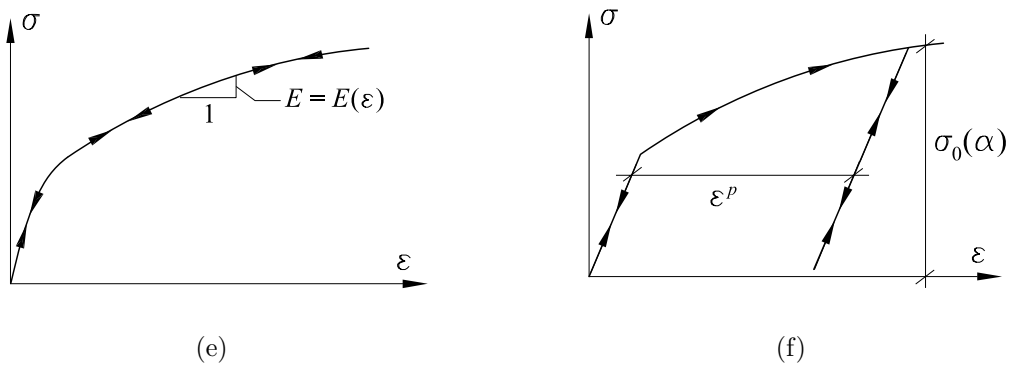(e)   (f)

Figure 1: Material nonlinearity: nonlinear elasticity (a) and elasto-plasticity (b).
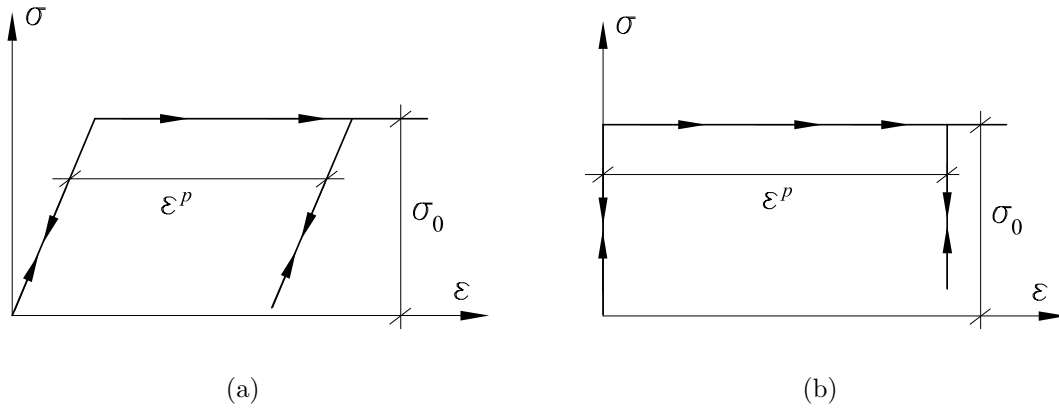
Figure 2: Material models: Linear elastic-perfectly plastic (a) and rigid-perfectly plastic (b).

1.1 Plasticity

Material nonlinearity itself may be subdivided into some fundamentally different categories. In nonlinear elasticity the stress-strain relation is nonlinear but otherwise the behaviour follows that of linear elasticity, that is, no distinction is made between loading and unloading except for the sign. This is in contrast to what is that case with plastic or elasto-plastic materials, where irreversible strains occur. The difference between two such materials is illustrated in Figure 1. For low stress levels both materials follow a linear stress-strain relation. This is followed by a decrease in stiffness as the stress increases. If however, the stress is reduced the nonlinear elastic material will follow the same stress-strain curve as in loading, whereas unloading of the elasto-plastic material leads to a new branch on the $\sigma - \varepsilon$ curve where the material is again elastic, often with a stiffness equal to the initial elastic stiffness. Furthermore, it is clear that when the material is completely unloaded, an irreversible plastic strain ε^p remains.

The curved part above the elastic limit in Figure 1 (b) suggests that the stress-strain relation may be of a rather complicated nature and in general this may indeed be true. However, in practice, a number of approximations can be made. One of the most common
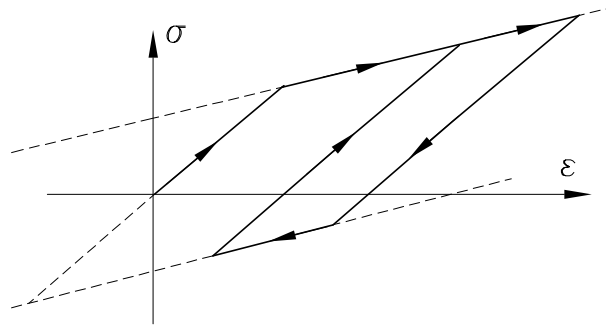
Figure 3: Hardening plasticity: Bauschinger effect.

4

approximations is to assume linear elasticity below some limit above which the material is perfectly plastic, i.e. the stress remains constant. This is shown in Figure 2 (a). Another common approximation is the rigid-perfectly plastic material model. Here the elastic strains are ignored altogether and deformations then take place only when the stress reaches a certain level, Figure 2 (b).

Alternatively, a certain degree of hardening may be introduced such that yield stress $\sigma_0(\alpha)$, now depending on one or more hardening parameters, is continuously increased as the plastic loading progresses. One of the most common types of hardening is described by the so-called Baushcinger effect which can be observed in metals. This model involves an increase in the yield stress under plastic loading while at the same time the negative limit is lifted such that the permissible stress range remains constant. This is illustrated in Figure 3.

For a more complete discussion of the physical foundations of plasticity we refer to Chen and Han [1].

2 Plasticity Theory

The theory of plasticity makes use of some fundamental concepts: the yield criterion defining the limit at which the material becomes plastic, the flow rule describing the relationship between stresses and strains once the material has become plastic, and the consistency condition which prevents stresses from exceeding the yield limit.

In metal plasticity the theory necessary for describing plastic flow is particularly simple since metals are generally incompressible, insensitive to the influence of hydrostatic pressure and furthermore, with very good approximation, follow so-called associated flow rules. For other materials such a rocks, soils, concrete, and other granular materials the conditions are somewhat more complicated, but nevertheless capable of being modeled within the framework of plasticity theory. In the following the basic plasticity theory is described and examples demonstrating the above mentioned differences are given.

2.1 The yield criterion

As described in the Section 1 an elastic limit exists above which the material becomes plastic. In the uniaxial stress test it is easy to define this limit by the uniaxial yield stress σ_0. However, when several stress components are present and acting simultaneously, the question of when the material becomes plastic is less easily answered. Generally, a scalar yield function is defined as

$$f(\sigma_x, \sigma_y, \sigma_z, \tau_{xy}, \tau_{yz}, \tau_{zx}, \alpha_1, ..., \alpha_n) = 0 \tag{1}$$

or

$$f(\boldsymbol{\sigma}, \boldsymbol{\alpha}) = 0 \tag{2}$$

where $\boldsymbol{\sigma}$ contains the six independent stress components and $\boldsymbol{\alpha}$ defines n specific material parameters. The yield function (2) defines a surface in stress space, in Figure 4 illustrated in the two-dimensional case. The convention is to define the yield function such that $f < 0$ inside the yield surface and $f > 0$ outside it. The boundary $f = 0$ defines the elastic limit, i.e. the set of maximal permissible stresses. Thus, the state of stress is given
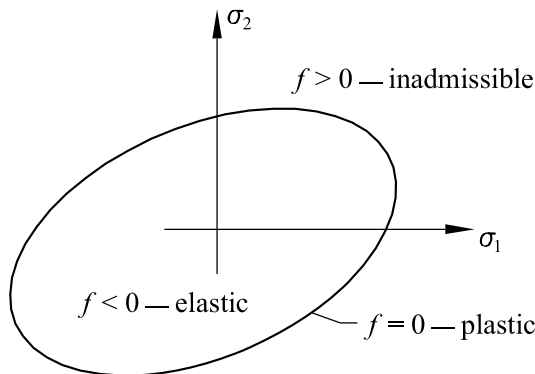

Figure 4: Yield surface.

6

by a point either inside or on the yield surface. In plastic loading with perfect plasticity the state of stress can only be altered by a redistribution between the different stress components such that the stress point can be imagined as sliding along the yield surface. If hardening is present the stress point will still remain on the boundary defined by $f = 0$ but this boundary will be shifted according to the relevant hardening rules as the loading progresses. In unloading the stress point moves from the boundary of the yield surface to the inside and thus, immediately recovers its elastic properties.

2.2 Specific yield criteria

When defining the yield criterion care must be taken to ensure that a rotation of the coordinate system does not influence the conditions at which the material yields. For the uniaxial stress test a yield criterion may be formulated as

$$\sigma_x - \sigma_0 = 0 \tag{3}$$

This criterion, however, only makes sense under the assumption that that the material is aligned in a coordinate system where all other stress components than σ_x are equal to zero. Therefore, it is convenient to define yield criteria in terms of certain invariants, that is, quantities which are not affected by a rotation of the coordinate system. These could be the magnitude of the principal stresses. Thus, rather than using (1) it is more convenient to write the yield criterion as

$$f(\sigma_1, \sigma_2, \sigma_3, \alpha_1, ..., \alpha_n) = 0 \tag{4}$$

where $\sigma_1, \sigma_2, \sigma_3$ are the principal stresses. In many cases, however, it turns out to be even more convenient to use another set of invariants, namely the invariants of the stress tensor I_1, I_2 and I_3 and the invariants of the deviatoric stress tensor J_1, J_2 and J_3, see e.g. Chen and Han [1]. Furthermore, of these invariants it turns out that especially I_1 and J_2 are useful. The invariants are obtained by separating the state of stress into two components as

$$\sigma_{ij} = \begin{bmatrix} \sigma_1 - p & 0 & 0 \\ 0 & \sigma_1 - p & 0 \\ 0 & 0 & \sigma_1 - p \end{bmatrix} + \begin{bmatrix} p & 0 & 0 \\ 0 & p & 0 \\ 0 & 0 & p \end{bmatrix} \tag{5}$$

or

$$\sigma_{ij} = s_{ij} + p\delta_{ij} \tag{6}$$

where the mean stress is defined as

$$p = \tfrac{1}{3}(\sigma_1 + \sigma_2 + \sigma_3) = \tfrac{1}{3}(\sigma_x + \sigma_y + \sigma_z) = \tfrac{1}{3}I_1 \tag{7}$$

The second invariant of s_{ij} is given by

$$\begin{aligned} J_2 &= \tfrac{1}{6}\left[(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2\right] \\ &= \tfrac{1}{6}\left[(\sigma_x - \sigma_y)^2 + (\sigma_y - \sigma_z)^2 + (\sigma_z - \sigma_x)^2\right] + \tau_{xy}^2 + \tau_{yz}^2 + \tau_{zx}^2 \end{aligned} \tag{8}$$
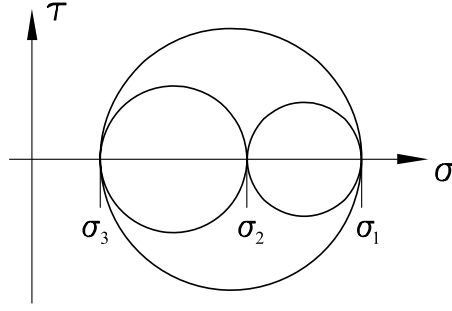
7

Figure 5: Mohr's circles for triaxial state of stress.

Whereas $p\delta_{ij}$ represents a state of hydrostatic stress the deviatoric part s_{ij} can be shown to represent a state of pure shear, and thus, the effects on yielding of shear and hydrostatic pressure can be effectively separated. Furthermore, the principal shear stresses are given by, see Figure 5,

$$\tau_{12} = \tfrac{1}{2}|\sigma_1 - \sigma_2|, \qquad \tau_{23} = \tfrac{1}{2}|\sigma_2 - \sigma_3|, \qquad \tau_{31} = \tfrac{1}{2}|\sigma_3 - \sigma_1| \qquad (9)$$

which, except for the factor $\tfrac{1}{2}$ are the terms contained in J_2. Thus, yield criteria making use if J_2 actually predict yielding to occur as a result of some combination of the principal shear stresses exceeding a certain limit.

In the following, both examples of yield criteria developed directly from the principal stresses as well as from I_1 and J_2 are given.

2.2.1 The Tresca criterion

In 1864, after having subjected metal specimens to a combination of stresses, Tresca proposed that yielding occurs as a result of the maximal shear stress reaching a critical value. This was formulated as

$$
\begin{aligned}
\tau_{max} &= \max(\tfrac{1}{2}|\sigma_1 - \sigma_2|, \tfrac{1}{2}|\sigma_2 - \sigma_3|, \tfrac{1}{2}|\sigma_3 - \sigma_1|) \\
&= \max(\tau_{12}, \tau_{23}, \tau_{31}) = k
\end{aligned}
\qquad (10)
$$

where the material parameter k can be determined by the simple tension test as

$$k = \tfrac{1}{2}\sigma_0 \qquad (11)$$

In plane stress one of the principal stresses, say σ_3, is equal to zero, and the yield surface may then be plotted in $\sigma_1 - \sigma_2$ space as shown in Figure 6. The yield surface is defined by six different expressions: When both σ_1 and σ_2 are greater than zero the limiting conditions are simply

$$\tfrac{1}{2}|\sigma_1| = \tfrac{1}{2}\sigma_1 = k \quad \text{(line AB)} \qquad (12)$$

and

$$\tfrac{1}{2}|\sigma_2| = \tfrac{1}{2}\sigma_2 = k \quad \text{(line BC)} \qquad (13)$$
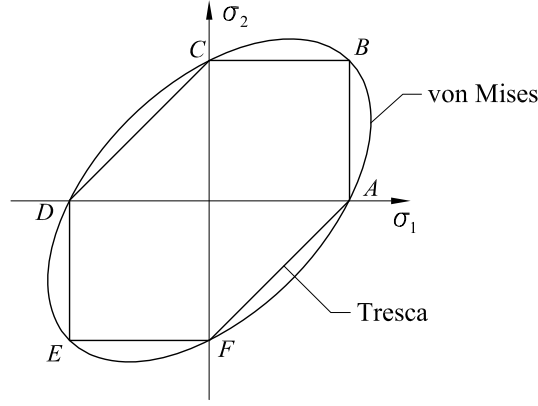
8

Figure 6: Tresca and von Mises criteria in plane stress.

Similarly, when both σ_1 and σ_2 are negative, we have

$$\tfrac{1}{2}|\sigma_1| = -\tfrac{1}{2}\sigma_1 = k \quad \text{(line DE)} \tag{14}$$

and

$$\tfrac{1}{2}|\sigma_2| = -\tfrac{1}{2}\sigma_2 = k \quad \text{(line EF)} \tag{15}$$

When $\sigma_1 < 0$ and $\sigma_2 > 0$ the stresses are limited by

$$\tfrac{1}{2}|\sigma_1 - \sigma_2| = \tfrac{1}{2}(\sigma_2 - \sigma_1) = k \quad \text{(line CD)} \tag{16}$$

and finally when $\sigma_1 > 0$ and $\sigma_2 < 0$ the criterion for yielding is

$$\tfrac{1}{2}|\sigma_1 - \sigma_2| = \tfrac{1}{2}(\sigma_1 - \sigma_2) = k \quad \text{(line CD)} \tag{17}$$

2.2.2 The von Mises criterion

Although Tresca's yield criterion does predict the yielding of metals quite well, an inherent weakness is that only the maximal shear stress is considered whereas the influence of the two lesser shear stresses is ignored. A more accurate criterion taking the all three principal shear stresses into account is that of von Mises. In this case the yield function is given by

$$f(J_2) = \sqrt{J_2} - k \tag{18}$$

or in terms of the so-called equivalent stress σ_e as

$$f(\boldsymbol{\sigma}) = \sigma_e - \sigma_0 \tag{19}$$

where

$$\begin{aligned}
\sigma_e &= \left(\tau_{12}^2 + \tau_{23}^2 + \tau_{31}^2\right)^{\frac{1}{2}} \\
&= \left[\tfrac{1}{2}(\sigma_1 - \sigma_2)^2 + \tfrac{1}{2}(\sigma_2 - \sigma_3)^2 + \tfrac{1}{2}(\sigma_3 - \sigma_1)^2\right]^{\frac{1}{2}} \\
&= \left[\tfrac{1}{2}(\sigma_x - \sigma_y)^2 + \tfrac{1}{2}(\sigma_y - \sigma_z)^2 + \tfrac{1}{2}(\sigma_z - \sigma_x)^2 + 3\tau_{xy}^2 + 3\tau_{yz}^2 + 3\tau_{zx}^2\right]^{\frac{1}{2}}
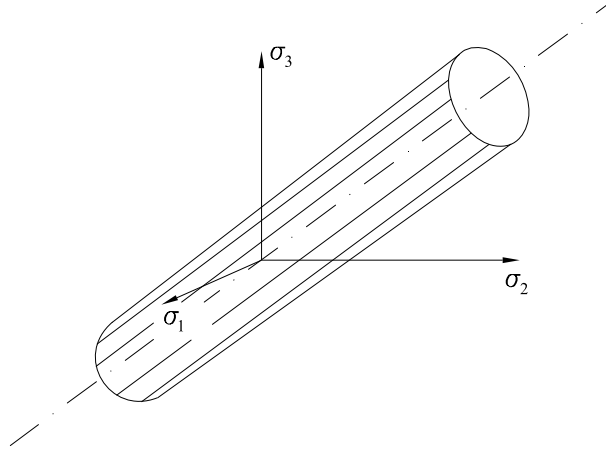\end{aligned} \tag{20}$$

9

Figure 7: von Mises criterion in principal stress space.

The similarities to the Tresca criterion are evident, and when plotted in plane stress, Figure 6, the difference is seen to be rather small. When the principal stress are identical or when one is equal to zero, the two criteria coincide. In contrast to the Tresca criterion, von Mises' criterion predicts that the value of one principal stress may exceed σ_0 provided that the other principal stress is adjusted accordingly.

As with the Tresca criterion, the von Mises criterion is independent of hydrostatic pressure. When plotted in three dimensional principal stress space the yield surface depicts a cylinder parallel to the hydrostatic axis $\sigma_1 = \sigma_2 = \sigma_3$ as shown in Figure 7.

2.2.3 The Drucker-Prager criterion

Whereas hydrostatic pressure independence is a realistic assumption for metals, it fails for other materials such as concrete and soils. Therefore Drucker and Prager formulated a modified von Mises criterion by adding a mean stress term

$$f(I_1, J_2) = \sqrt{J_2} + \alpha I_1 - k \tag{21}$$

or

$$f(\boldsymbol{\sigma}) = \sigma_e + \alpha \sigma_m - \sigma_0 \tag{22}$$

where the mean stress is given by

$$\sigma_m = I_1 = \tfrac{1}{3}(\sigma_1 + \sigma_2 + \sigma_3) = \tfrac{1}{3}(\sigma_x + \sigma_y + \sigma_z) \tag{23}$$

This criterion is illustrated in Figure 8. Compared to the von Mises criterion, there is a limit for positive (tensile) mean stresses whereas the material is strengthened by superposition of a negative (compressive) mean stress.

2.3 Loading/unloading conditions

As already touched upon earlier the yield condition defines not only the set of permissible stresses, but also the conditions for which plastic deformations can continue to occur.
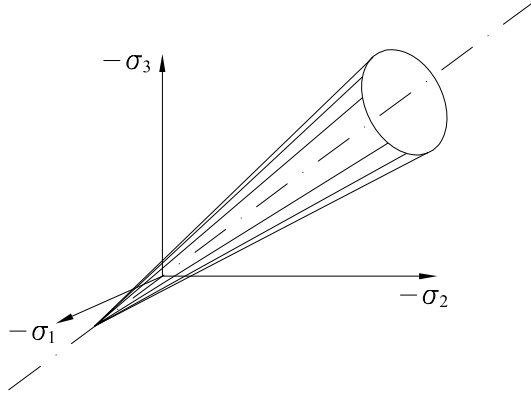
Figure 8: Drucker-Prager criterion in principal stress space.

Whereas all elastic stress states are located inside the yield surface and defined uniquely by the elastic strain, plastic deformations can occur as long as the stress point is located on the yield surface. Thus, during loading in perfect plasticity, a stress point may remain in one fixed position on the yield surface or slide along it with redistribution of stresses among the different components. Mathematically, the conditions for plastic loading can be written as

$$f(\boldsymbol{\sigma} + d\boldsymbol{\sigma}) = f(\boldsymbol{\sigma}) + \nabla f^T d\boldsymbol{\sigma} = 0 \tag{24}$$

where

$$\nabla f = [\partial f/\partial \sigma_x, ..., \partial f/\partial \tau_{zx}]^T \tag{25}$$

is the normal to the yield surface and $d\boldsymbol{\sigma}$ is a stress increment, see Figure 9. Since $f(\boldsymbol{\sigma}) = 0$

$$df = \nabla f^T d\boldsymbol{\sigma} = 0 \tag{26}$$

In other words, (26) states that during plastic loading the change in stress, if any, occurs tangential to the yield surface. This is the so-called consistency condition which, as shall be shown later, is a key ingredient in the general theory.

In unloading the state of stress immediately becomes elastic which can be written as

$$df = \nabla f^T d\boldsymbol{\sigma} < 0 \tag{27}$$

Figure 9: Plastic loading and unloading.

Figure 10: Plastic strain increment.

That is, the angle between the normal to the yield surface and the stress increment must be greater than 90°.

2.4 Plastic strains

When a structure is loaded from an neutral state elastic strains first appear. The relationship between these and the stresses is given by e.g. Hooke's law or similar, possibly nonlinear, elastic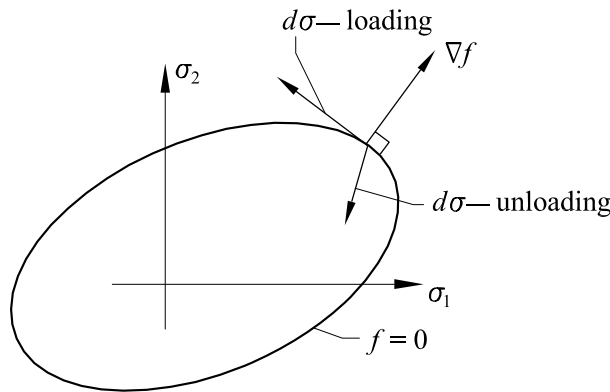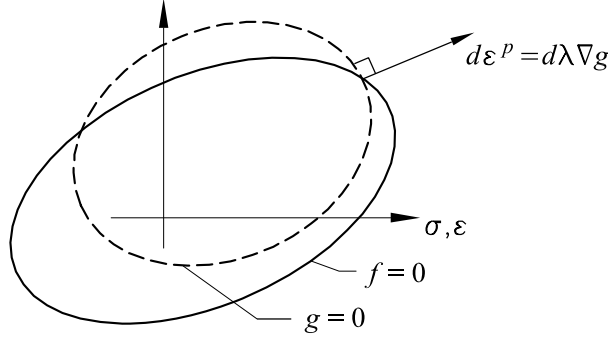 stress-strain relations. As the loading increases and the material begins to yield plastic strains will develop. At some point the load carrying capacity of the structure becomes exhausted, e.g. for frame structures when a sufficient number of plastic hinges has been formed, and at this point the plastic strains become infinitely large, corresponding to collapse of the structure. Because of this, and because, as mentioned in Section 1, there is no unique relation between stresses and total strains in the plastic range, it is convenient to consider only plastic strain increments[1] $d\boldsymbol{\varepsilon}^p$. The total strain increment is assumed to be the sum of the elastic strains increment and the plastic strain increment

$$d\boldsymbol{\varepsilon} = d\boldsymbol{\varepsilon}^e + d\boldsymbol{\varepsilon}^p \tag{28}$$

where as mentioned the elastic strain increment is related to the stresses by e.g. Hooke's law. As for the plastic strain increment, we can without loss of generality assume that this can be derived from a plastic potential g as

$$d\boldsymbol{\varepsilon}^p = d\lambda \frac{\partial g}{\partial \boldsymbol{\sigma}} = d\lambda \nabla g \tag{29}$$

where $d\lambda$ is a positive scalar of proportionality. Similarly to the yield criterion $f = 0$, the equation $g = 0$ can be thought of as describing a surface in stress space. Since ∇g is a vector normal to this surface, the strain increment can be plotted as a vector normal to the surface with a length determined by $d\lambda$, see Figure 10.

Generally, the expression for g must be determined experimentally similarly to the way the yield function is established. This is, however, quite a demanding task, and as a first

---

[1]The term strain rate and the symbol $\dot{\boldsymbol{\varepsilon}}^p$ is also commonly used

guess at g it would be reasonable to assume that $g = f$. The plastic strain increment is then determined as

$$d\boldsymbol{\varepsilon}^p = d\lambda \frac{\partial f}{\partial \boldsymbol{\sigma}} = d\lambda \nabla f \tag{30}$$

This is the so-called associated flow rule, in contrast to the nonassociated flow rule (29) where the strains are not connected with the yield criterion. Use of the yield criterion to derive the plastic strain increments is also refered to as the normality rule. The use of the associated flow rule is a cornerstone in the so-called mathematical theory of plasticity which was formulated around 1950 by among others Hill, Drucker and Prager. This mathematical theory of plasticity contains some very attractive results such as the necessity of a convex yield surface and the existence of the limit theorems which have been used extensively in engineering computations. Moreover, the physical interpretation of the results seem to comply so well with the basic thermodynamical requirement of energy conservation that the associated flow rule in some circles has almost the status of a fundamental law of nature. It can be shown, however, that this is not the case, and although very attractive, associated plasticity often fails to describe the experimentally observed results for other materials than metals.

For the yield criteria discussed previously the plastic strain increments can now be determined using the associated flow rule. For the von Mises criterion the plastic strain increments are

$$d\boldsymbol{\varepsilon}^p = \begin{bmatrix} d\varepsilon_x^p \\ d\varepsilon_y^p \\ d\varepsilon_z^p \\ d\gamma_{xy}^p \\ d\gamma_{yz}^p \\ d\gamma_{zx}^p \end{bmatrix} = d\lambda \frac{1}{2\sigma_e} \begin{bmatrix} 2\sigma_x - \sigma_y - \sigma_z \\ 2\sigma_y - \sigma_z - \sigma_x \\ 2\sigma_z - \sigma_x - \sigma_y \\ 6\tau_{xy} \\ 6\tau_{yz} \\ 6\tau_{zx} \end{bmatrix} \tag{31}$$

where $d\lambda$ again determines the magnitude of the increment.
Given a particular state of strain the relative volume change can be determined as

$$\frac{\Delta V}{V} = (1 + \varepsilon_x)(1 + \varepsilon_y)(1 + \varepsilon_z) - 1 \simeq \varepsilon_x + \varepsilon_y + \varepsilon_z \tag{32}$$

and thus, the associated von Mises flow rule predicts that no volumetric changes occur as a result of plastic straining,[2] which for metals is in good agreement with what can be observed experimentally.
In contrast, soils, concrete and other granular materials do exhibit a volumetric dilatation during plastic flow. This is reflected in the Drucker-Prager criterion where the plastic

---

[2]This is not to be confused with the total volumetric change. Usually the elastic contribution $\varepsilon_x^e + \varepsilon_y^e + \varepsilon_z^e$ will be different from zero.

strain increments are given by

$$d\boldsymbol{\varepsilon}^p = \begin{bmatrix} d\varepsilon_x^p \\ d\varepsilon_y^p \\ d\varepsilon_z^p \\ d\gamma_{xy}^p \\ d\gamma_{yz}^p \\ d\gamma_{zx}^p \end{bmatrix} = d\lambda \left(\frac{1}{2\sigma_e} \begin{bmatrix} 2\sigma_x - \sigma_y - \sigma_z \\ 2\sigma_y - \sigma_z - \sigma_x \\ 2\sigma_z - \sigma_x - \sigma_y \\ 6\tau_{xy} \\ 6\tau_{yz} \\ 6\tau_{zx} \end{bmatrix} + \begin{bmatrix} \frac{1}{3}\alpha\sigma_x \\ \frac{1}{3}\alpha\sigma_y \\ \frac{1}{3}\alpha\sigma_z \\ 0 \\ 0 \\ 0 \end{bmatrix} \right) \tag{33}$$

Here the relative change in volume is

$$\frac{\Delta V}{V} = d\varepsilon_x^p + d\varepsilon_y^p + d\varepsilon_z^p = d\lambda \tfrac{1}{3}\alpha(\sigma_x + \sigma_y + \sigma_z) \tag{34}$$

which is not necessarily equal to zero. For soils, however, the volumetric dilatation predicted by the associated Drucker-Prager flow rule is often somewhat larger than can be verified experimentally. Therefore, a nonassociated flow rule can be used, i.e. the elastic limit is still defined by the Drucker-Prager criterion whereas the flow rule is defined by some other function, e.g.

$$f(\boldsymbol{\sigma}) = \sqrt{J_2} + \alpha I_1 - k \tag{35}$$

$$g(\boldsymbol{\sigma}) = \sqrt{J_2} + \beta I_1 - k \tag{36}$$

where β should be smaller than α. The situation is illustrated in Figure 11. Another
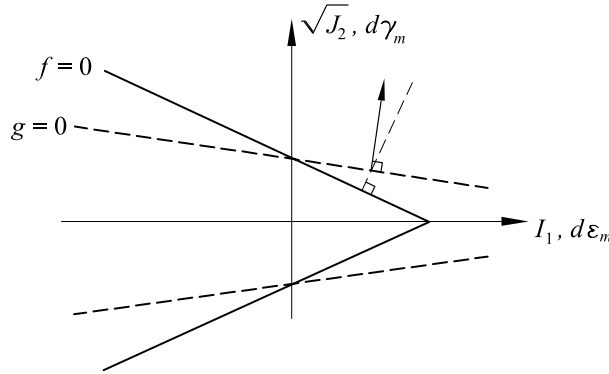
Figure 11: Drucker-Prager plasticity with nonassociated flow rule.

example of nonassociated plasticity can be illustrated by an example from elementary mechanics, namely the sliding on a rigid block on a rigid frictional surface as shown in Figure 12 (a). When the dragging force exceeds the friction force the block begins to slide. That is, at the instant of displacement the forces are related by

$$f = \tau - \mu\sigma = 0 \tag{37}$$

where μ is the coefficient of friction. This problem can be interpreted as a plasticity problem with the possibility of yielding in a thin layer in the interface between the block
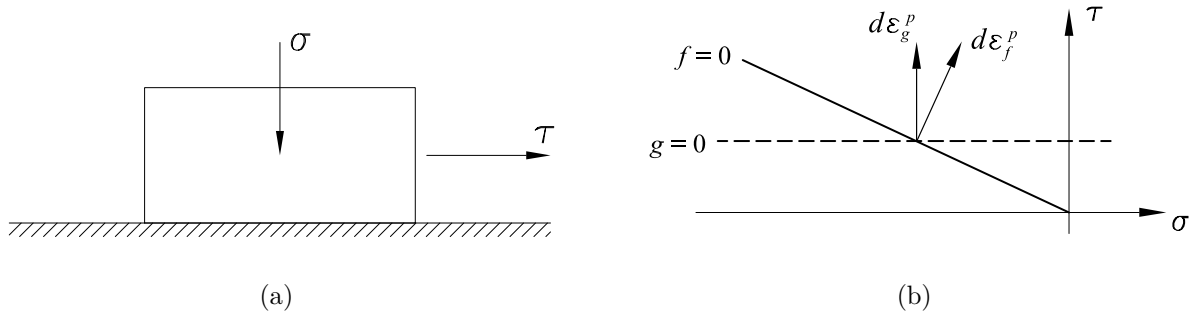
Figure 12: Rigid block on frictional surface (a) and corresponding yield criterion and flow rule (b).

Figure 13: Mechanisms corresponding to associated (a) and nonassociated (b) flow rules.

and the surface. The yield criterion is given by (37) and shown in Figure 12 (b). The 'collapse mechanism' can now be determined by computing the plastic strains[3]. If an associated flow rule is used the strain vector $d\varepsilon_f^p$ will be normal the the yield surface $f = 0$ with both a normal and a shear strain component. The corresponding mechanism is shown in Figure 13 (a). Due to the presence of a normal strain component a volume change occurs in the block-surface interface. This is in contrast to what would be expected. Rather than the block-surface interface layer expanding a mechanism as shown in Figure 13 (b) would be expected. This mechanism corresponds to zero normal strain, and the flow rule g should then be as shown in 12 (b).

2.5 Incremental stress-strain relations

In computational elasto-plastic analysis one usually proceeds by applying a load increment which produces a displacement increment, and thus a total strain increment. The stress increment corresponding to this increment in total strain can be determined by a constitutive relation similar to the relation one has in elasticity, i.e.

$$d\boldsymbol{\sigma} = \boldsymbol{D}^{ep} d\boldsymbol{\varepsilon} \tag{38}$$

---

[3]Note that for this problem the elastic strains are assumed equal to zero corresponding to the rigid perfectly plastic material model illustrated in Figure 2 (b)

where $\boldsymbol{D}^{ep}$ is the elasto-plastic constitutive matrix. Such a relation was first derived and used in an finite element context by Zienkiewicz *et al.* [2]. The elasto-plastic constitutive matrix can be derived by considering the basic relations discussed in the foregoing.

The total strain increment is given as the sum of the elastic strain increment and the plastic strain increment

$$d\varepsilon = d\varepsilon^e + d\varepsilon^p \tag{39}$$

Hooke's law gives the relation between stresses and elastic strains as

$$\boldsymbol{\sigma} = \boldsymbol{D}\varepsilon^e = \boldsymbol{D}(\varepsilon - \varepsilon^p) \tag{40}$$

or in rate form as

$$d\boldsymbol{\sigma} = \boldsymbol{D}(d\varepsilon - d\varepsilon^p) \tag{41}$$

where $\boldsymbol{D}$ is the elastic constitutive matrix. The plastic strain increment is determined by the flow rule (29) as

$$d\varepsilon^p = d\lambda \frac{\partial g}{\partial \boldsymbol{\sigma}} \tag{42}$$

Thus, the stress increment is given by

$$d\boldsymbol{\sigma} = \boldsymbol{D}d\varepsilon - d\lambda \boldsymbol{D}\frac{\partial g}{\partial \boldsymbol{\sigma}} \tag{43}$$

The expression for the stress increment is now substituted into the consistency condition (26)

$$\left(\frac{\partial f}{\partial \boldsymbol{\sigma}}\right)^T \left(\boldsymbol{D}d\varepsilon - d\lambda \boldsymbol{D}\frac{\partial g}{\partial \boldsymbol{\sigma}}\right) = 0 \tag{44}$$

Solving this equation for the scalar $d\lambda$ one obtains

$$d\lambda = \frac{\left(\frac{\partial f}{\partial \boldsymbol{\sigma}}\right)^T \boldsymbol{D}\, d\varepsilon}{\left(\frac{\partial f}{\partial \boldsymbol{\sigma}}\right)^T \boldsymbol{D}\, \frac{\partial g}{\partial \boldsymbol{\sigma}}} \tag{45}$$

Finally, $d\lambda$ is substituted back into (43) to yield

$$d\boldsymbol{\sigma} = \boldsymbol{D}\left(d\varepsilon - \frac{\left(\frac{\partial f}{\partial \boldsymbol{\sigma}}\right)^T \boldsymbol{D}\, d\varepsilon\, \frac{\partial g}{\partial \boldsymbol{\sigma}}}{\left(\frac{\partial f}{\partial \boldsymbol{\sigma}}\right)^T \boldsymbol{D}\, \frac{\partial g}{\partial \boldsymbol{\sigma}}}\right) \tag{46}$$

This can be rearranged to give the elasto-plastic constitutive relation

$$d\boldsymbol{\sigma} = \left(\boldsymbol{D} - \frac{\boldsymbol{D}\frac{\partial g}{\partial \boldsymbol{\sigma}}\left(\frac{\partial f}{\partial \boldsymbol{\sigma}}\right)^T \boldsymbol{D}}{\left(\frac{\partial f}{\partial \boldsymbol{\sigma}}\right)^T \boldsymbol{D}\, \frac{\partial g}{\partial \boldsymbol{\sigma}}}\right) d\varepsilon \tag{47}$$

and thus, the elasto-plastic constitutive matrix introduced in (38) is

$$\boldsymbol{D}^{ep} = \boldsymbol{D} - \frac{\boldsymbol{D}\dfrac{\partial g}{\partial \boldsymbol{\sigma}}\left(\dfrac{\partial f}{\partial \boldsymbol{\sigma}}\right)^{T}\boldsymbol{D}}{\left(\dfrac{\partial f}{\partial \boldsymbol{\sigma}}\right)^{T}\boldsymbol{D}\,\dfrac{\partial g}{\partial \boldsymbol{\sigma}}} \tag{48}$$

The elasto-plastic constitutive relation defines the stress increment uniquely once the total strain increment and the current state of stress is known, whereas the a strain increment cannot be determined uniquely on the the basis of a stress increment, i.e. $\boldsymbol{D}^{ep}$ is singular. When used in finite element formulations (48) defines a nonlinear relation between stress and strain increments since the evaluation of the current stress must naturally be influenced by the magnitude of the stress increment. In this way the use of (48) leads to a classical type of finite element nonlinearity where the current state and an increment is known, but where the effect of the increment depends on the state that the increment gives rise to, i.e. an iterative procedure must be applied.

2.6   Hardening

As already discussed most materials exhibit some degree of hardening as an accompaniment to plastic straining. In general this means that the shape and size of the yield surface changes during plastic loading. This change may be rather arbitrary and extremely difficult to describe accurately. Therefore, hardening is often described by a combination of two specific types of hardening, namely isotropic hardening and kinematic hardening, see Figures 14 (a)-(b). In the following isotropic hardening related to the von Mises criterion is treated in some detail. For the von Mises criterion isotropic hardening implies an increase in the yield strength during plastic loading such that the yield criterion may be written as

$$f(\boldsymbol{\sigma}) - \sigma_0(\alpha) = 0 \tag{49}$$

(a)   (b)
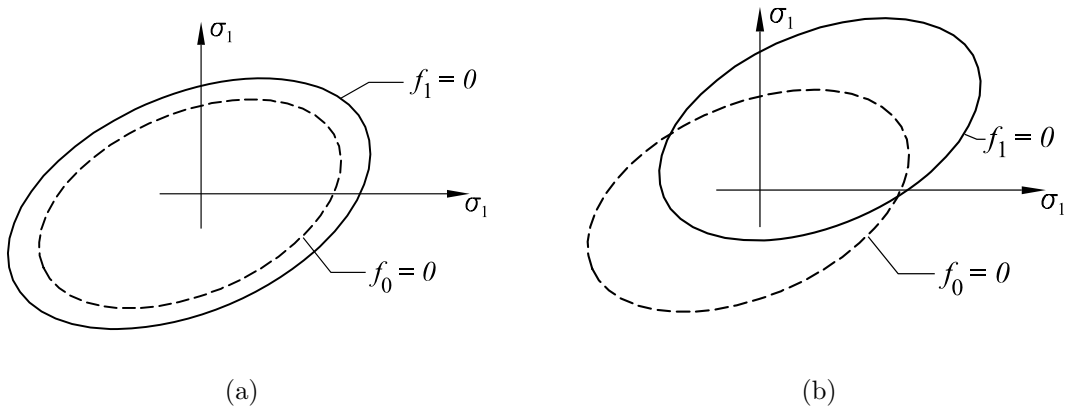
Figure 14: Isotropic (a) and kinematic (b) hardening.
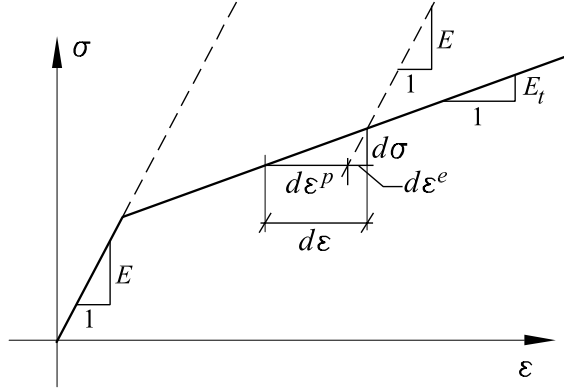
Figure 15: Uniaxial tension test with hardening.

where α is a hardening parameter. The parameter α can be calibrated from a uniaxial tension test. Suppose that from such a test a stress-strain curve as shown in Figure 15 has been measured. Considering the part of curve above the elastic limit the stress can be related to the strain by the introduction of a tangent modulus E_t as

$$d\sigma = E_t d\varepsilon \tag{50}$$

The representation of the total strain as the sum of the elastic and plastic strain leads to

$$d\sigma = E_t(d\varepsilon^e + d\varepsilon^p) \tag{51}$$

For the elastic part of the strain increment Hooke's law is used to give

$$d\sigma = E_t \left(\frac{d\sigma}{E} + d\varepsilon^p \right) \tag{52}$$

This relates the stress increment to the plastic strain increment by

$$\left(\frac{1}{E_t} - \frac{1}{E} \right) d\sigma = d\varepsilon^p \tag{53}$$

or

$$d\sigma = H d\varepsilon^p \tag{54}$$

where

$$H = \frac{E_t}{1 - E_t/E} \tag{55}$$

Assuming isotropic hardening an increase in stress above the elastic limit is equivalent to an increase in yield stress, and hence

$$\frac{d\sigma}{d\varepsilon^p} = \frac{d\sigma_0}{d\varepsilon^p} = H \tag{56}$$

This means that α in (49) should be replaced by ε^p. In the multidimensional case, however, some equivalent plastic strain must be used. For the von Mises criterion a suitable equivalent plastic strain is

$$d\varepsilon_{eq}^p = \sqrt{\tfrac{2}{3}[(d\varepsilon_1^p)^2 + (d\varepsilon_2^p)^2 + (d\varepsilon_3^p)^2]} \tag{57}$$

18

By using the associated flow rule this equivalent plastic strain can be shown be be equal to the plastic multiplier

$$d\varepsilon_{eq}^p = d\lambda \tag{58}$$

For the uniaxial tension test with $d\varepsilon_2^p = d\varepsilon_3 = -\frac{1}{2}d\varepsilon_1^p$ the equivalent plastic strain is then given by $d\varepsilon_{eq}^p = d\varepsilon_1^p = d\lambda$.

The relation (56) can now be rewritten as

$$\frac{d\sigma_0}{d\varepsilon_{eq}^p} = \frac{d\sigma_0}{d\lambda} = H \tag{59}$$

Integration of (59) yields

$$\sigma_0 = \int H d\varepsilon_{eq}^p \tag{60}$$

That is, at any given instant the yield strength is determined by the prior plastic strain history. This means that isotropic hardening is irreversible; once the material has experienced a certain degree of hardening the yield limit is shifted permanently. As this may not be in accordance with reality the isotropic hardening can be supplemented with kinematic hardening which is, however, a somewhat more demanding model to calibrate.

Isotropic hardening is easily included into the elasto-plastic stress-strain relation (47) by considering the appropriate consistency condition. For (49) this is

$$\frac{\partial f}{\partial \boldsymbol{\sigma}} d\boldsymbol{\sigma} - \frac{\partial \sigma_0}{\partial \alpha} d\alpha = 0 \tag{61}$$

or

$$\frac{\partial f}{\partial \boldsymbol{\sigma}} d\boldsymbol{\sigma} - H d\lambda = 0 \tag{62}$$

This leads to the following elasto-plastic constitutive relation

$$d\boldsymbol{\sigma} = \left(\boldsymbol{D} - \frac{\boldsymbol{D}\dfrac{\partial g}{\partial \boldsymbol{\sigma}}\left(\dfrac{\partial f}{\partial \boldsymbol{\sigma}}\right)^T \boldsymbol{D}}{H + \left(\dfrac{\partial f}{\partial \boldsymbol{\sigma}}\right)^T \boldsymbol{D}\,\dfrac{\partial g}{\partial \boldsymbol{\sigma}}} \right) d\boldsymbol{\varepsilon} \tag{63}$$

Here it is clearly seen that for H tending to infinity the usual linear elastic constitutive matrix $\boldsymbol{D}$ is recovered while $H = 0$ of course corresponds to the elastic-perfectly plastic case (47). It should be noted that while for perfect plasticity $\boldsymbol{D}^{ep}$ is singular, the introduction of hardening results in a unique relationship between infinitesimal stress and strain increments such that this singularity is removed. The inverse of $\boldsymbol{D}^{ep}$ can be found by use of the Sherman-Morrison formula which states that

$$(\boldsymbol{A} - \boldsymbol{u}\boldsymbol{v}^T)^{-1} = \boldsymbol{A}^{-1} + \frac{\boldsymbol{A}^{-1}\boldsymbol{u}\boldsymbol{v}^T\boldsymbol{A}^{-1}}{1 - \boldsymbol{v}^T\boldsymbol{A}^{-1}\boldsymbol{u}} \tag{64}$$

where $\boldsymbol{A}$ is an $n \times n$ matrix and $\boldsymbol{v}$ and $\boldsymbol{u}$ are vectors of length n. The inverse of $\boldsymbol{D}^{ep}$ can then be written as

$$(\boldsymbol{D}^{ep})^{-1} = \boldsymbol{D}^{-1} + \frac{1}{H}\frac{\partial g}{\partial \boldsymbol{\sigma}}\left(\frac{\partial f}{\partial \boldsymbol{\sigma}}\right)^T \tag{65}$$

which is clearly finite only for values of H different from zero.

2.7 Plane stress versus plane strain

The yield criteria discussed in the above were all formulated with reference to the most general triaxial stress state. Under certain circumstances, however, the number of variables can be reduced. This is for example the case in plane stress, plane strain and under axisymmetric conditions. In the following the two former states are treated.

In plane stress only in-plane stresses are considered. This means that

$$\sigma_z = \tau_{yz} = \tau_{zx} = 0 \tag{66}$$

The corresponding plane stress yield criterion is then obtained by simply deleting the above stress components from the general triaxial yield criterion. In this way von Mises' yield criterion can be written as

$$f = \sqrt{\sigma_x^2 + \sigma_y^2 - \sigma_x\sigma_y + 3\tau_{xy}^2} - \sigma_0 \tag{67}$$

Here it should be noted, however, that when computing the plastic strain increments reference must again be made to the general triaxial criterion, i.e. $d\varepsilon_z^p$, $d\gamma_{yz}^p$, $d\gamma_{zx}^p$ are not necessarily equal to zero, and the same is of course the case with respect to the elastic strains.

In plane strain we have

$$\varepsilon_z = \gamma_{yz} = \gamma_{zx} = 0 \tag{68}$$

If this is to be valid for all strains it must hold that

$$d\varepsilon_z^e = d\gamma_{yz}^e = d\gamma_{zx}^e = 0 \tag{69}$$

and

$$d\varepsilon_z^p = d\gamma_{yz}^p = d\gamma_{zx}^p = 0 \tag{70}$$

In associated plasticity the relation between the plastic strain increments and the yield function now makes it possible to determine a two-dimensional yield criterion as in the case of plane stress. Again von Mises' criterion is considered, and according to (31) the plastic strain increments are here given by

$$
\begin{aligned}
d\varepsilon_z^p &= d\lambda \frac{1}{2\sigma_e}(2\sigma_z - \sigma_x - \sigma_y) \\
d\gamma_{yz}^p &= d\lambda \frac{1}{2\sigma_e}6\tau_{yz} \\
d\gamma_{zx}^p &= d\lambda \frac{1}{2\sigma_e}6\tau_{zx}
\end{aligned}
\tag{71}
$$

Whereas it follows directly from the plastic shear strain increments that the corresponding shear stresses should be deleted from the yield criterion, the expression for $d\varepsilon_z^p$ gives a condition for the normal stresses as

$$\sigma_z = \tfrac{1}{2}(\sigma_x + \sigma_y) \tag{72}$$

20

This is now substituted into the general von Mises criterion to yield the following plane strain criterion

$$\sqrt{\tfrac{3}{4}(\sigma_x - \sigma_y)^2 + 3\tau_{xy}^2} - \sigma_0 = 0 \tag{73}$$

If $\sigma_y = \tau_{xy} = 0$ the uniaxial yield strength is

$$\sigma_x^{max} = \tfrac{2}{\sqrt{3}}\sigma_0 \simeq 1.15\sigma_0 \tag{74}$$

The deviation between this value and what is obtained in plane stress stems from the fact that σ_z is different from zero, in fact, (72),

$$\sigma_z = \tfrac{1}{2}(\sigma_x + \sigma_y) = \tfrac{1}{\sqrt{3}}\sigma_0 \tag{75}$$

In this way σ_z is 'chosen' as having an optimal value such that the yield strength becomes as large as possible. Similarly, if $\sigma_x = \sigma_y$ and $\tau_{xy} = 0$ the magnitude of σ_x and σ_y is unlimited, corresponding to σ_z being chosen such that a hydrostatic pressure appears. That σ_z is always chosen in an optimal way can be verified by realizing that the solution to the problem

$$\max_{\sigma_z} \sigma_e \tag{76}$$

is found in exactly the same way as by use of the associated flow rule, namely by setting the derivative of σ_e with respect to σ_z equal to zero and solving for σ_z.

3   Finite element implementation

The implementation of the elasto-plastic constitutive relations in a finite element context requires the consideration of two different levels, in the following called the global level and the material level respectively. On the global level equilibrium must be satisfied as in any other linear or nonlinear finite element computation, whereas on the material level the plasticity relations must be satisfied.

Performing the finite element discretization in the usual way, see e.g. Zienkiewicz and Taylor [3], one ends up with the following equation system

$$\boldsymbol{K}\boldsymbol{u} = \boldsymbol{f} \tag{77}$$

where $\boldsymbol{f}$ is the applied force and $\boldsymbol{K}$ is the stiffness matrix given by

$$\boldsymbol{K} = \int_{\Omega} \boldsymbol{B}^T \hat{\boldsymbol{D}}^{ep} \boldsymbol{B} \, d\Omega \tag{78}$$

Since the constitutive matrix $\hat{\boldsymbol{D}}^{ep}$ depends on the current state of stress the stiffness matrix is nonlinear. Moreover, since the response of the system depends of the deformation history, an incremental relationship between displacement and force is needed, i.e.

$$\boldsymbol{K}_t \Delta \boldsymbol{u} = \Delta \boldsymbol{f} \tag{79}$$

Here $\boldsymbol{K}_t$ is the tangent stiffness matrix, i.e. a linearization of the exact nonlinear stiffness matrix (78). The computations now proceed by applying a load load increment $\Delta \boldsymbol{f}$ and computing the corresponding displacement increment from (79). This is illustrated in Figure 16.

The strain increment is computed in the usual way as

$$\Delta \boldsymbol{\varepsilon} = \boldsymbol{B} \Delta \boldsymbol{u} \tag{80}$$

Next, the stresses are computed. This is done via the elasto-plastic constitutive relation (48)

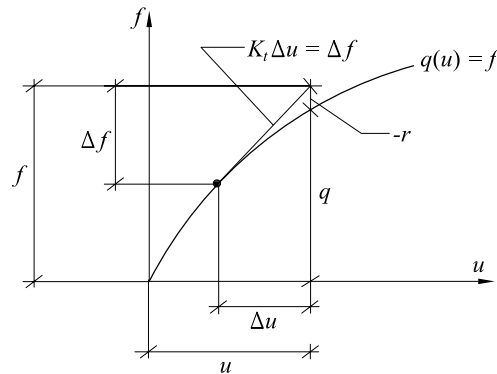$$d\boldsymbol{\sigma} = \boldsymbol{D}^{ep} d\boldsymbol{\varepsilon} \tag{81}$$

Figure 16: Load step in elasto-plastic computation

This is nonlinear relation since $\boldsymbol{D}^{ep}$ depends on the current stress state and generally, iterative procedures must thus be used. These are treated in detail in the following Section. Assuming that the stresses have been computed the internal force vector can then be found as

$$q = \int_\Omega \boldsymbol{B}^T \boldsymbol{\sigma} \, d\Omega \tag{82}$$

This must be balanced by the total applied load, i.e. the residual

$$\boldsymbol{r} = \boldsymbol{q} - \boldsymbol{f} \tag{83}$$

must vanish. If the residual is different from zero it is applied as an external load following the well-known Newton-Raphson procedure. This then gives a new strain increment and a corresponding new stress increment which must be determined via the nonlinear elasto-plastic constitutive relation (48), a new residual is computed and so on until the residual becomes sufficiently small. The procedure can be sketched as follows

1. Apply load increment Δf and find displacement and strain increments $\Delta \boldsymbol{u}$ and $\Delta \boldsymbol{\varepsilon}$.

2. Determine stress increment $\Delta \boldsymbol{\sigma}$ from (48).

3. Compute residual $\boldsymbol{r}$.

4. If $||\boldsymbol{r}|| > \text{TOL}$ set $\Delta \boldsymbol{f} = -\boldsymbol{r}$ and return to 1.

Thus, the computation of a load step requires a global iterative procedure where the out of balance force, or residual, must vanish (steps 1 to 4), as well as a procedure to compute the stress increments (step 2). The latter of these can of course be modified and substituted independent of the iterations on the global level, and in contrast to the global equilibrium iterations the stress update is performed independently in each Gauss point. In the following a number of procedures to compute the stress increment given a strain increment are discussed.

3.1 Integration of the constitutive relations

3.1.1 Explicit integration

The equation (48) gives the relation between infinitesimal stress and strain increments as

$$d\boldsymbol{\sigma} = \boldsymbol{D}^{ep}(\boldsymbol{\sigma})d\boldsymbol{\varepsilon} \tag{84}$$

Considering finite increments $\Delta \boldsymbol{\sigma}$ and $\Delta \boldsymbol{\varepsilon}$ we wish to integrate (84) such that the stress increment is found as

$$\int_{\boldsymbol{\sigma}_j}^{\boldsymbol{\sigma}_j+\Delta\boldsymbol{\sigma}} d\boldsymbol{\sigma} = \Delta\boldsymbol{\sigma} = \int_{\boldsymbol{\varepsilon}_j}^{\boldsymbol{\varepsilon}_j+\Delta\boldsymbol{\varepsilon}} \boldsymbol{D}^{ep}d\boldsymbol{\varepsilon} \tag{85}$$

where index j refers to the initial state. The problem with this integration is that $\boldsymbol{D}^{ep}$ depends on the current stress. Therefore, an approximate solution procedure has to be employed. The simplest possible approximation to (85) is

$$\Delta\boldsymbol{\sigma} = \boldsymbol{D}^{ep}(\boldsymbol{\sigma}_j)\Delta\boldsymbol{\varepsilon} \tag{86}$$
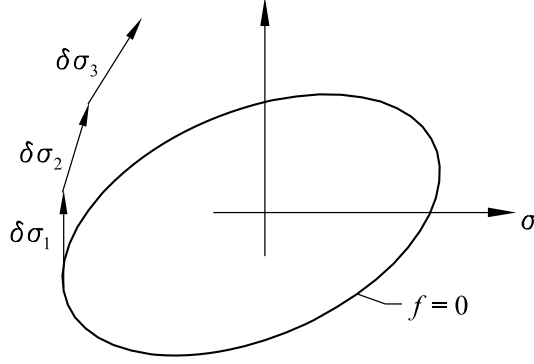
Figure 17: Explicit integration of the constitutive relation.
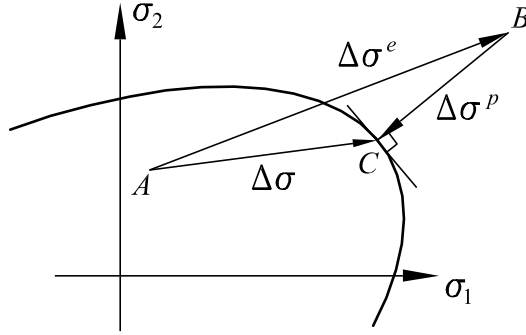
Figure 18: Basic principle of stress update.

The new stresses are then updated as

$$\boldsymbol{\sigma}_{j+1} = \boldsymbol{\sigma}_j + \Delta\boldsymbol{\sigma} \qquad (87)$$

Alternatively, the strain increment could be divided into a number of equal subincrements

$$\delta\boldsymbol{\varepsilon} = \frac{\Delta\boldsymbol{\varepsilon}}{N} \qquad (88)$$

and the stresses would then be updated as

$$\delta\boldsymbol{\sigma} = \boldsymbol{D}^{ep}(\boldsymbol{\sigma}_j)\delta\boldsymbol{\varepsilon}, \qquad \boldsymbol{\sigma}_{j+1} = \boldsymbol{\sigma}_j + \delta\boldsymbol{\sigma} \qquad (89)$$

If the number of subincrements N is significant this of course improves the accuracy over (86) but there is still the possibility that errors may accumulate and the stresses drift away from the yield surface, thus violating the yield condition as illustrated in Figure 17.

3.1.2 Return algorithms

One way of avoiding this is, instead of operating directly on $\boldsymbol{D}^{ep}$, to solve step by step the equations involved in the derivation of $\boldsymbol{D}^{ep}$. The basic equation needed is the expression for the stress increment given a total strain increment, (43),

$$d\boldsymbol{\sigma} = \boldsymbol{D}d\boldsymbol{\varepsilon} - d\lambda\boldsymbol{D}\frac{\partial g}{\partial \boldsymbol{\sigma}} \qquad (90)$$

This expression is now integrated to yield

$$\Delta\boldsymbol{\sigma} = \boldsymbol{D}\Delta\boldsymbol{\varepsilon} - \int_{\lambda}^{\lambda+\Delta\lambda} \boldsymbol{D}\frac{\partial g}{\partial\boldsymbol{\sigma}}d\lambda \tag{91}$$

or

$$\Delta\boldsymbol{\sigma} = \Delta\boldsymbol{\sigma}^e - \Delta\boldsymbol{\sigma}^p \tag{92}$$

where

$$\Delta\boldsymbol{\sigma}^e = \boldsymbol{D}\Delta\boldsymbol{\varepsilon} \qquad \text{and} \qquad \Delta\boldsymbol{\sigma}^p = \int_{\lambda}^{\lambda+\Delta\lambda} \boldsymbol{D}\frac{\partial g}{\partial\boldsymbol{\sigma}}d\lambda \tag{93}$$

Here $\Delta\boldsymbol{\sigma}^e$ is usually called the elastic predictor and $\Delta\boldsymbol{\sigma}^p$ the plastic corrector. Equation (91) implies an integration of the stresses along a path from the initial stress point to the final stress point. Since this path is of course unknown an approximation to the plastic corrector has to be used. This approximation may be written as

$$\Delta\boldsymbol{\sigma}^p \simeq \Delta\lambda\boldsymbol{D}\left(\frac{\partial g}{\partial\boldsymbol{\sigma}}\right)_P \tag{94}$$

where index P denotes a point on the integration path. The question is now how this point should be chosen. An obvious choice is the final stress point as shown in Figure 18, such that (91) is written as

$$\Delta\boldsymbol{\sigma} = \boldsymbol{D}\Delta\boldsymbol{\varepsilon} - \Delta\lambda\boldsymbol{D}\left(\frac{\partial g}{\partial\boldsymbol{\sigma}}\right)_C \tag{95}$$

Thus, from a point A lying inside or on the yield surface an elastic predictor is applied. This leads to a state of stress outside the yield surface, point B. To fulfill the yield condition a plastic corrector is applied, thus returning the stresses to the yield surface, point C. The plastic corrector is determined by two quantities, the scalar $\Delta\lambda$ giving the magnitude, and the gradient of the loading surface $(\partial g/\partial\boldsymbol{\sigma})_C$ giving the direction. The magnitude, i.e. $\Delta\lambda$, is determined such that the yield condition is fulfilled

$$f(\boldsymbol{\sigma}_C - \Delta\boldsymbol{\sigma}^p) = 0 \tag{96}$$

where $\Delta\boldsymbol{\sigma}^p$ is given by

$$\Delta\boldsymbol{\sigma}^p = \Delta\lambda\boldsymbol{D}\left(\frac{\partial g}{\partial\boldsymbol{\sigma}}\right)_C \tag{97}$$

Considering a Taylor expansion of the yield function around point B we have

$$f(\boldsymbol{\sigma}_C) = f(\boldsymbol{\sigma}_B - \Delta\boldsymbol{\sigma}^p) \simeq f(\boldsymbol{\sigma}_B) - \left(\frac{\partial f}{\partial\boldsymbol{\sigma}}\right)_B^T \Delta\boldsymbol{\sigma}^p = 0 \tag{98}$$

Inserting the expression for $\Delta\boldsymbol{\sigma}^p$ into this and solving for $\Delta\lambda$ gives a step size of

$$\Delta\lambda = \frac{f(\boldsymbol{\sigma}_B)}{\left(\dfrac{\partial f}{\partial\boldsymbol{\sigma}}\right)_B^T \boldsymbol{D}\left(\dfrac{\partial g}{\partial\boldsymbol{\sigma}}\right)_C} \tag{99}$$

If the gradients of f and g are constant between points B and C it is possible to obtain a closed form solution. This is the case for circular yield surfaces such as the von Mises criterion in triaxial stress, where the method was first used by Krieg and Krieg [4] and named the radial return algorithm. If the gradients are not constant along the line connecting B and C the same general principles still apply, but iterations must now be performed. The procedure is outlined in Table 1. It should be noticed that the requirement that $\Delta\boldsymbol{\sigma}^p$ be a vector between points B and C and normal to the loading surface in point C is not fulfilled, see Figure 19.

3.1.3 Implicit methods

In the precious Section it was demonstrated how, if the gradients of f and g are not constant between points B and C, the application of the radial return method leads to violation of the condition that total plastic correction increment is normal to the loading surface. This requirement can, however, be met if a full Newton-Raphson iterative procedure is applied. The equations to be solved are

$$\boldsymbol{p} = \Delta\boldsymbol{\sigma} - \boldsymbol{D}\Delta\boldsymbol{\varepsilon} + \boldsymbol{D}\Delta\lambda\nabla g(\boldsymbol{\sigma}_A + \Delta\boldsymbol{\sigma}) = \boldsymbol{0} \tag{100}$$

Initial point $\boldsymbol{\sigma}^0$

Iterations $j = 0, \ldots, n$

$\quad \Delta\lambda = \dfrac{f(\boldsymbol{\sigma}^j)}{((\partial f/\partial\boldsymbol{\sigma})^j)^T \boldsymbol{D}(\partial g/\partial\boldsymbol{\sigma})^j}$

$\quad \Delta\boldsymbol{\sigma}^p = \Delta\lambda\boldsymbol{D}(\partial g/\partial\boldsymbol{\sigma})^j$

$\quad \boldsymbol{\sigma}^{j+1} = \boldsymbol{\sigma}^j - \Delta\boldsymbol{\sigma}^p$

Until $f(\boldsymbol{\sigma}^{j+1}) < \epsilon$

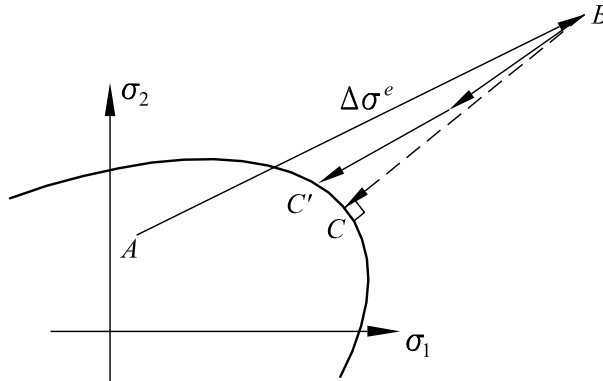---

Table 1: Radial return algorithm.

Figure 19: Radial return method for non-constant gradient between points B and C.

where ∇g is the gradient of g. Furthermore, the yield condition must be satisfied

$$q = f(\boldsymbol{\sigma}_A + \Delta\boldsymbol{\sigma}) = 0 \tag{101}$$

These two nonlinear equations are now solved using Newton's method. With this method a set of nonlinear equations

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0} \tag{102}$$

is solved iteratively as

$$\boldsymbol{J}^j \Delta\boldsymbol{x}^j = -\boldsymbol{f}(\boldsymbol{x}^j), \qquad \boldsymbol{x}^{j+1} = \boldsymbol{x}^j + \Delta\boldsymbol{x}^{j+1} \tag{103}$$

where $\boldsymbol{J}$ is the Jacobian of $\boldsymbol{f}$

$$\boldsymbol{J} = \begin{bmatrix} \partial f_1/\partial x_1 & \cdots & \partial f_1/\partial x_n \\ \vdots & \ddots & \vdots \\ \partial f_n/\partial x_1 & \cdots & \partial f_n/\partial x_n \end{bmatrix} = \nabla\boldsymbol{f}^T \tag{104}$$

and j refers to the iteration number.
Thus, the solution of (100)-(101) is realized as

$$\begin{bmatrix} \boldsymbol{I} + \Delta\lambda\boldsymbol{D}\nabla^2 g & \boldsymbol{D}\nabla g \\ \nabla f^T & 0 \end{bmatrix}^j \begin{bmatrix} \delta\boldsymbol{\sigma}^j \\ \delta\lambda^j \end{bmatrix} = \begin{bmatrix} -\boldsymbol{p}^j \\ -q^j \end{bmatrix} \tag{105}$$

$$\begin{aligned} \Delta\boldsymbol{\sigma}^{j+1} &= \Delta\boldsymbol{\sigma}^j + \delta\boldsymbol{\sigma}^j \\ \Delta\lambda^{j+1} &= \Delta\lambda^j + \delta\lambda^j \end{aligned} \tag{106}$$

where it is seen that the curvature of the loading surface is now also taken into consideration by the term $\nabla^2 g$.
If the iterations are started from $\Delta\boldsymbol{\sigma}^0 = \boldsymbol{D}\Delta\boldsymbol{\varepsilon}$ and $\Delta\lambda = 0$ the first increment must satisfy

$$\begin{bmatrix} \boldsymbol{I} & \boldsymbol{D}\nabla g(\boldsymbol{\sigma}_B) \\ \nabla f(\boldsymbol{\sigma}_B)^T & 0 \end{bmatrix}^0 \begin{bmatrix} \delta\boldsymbol{\sigma}^0 \\ \delta\lambda^0 \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ -f(\boldsymbol{\sigma}_B) \end{bmatrix} \tag{107}$$

The solutions to this system is

$$\delta\lambda^0 = \frac{f(\boldsymbol{\sigma}_B)}{\nabla f(\boldsymbol{\sigma}_B)^T \boldsymbol{D}\nabla g(\boldsymbol{\sigma}_B)}, \qquad \delta\boldsymbol{\sigma}^0 = -\delta\lambda^0 \boldsymbol{D}\nabla g(\boldsymbol{\sigma}_B) \tag{108}$$

which corresponds exactly to what was obtained with the radial return algorithm. Thus, if the first order Taylor approximation of the residuals is exact, i.e. the gradients of f and g are constant, only one Newton step is required. The situation where more than one step is needed is depicted in Figure 20. As shown, subincrements $\delta\boldsymbol{\sigma}$ and $\delta\lambda$ are computed in each iteration after which the physical stress increment $\Delta\boldsymbol{\sigma}$ is updated. At convergence the stress state is updated as $\boldsymbol{\sigma}_C = \boldsymbol{\sigma}_A + \Delta\boldsymbol{\sigma}$, and the final stress state is thus independent of the path defined by the subincrements.
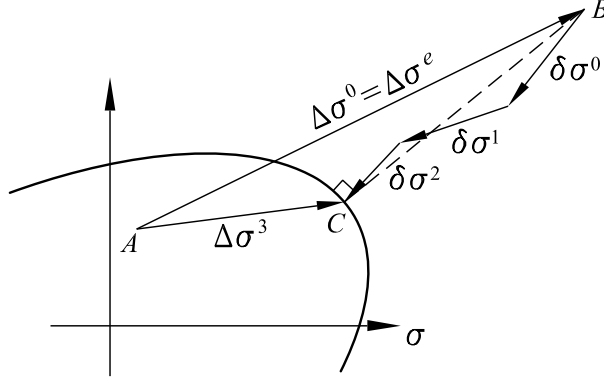
Figure 20: Implicit return algorithm.

3.1.4 Accuracy of return algorithms

The next question which arises is to how accurate the above algorithm is in relation to the exact incremental elasto-plastic constitutive stress-strain relation (47).

The basic principle was to approximate the infinitesimal stress-strain relation

$$d\boldsymbol{\sigma} = \boldsymbol{D}(d\boldsymbol{\varepsilon} - d\boldsymbol{\varepsilon}^p) \tag{109}$$

in terms of finite increments

$$\Delta\boldsymbol{\sigma} = \boldsymbol{D}(\Delta\boldsymbol{\varepsilon} - \Delta\boldsymbol{\varepsilon}^p) = \Delta\boldsymbol{\sigma}^e - \Delta\lambda\boldsymbol{D}\left(\frac{\partial g}{\partial\boldsymbol{\sigma}}\right)_P \tag{110}$$

In the previous Section the point P was taken as the final point C. However, a more refined choice is to represent the gradient by some average across the increment. That is, if yielding has already occurred once before, the gradient should be taken as an average between the previously established point and the new unknown point. Alternatively, if we start from an elastic state the gradient used could be an average between the gradient in the intersection between the yield surface and the elastic predictor and the gradient in the final stress state. This can be expressed as

$$\nabla g_{av} = (1 - \theta)\nabla g(\boldsymbol{\sigma}_{A'}) + \theta\nabla g(\boldsymbol{\sigma}_C) \tag{111}$$

where $\boldsymbol{\sigma}_C = \boldsymbol{\sigma}_A + \Delta\boldsymbol{\sigma}$ is the point to be found and $\boldsymbol{\sigma}_{A'}$ is a previously established point. Clearly, $\theta = 1$ corresponds to the algorithm discussed in the above. Alternatively, the gradient may computed using an average stress over the increment, i.e.

$$\nabla g_{av} = \nabla g\left[(1 - \theta)\boldsymbol{\sigma}_{A'} + \theta(\boldsymbol{\sigma}_C)\right] \tag{112}$$

This, however, corresponds to evaluation of the gradient in a point which lies inside the yield surface. Again, $\theta = 1$ corresponds to the first implicit algorithm. Both averaging schemes are easily implemented in an iterative procedure similar to the one shown above. Experience has shown that for small i-ncrements $\theta \simeq 0.5$ gives the highest accuracy, while for larger increments θ should be chosen as being slightly higher. The average gradient
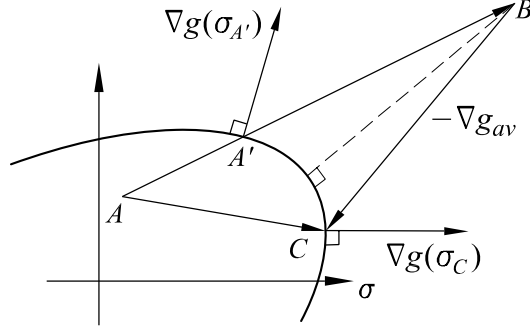
Figure 21: Return along average gradient. Dashed line corresponds to $\theta = 1$

rule (111) is illustrated in Figure 21. It should be noted that in contrast to the case where $\theta = 1$ it is necessary to determine the intersection between the yield surface and the equivalent elastic stress vector. This can in some cases be done analytically or else requires an iterative procedure in itself.

3.2 Solution of nonlinear finite element equations

Since the elasto-plastic constitutive matrix $\boldsymbol{D}^{ep}$ depends on the current state of stress implementation of plasticity in a finite element context generate nonlinear systems of equations. These can be solved by the methods discussed in the following.

3.2.1 Newton's method

Newton's method[4], which has already been discussed, is the most basic method for the solution of nonlinear equations. The method starts by considering a linearization of the equations around the current point after which an approximate solution to the linearized system is computed. This leads to a solution which does not necessarily satisfy the nonlinear equations but rather generates a residual. With Newton's method this is accounted for by once again linearizing the equations, now around the new point after which the procedure is repeated. As a example, consider the scalar equation of one variable

$$f(x) = 0 \tag{113}$$

We will define the residual as

$$r = f(x) \tag{114}$$

implying that when the solution is found r vanishes. Starting from a point x^0 a linearization is made

$$f(x) \simeq f(x^0) + f'(x^0)(x - x^0) = 0 \tag{115}$$

or

$$f(x) \simeq r(x^0) + f'(x^0)\Delta x^0 = 0 \tag{116}$$

---

[4]Newton's method after Isaac Newton (1643-1727) who first described the method is also commonly referred to as the Newton-Raphson method after Joseph Raphson (1648-1715) who first published it.
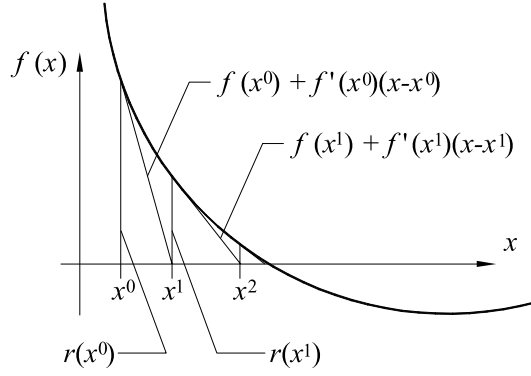
Figure 22: Newton's method in one dimension

where $\Delta x^0 = x - x^0$. The solution to this equation is

$$\Delta x^0 = -(f'(x^0))^{-1} r(x^0) \tag{117}$$

and the current point is then updated as

$$x^1 = x^0 + \Delta x^0 \tag{118}$$

A new residual is computed and the procedure repeated until convergence. The method is illustrated in Figure 22. In multiple dimensions the solution to a set of equations

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0} \tag{119}$$

where $\boldsymbol{f} = [f_1(\boldsymbol{x}), ..., f_n(\boldsymbol{x})]^T$ and $\boldsymbol{x} = [x_1, ..., x_n]^T$ is required. Here the method can be generalized such that the iterative solution takes the form

$$\Delta \boldsymbol{x}^j = -\boldsymbol{J}^{-1} \boldsymbol{r}(\boldsymbol{x}^j), \quad \boldsymbol{x}^{j+1} = \boldsymbol{x}^j + \Delta \boldsymbol{x}^j \tag{120}$$

where the Jacobian is given in (104). A thorough discussion of the limitations and possible extensions to Newton's method, both in one as well as in several dimensions, can be found in Nash and Sofer [5].

Turning the attention to the nonlinear finite element equations the end goal is here to find solutions satisfying equilibrium. That is, the force resultant of the internal stresses must balance the externally applied loads. This is expressed as

$$\boldsymbol{r}_{k+1} = \boldsymbol{q}_{k+1} - \boldsymbol{f}_{k+1} = \boldsymbol{0} \tag{121}$$

where index k refers to an already established equilibrium state such that $k+1$ is the new equilibrium state that we wish to find. The internal force vector $\boldsymbol{q}^{k+1}$ may be written as

$$\boldsymbol{q}_{k+1} = \boldsymbol{q}(\boldsymbol{u}_k + \Delta \boldsymbol{u}_k) \tag{122}$$

and the external force vector as

$$\boldsymbol{f}_{k+1} = \boldsymbol{f}_k + \Delta \boldsymbol{f}_k \tag{123}$$

30

where $\Delta \boldsymbol{f}_k$, in contrast to $\Delta \boldsymbol{u}_k$, is considered known.

Applying Newton's method we have

$$\left(\frac{d\boldsymbol{q}}{d(\Delta \boldsymbol{u}_k)}\right)_j^T \delta \boldsymbol{u}_k^j = -\boldsymbol{q}(\boldsymbol{u}_k + \Delta \boldsymbol{u}_k^j) + (\boldsymbol{f}_k + \Delta \boldsymbol{f}_k) \tag{124}$$

where index k refers to the load step number, and j is the iteration number. For convenience in notation the indices on $\Delta \boldsymbol{u}$ are temporarily dropped in the following.

The Jacobian, or tangent stiffness matrix, is determined as

$$\begin{aligned}
\boldsymbol{K}_t &= \left(\frac{d\boldsymbol{q}}{d(\Delta \boldsymbol{u})}\right)^T \\
&= \int_\Omega \boldsymbol{B}^T \left(\frac{d(\boldsymbol{\sigma}(\boldsymbol{u}_k + \Delta \boldsymbol{u}))}{d(\Delta \boldsymbol{u})}\right)^T d\Omega \\
&= \int_\Omega \boldsymbol{B}^T \left(\frac{d(\boldsymbol{\sigma}_k + \Delta \boldsymbol{\sigma})}{d(\Delta \boldsymbol{u})}\right)^T d\Omega \\
&= \int_\Omega \boldsymbol{B}^T \left(\frac{d(\Delta \boldsymbol{\sigma})}{d(\Delta \boldsymbol{u})}\right)^T d\Omega \\
&= \int_\Omega \boldsymbol{B}^T \left(\frac{d(\Delta \boldsymbol{\sigma})}{d(\Delta \boldsymbol{\varepsilon})}\right)^T \boldsymbol{B} d\Omega
\end{aligned} \tag{125}$$

From the incremental elasto-plastic constitutive relation we have

$$d\boldsymbol{\sigma} = \boldsymbol{D}^{ep} d\boldsymbol{\varepsilon} \tag{126}$$

or since $d\boldsymbol{\sigma} = d(\boldsymbol{\sigma}_k + \Delta \boldsymbol{\sigma}) = d(\Delta \boldsymbol{\sigma})$ and $d\boldsymbol{\varepsilon} = d(\boldsymbol{\varepsilon}_k + \Delta \boldsymbol{\varepsilon}) = d(\Delta \boldsymbol{\varepsilon})$

$$d(\Delta \boldsymbol{\sigma}) = \boldsymbol{D}^{ep} d(\Delta \boldsymbol{\varepsilon}) \tag{127}$$

and thus,

$$\boldsymbol{K}_t = \int_\Omega \boldsymbol{B}^T \boldsymbol{D}^{ep} \boldsymbol{B} d\Omega \tag{128}$$

However, in determining the stress increment corresponding to a given strain increment at the material level the elasto-plastic constitutive matrix was not used directly; in fact, it may be argued that $\boldsymbol{D}^{ep}$ is an unnecessary quantity. Rather, an elastic predictor was applied followed by a plastic corrector such that the consistency condition was satisfied. Using this principle a slightly different stiffness matrix consistent with the material point stress updating can be determined, see Section 3.3.

Rewriting (124) we have

$$\boldsymbol{K}_t^{k,j} \delta \boldsymbol{u}_k^j = -\boldsymbol{q}(\boldsymbol{u}_k + \Delta \boldsymbol{u}_k^j) + (\boldsymbol{f}_k + \Delta \boldsymbol{f}_k) \tag{129}$$

where again k refers to the global load step and j to the iterations. In order to start the iterations an initial value of $\Delta\boldsymbol{u}_k^0$ is needed. This is usually done by setting $\Delta\boldsymbol{u}_k^0 = \boldsymbol{0}$. In the first iteration the system to be solved is then

$$\boldsymbol{K}_t^{k,0}\delta\boldsymbol{u}_k^0 = -\boldsymbol{q}(\boldsymbol{u}_k) + (\boldsymbol{f}_k + \Delta\boldsymbol{f}_k) \tag{130}$$

Since at the last converged step k the internal force vector is balance by the external load, $-\boldsymbol{q}(\boldsymbol{u}_k) + \boldsymbol{f}_k = \boldsymbol{0}$, we have

$$\boldsymbol{K}_t^{k,0}\delta\boldsymbol{u}_k^0 = \Delta\boldsymbol{f}_k \tag{131}$$

This equation system is now solved and the displacement increment updated as

$$\Delta\boldsymbol{u}_k^1 = \boldsymbol{\Delta}bsu_k^0 + \delta\boldsymbol{u}_k^0 \tag{132}$$

The strain increment is then determined as

$$\Delta\boldsymbol{\varepsilon}_k^1 = \boldsymbol{B}\Delta\boldsymbol{u}_k^1 \tag{133}$$

With this strain increment it is now possible, by the methods discussed in Section 3.1, to compute a stress increment, and thereby the total stress $\boldsymbol{\sigma}_{k+1}^1$. This is then in turn used to compute a new internal force vector and thus a new right hand side of (129). Furthermore the tangent stiffness matrix is updated. The complete iterative procedure is outlined in Table 2. The graphical interpretation of the method for a problem with one displacement variable is shown in Figure 23 (a) where the subscripts k indicating the load step number have been dropped. In practice the so-called modified Newton-Raphson
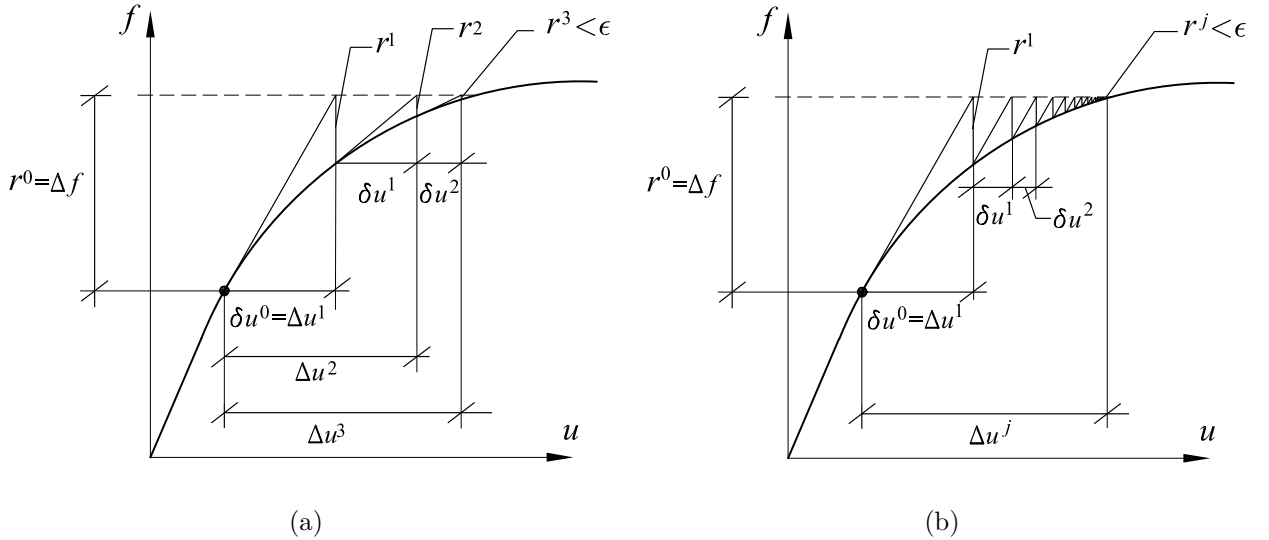

Figure 23: Solution of one-dimensional nonlinear equation by Newton-Raphson (a) and modified Newton-Raphson (b) methods.

method is often used. The modification consists of computing the tangent stiffness matrix

Initialize $\boldsymbol{u}_0 = \boldsymbol{0}, \boldsymbol{f}_0 = 0$
Load steps $k = 0, \ldots, k_{max}$
Force increment: $\Delta \boldsymbol{f}_k$
Initialize: $\Delta \boldsymbol{u}_k^0 = \boldsymbol{0}, \Delta \lambda_k^0 = 0$
Iterations $j = 0, \ldots$
Tangent stiffness: $\boldsymbol{K}_t^{k,j}$
Residual: $\boldsymbol{r}_k^j = \boldsymbol{q}(\boldsymbol{u}_k + \Delta \boldsymbol{u}_k^j) - (\boldsymbol{f}_k + \Delta \boldsymbol{f}_k)$
Solve: $\boldsymbol{K}_t^{k,j} \delta \boldsymbol{u}_k^j = -\boldsymbol{r}_k^j$
Update: $\Delta \boldsymbol{u}_k^{j+1} = \Delta \boldsymbol{u}_k^j + \delta \boldsymbol{u}_k^j$
$\Delta \boldsymbol{\varepsilon}_k^{j+1} = \boldsymbol{B} \Delta \boldsymbol{u}_k^{j+1}$
Material point iterations
Compute: $\boldsymbol{\sigma}_{k+1}^{j+1}, \Delta \lambda_k^{j+1}$
Until $\|\boldsymbol{r}_k^j\| < \epsilon \|\boldsymbol{f}_k + \Delta \boldsymbol{f}_k\|$
Update: $\boldsymbol{u}_{k+1} = \boldsymbol{u}_k + \Delta \boldsymbol{u}_k^{j+1}$
$\boldsymbol{f}_{k+1} = \boldsymbol{f}_k + \Delta \boldsymbol{f}_k$
End load step

Table 2: Newton-Raphson method for elasto-plastic problems.

only once in the beginning of each load step rather than in each iteration as shown in Figure 23 (b). If relatively small steps are used, and this is often necessary both for the full and modified schemes, the difference between the two methods is quite small, and the additional number of iterations required with the modified scheme is usually more than compensated for by avoiding the often costly procedure of computing and factorizing the tangent matrix in each iteration. The two schemes can also be combined such that if the convergence becomes too slow a switch is made to the full Newton-Raphson procedure.

It should be mentioned that the modified Newton-Raphson scheme is used almost exclusively in the field of structural mechanics. For other nonlinear problems especially those involving a time dependence such as e.g. transient nonlinear diffusion type problems it is usually necessary to employ the full Newton-Raphson procedure, and even then one may have to resort to supplementary procedures such as line search to obtain convergence.

3.2.2 General solution procedures

Starting with Newton's method as the basic procedure for solving nonlinear equations a number of more advanced methods have been developed for structural analysis. These have the common feature that in each load step not only the displacement increment but also some optimal magnitude of the load increment is found. Newton's method in the form considered in the above aims at determining a displacement increment for a fixed
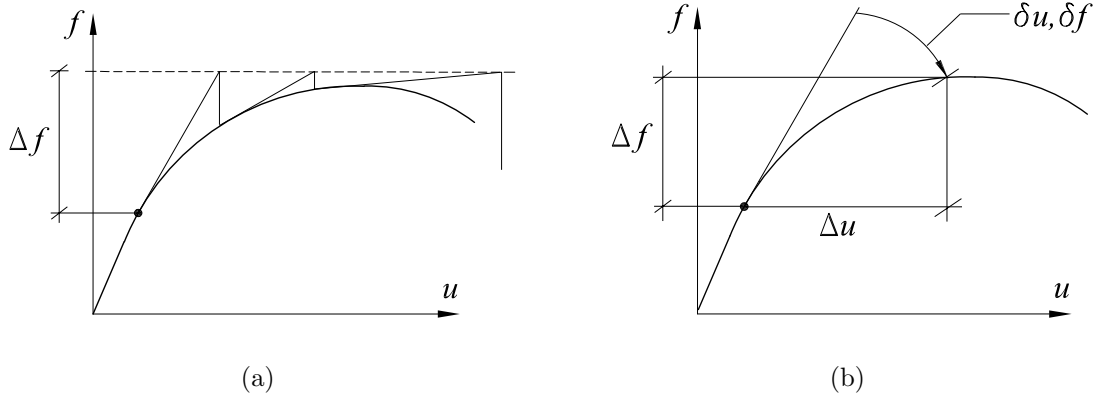
Figure 24: Failure of Newton-Raphson method (a) and general principle in simultaneous iterative adjustment of both load and displacement.

load increment such that the system is in equilibrium, that is we require the solution of

$$\boldsymbol{r} = \boldsymbol{q}(\boldsymbol{u}_k + \Delta \boldsymbol{u}_k^j) - (\boldsymbol{f}_k + \Delta \boldsymbol{f}_k) = \boldsymbol{0} \tag{134}$$

However, if it turns out to be more convenient a solution to

$$\boldsymbol{r} = \boldsymbol{q}(\boldsymbol{u}_k + \Delta \boldsymbol{u}_k^j) - (\boldsymbol{f}_k + \xi \Delta \boldsymbol{f}_k) = \boldsymbol{0} \tag{135}$$

where ξ is a positive scaling factor, would also be satisfactory. The motivation for considering such a system is two fold. First of all, with Newton's method the load continuously increases which means that if the load-displacement curve has a peak the method fails as illustrated in Figure 24 (a). Furthermore, in plasticity the load-displacement curve flattens out near collapse and in order to trace this part of the curve the load increment has to be adjusted, either manually in the beginning of each load step, or automatically in the iterations as illustrated in 24 (a). Secondly, it turns out that with some sort of load control the total number of iterations can usually be decreased quite significantly as opposed to the situation where the load increment is fixed. This feature is similar to the improvements one can obtain from using line search. Generally, Newton's method in multiple dimensions works well if the initial point is close to the solution. Otherwise convergence may be very slow or the solution may even diverge. For a system on nonlinear equations

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0} \tag{136}$$

iterative increments are found by solution of

$$\nabla \boldsymbol{f}_j^T \Delta \boldsymbol{x}^j = -\boldsymbol{f}(\boldsymbol{x}^j) \tag{137}$$

Instead of updating $\boldsymbol{x}$ in the usual way by simply adding the increment $\Delta \boldsymbol{x}^j$ to $\boldsymbol{x}^j$ it may be useful to introduce a damping parameter such that $\boldsymbol{x}$ is updated as

$$\boldsymbol{x}^{j+1} = \boldsymbol{x}^j + \xi \Delta \boldsymbol{x}^j \tag{138}$$

where the damping parameter ξ is determined on the basis of some optimality criterion ensuring that progress towards the solution is made in each iteration. A commonly used criterion states that ξ should be chosen as to minimize some norm of the residual, for example

$$\min_{\xi} \ \tfrac{1}{2}[\boldsymbol{f}(\boldsymbol{x}^j + \xi\Delta\boldsymbol{x}^j)]^T \boldsymbol{f}(\boldsymbol{x}^j + \xi\Delta\boldsymbol{x}^j) \tag{139}$$

The solution to this problem is found by setting the derivative with respect to ξ equal to zero. As an approximation a commonly applied criterion is

$$(\Delta\boldsymbol{x}^j)^T \boldsymbol{f}(\boldsymbol{x}^j + \xi\Delta\boldsymbol{x}^j) = 0 \tag{140}$$

from which ξ is determined and $\boldsymbol{x}$ updated accordingly. This condition also has the interpretation that the residual should be orthogonal to the increment.

In the current context of elasto-plasticity Krenk's orthogonal residual procedure [6] has proven to be rather efficient. Here the magnitude of the load increment is determined by a similar criterion, namely that the current residual should be orthogonal to the current displacement increment

$$(\Delta\boldsymbol{u}_k^j)^T[\boldsymbol{q}(\boldsymbol{u}_k + \Delta\boldsymbol{u}_k^j) - (\boldsymbol{f}_k + \xi\Delta\boldsymbol{f}_k)] = 0 \tag{141}$$

From this criterion ξ is determined as

$$\xi = \frac{(\Delta\boldsymbol{u}_k^j)^T[\boldsymbol{q}(\boldsymbol{u}_k + \Delta\boldsymbol{u}_k^j) - \boldsymbol{f}_k]}{(\Delta\boldsymbol{u}_k^j)^T\Delta\boldsymbol{f}_k} \tag{142}$$

This defines a new residual

$$\boldsymbol{r} = \boldsymbol{q}(\boldsymbol{u}_k + \Delta\boldsymbol{u}_k^j) - (\boldsymbol{f}_k + \xi\Delta\boldsymbol{f}_k) \tag{143}$$

which is then used to compute the increment $\delta\boldsymbol{u}_k^j$. Thus, the orthogonal residual procedure requires a minimum of extra computations compared to the classical Newton-Raphson method. In order to compute the initial value of ξ an initial displacement increment is needed. An obvious choice is

$$\Delta\boldsymbol{u}_k^0 = (\boldsymbol{K}_t^{k,0})^{-1}\Delta\boldsymbol{f}_k \tag{144}$$

corresponding to $\xi = 1$. It is also possible to define the magnitude of the initial displacement increment empirically of the basis of some reference increment and the number of iterations used in the previous load increment. More details about the orthogonal residual procedure can be found in Krenk [7, 8]. Other related methods such as Riks' arc length method are described in Crisfield [9].

3.3 Consistent tangent matrix

In Section 3.1 it was demonstrated how the updating of stresses by

$$\Delta\boldsymbol{\sigma} = \boldsymbol{D}^{ep}\Delta\boldsymbol{\varepsilon} \tag{145}$$

could lead to states of stresses not satisfying the yield condition. So rather than using (145) directly a procedure of applying an equivalent elastic stress increment followed by a plastic corrector was used. That is, the stresses were updated as

$$\Delta\boldsymbol{\sigma} = \boldsymbol{D}\Delta\boldsymbol{\varepsilon} - \Delta\lambda\boldsymbol{D}\nabla g(\boldsymbol{\sigma}_k + \Delta\boldsymbol{\sigma}) \tag{146}$$

where $\Delta\lambda$ was determined such that the new point would satisfy the yield criterion. Index k refers to a previously established point, i.e. point A in (100).

A tangent matrix based on this stress updating scheme can then be derived. To do this a function $\boldsymbol{\Pi}$ is defined as

$$\boldsymbol{\Pi} = -\Delta\boldsymbol{\sigma} + \boldsymbol{D}\Delta\boldsymbol{\varepsilon} - \Delta\lambda\boldsymbol{D}\nabla g(\boldsymbol{\sigma}_k + \Delta\boldsymbol{\sigma}) = \boldsymbol{0} \tag{147}$$

Differentiation gives

$$\begin{aligned}
d(\boldsymbol{\Pi}) &= \left(\frac{\partial\boldsymbol{\Pi}}{\partial(\Delta\boldsymbol{\varepsilon})}\right)^T d(\Delta\boldsymbol{\varepsilon}) + \left(\frac{\partial\boldsymbol{\Pi}}{\partial(\Delta\boldsymbol{\sigma})}\right)^T d(\Delta\boldsymbol{\sigma}) + \left(\frac{\partial\boldsymbol{\Pi}}{\partial(\Delta\lambda)}\right)^T d(\Delta\lambda) = \boldsymbol{0} \\
&= -d(\Delta\boldsymbol{\sigma}) + \boldsymbol{D}d(\Delta\boldsymbol{\varepsilon}) - \Delta\lambda\boldsymbol{D}\nabla^2 g(\boldsymbol{\sigma})d(\Delta\boldsymbol{\sigma}) - \boldsymbol{D}\nabla g(\boldsymbol{\sigma})d(\Delta\lambda) = \boldsymbol{0}
\end{aligned} \tag{148}$$

where $\boldsymbol{\sigma} = \boldsymbol{\sigma}_k + \Delta\boldsymbol{\sigma}$ is the final stress state. This now gives a relation between $d(\Delta\boldsymbol{\sigma})$ and $d(\Delta\boldsymbol{\varepsilon})$ as

$$[\boldsymbol{I} + \Delta\lambda\boldsymbol{D}\nabla^2 g(\boldsymbol{\sigma})]d(\Delta\boldsymbol{\sigma}) = \boldsymbol{D}d(\Delta\boldsymbol{\varepsilon}) - d(\Delta\lambda)\boldsymbol{D}\nabla g(\boldsymbol{\sigma}) \tag{149}$$

The factor $d(\Delta\lambda)$ may be determined via the consistency condition

$$\nabla f(\boldsymbol{\sigma})^T d(\Delta\boldsymbol{\sigma}) = 0 \tag{150}$$

Elimination of $d(\Delta\lambda)$ between (149) and (150) gives a relation between $d(\Delta\boldsymbol{\sigma})$ and $d(\Delta\boldsymbol{\varepsilon})$ as

$$d(\Delta\boldsymbol{\sigma}) = \boldsymbol{D}_c^{ep} d(\Delta\boldsymbol{\varepsilon}) \tag{151}$$

where

$$\boldsymbol{D}_c^{ep} = \boldsymbol{D}_c - \frac{\boldsymbol{D}_c \dfrac{\partial g}{\partial\boldsymbol{\sigma}}\left(\dfrac{\partial f}{\partial\boldsymbol{\sigma}}\right)^T \boldsymbol{D}_c}{\left(\dfrac{\partial f}{\partial\boldsymbol{\sigma}}\right)^T \boldsymbol{D}_c \dfrac{\partial g}{\partial\boldsymbol{\sigma}}} \tag{152}$$

and

$$\boldsymbol{D}_c^{-1} = \boldsymbol{D}^{-1} + \Delta\lambda\nabla^2 g(\boldsymbol{\sigma}) \tag{153}$$

Thus, the consistent tangent matrix differs only from (48) in the term $\boldsymbol{D}_c$.

In order to utilize the full potential of the consistent tangent matrix a full Newton-Raphson procedure, with update of the tangent matrix in each iteration, should be used. With the modified Newton-Raphson scheme some improvement may also be achieved provided that good estimate of $\Delta\lambda$ and the final stress state is made.

4 Example

In the following an elasto-plastic analysis of the beam structure shown in Figure 25 is given. The material is assumed to be elastic-perfectly plastic following the von Mises criterion with associated flow rule. The yield strength-stiffness ratio is $\sigma_0/E = 0.1\%$ corresponding approximately to mild steel, and Poison's ration is $\nu = 0.3$.
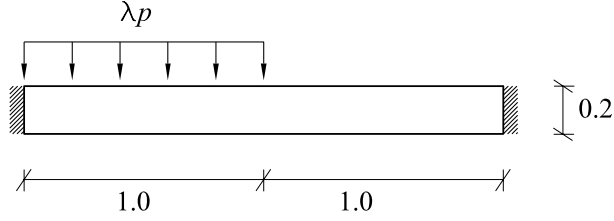

Figure 25: Beam structure.

The load is continuously increased to a point very near collapse after which the structure is unloaded completely. For the global iterations the orthogonal residual procedure is used and on the material level the stresses are updated by means of the radial return algorithm. The load-displacement curve for the vertical displacement of the center point is shown in Figure 26 where $\lambda = 1.0$ corresponds to the intensity of the load p at collapse and the displacement has been normalized with respect to the displacement the point of initial yielding. The structure first begins to yield at around $\lambda = 0.5$, that is for a load which is
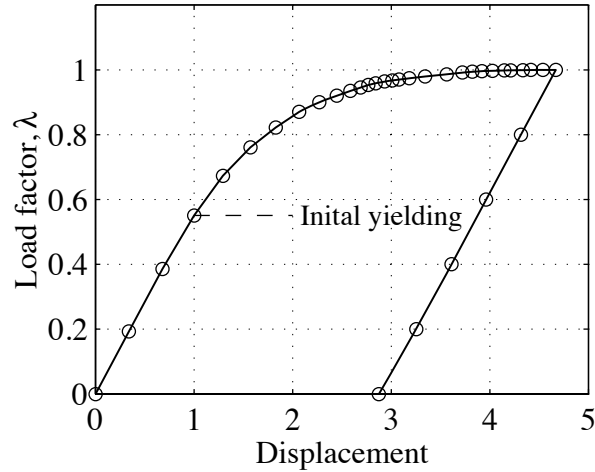

Figure 26: Load displacement curve.

approximately half the ultimate load. As the load is increased the structure begins to yield and the stiffness decreases. The evolution of the plastification is shown in Figures 27 (a)-(d). Just before the load carrying capacity becomes exhausted the structure is unloaded corresponding to the load being removed all together. As seen from the load-displacement curve the structure again reacts elastically with a stiffness corresponding to the initial

(a) Step 5: $\lambda = 0.673$

(b) Step 10: $\lambda = 0.920$

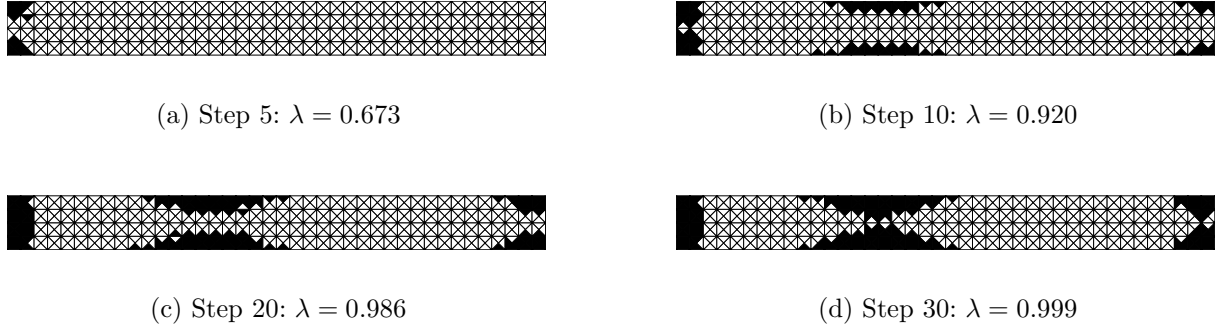(c) Step 20: $\lambda = 0.986$

(d) Step 30: $\lambda = 0.999$

Figure 27: Yielding of beam structure.

elastic stiffness. However, since the magnitude of this unloading, $\lambda = -1.0$, is twice that initially resulting in yielding there is the possibility that rather high stresses may develop upon unloading. In fact, near the left support the stresses again reach a magnitude corresponding to yielding. In Figure 28 the normal stresses in a point in the vicinity of the left support are traced. The Figure describes the same as the load-displacement curve: an initially linear elastic response followed by yielding where the ratio between the two stress components change. At unloading a path parallel to the initial elastic path is followed, and just before the load is removed completely the stresses reach a magnitude corresponding to yielding on the opposite' side of the yield surface. Thus, at the point where the load is zero, significant stresses still exist within the structure. These are of course in self-equilibrium, i.e. in equilibrium with zero external load.
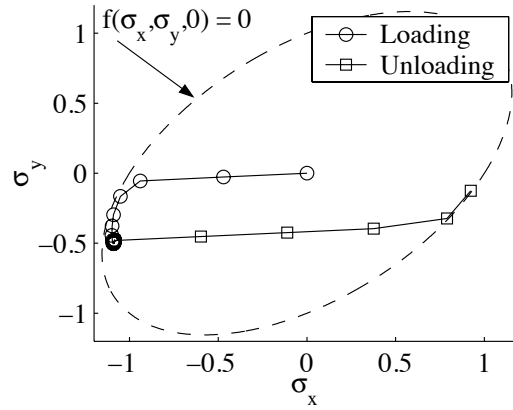

Figure 28: Normal stresses at point near left support.

5 Solutions of nonlinear FE–equations

5.1 Implementation of modified Newton-Raphson method for elasto-plasticity

$\boldsymbol{u}_1 = \boldsymbol{0}, \boldsymbol{\sigma}_1 = \boldsymbol{0}, \boldsymbol{f}_1 = \boldsymbol{0}$	*initial state*
`for k = 1:kmax`	*begin load increments*
$\quad \boldsymbol{K}_t = \texttt{tangentstiff}(\boldsymbol{\sigma}_k)$	
$\quad \Delta\boldsymbol{u} = \boldsymbol{K}_t^{-1}\Delta\boldsymbol{f}_k$	
$\quad$ `for j = 1:jmax`	*begin iterations*
$\quad\quad \Delta\boldsymbol{\varepsilon} = \texttt{strain}(\Delta\boldsymbol{u})$	
$\quad\quad \boldsymbol{\sigma}_{k+1} = \texttt{returns}(\boldsymbol{\sigma}_k, \Delta\boldsymbol{\varepsilon})$	*new stress state*
$\quad\quad \boldsymbol{q} = \texttt{intforce}(\boldsymbol{\sigma}_{k+1})$	
$\quad\quad \boldsymbol{r} = (\boldsymbol{f}_k + \Delta\boldsymbol{f}_k) - \boldsymbol{q}$	
$\quad\quad$ `if` $\|\boldsymbol{r}\| < \epsilon \|\boldsymbol{f}_k + \Delta\boldsymbol{f}_k\|$	*check convergence*
$\quad\quad\quad$ `break`	*convergence*
$\quad\quad$ `else`	*continue iterations*
$\quad\quad\quad \delta\boldsymbol{u} = \boldsymbol{K}_t^{-1}\boldsymbol{r}$	
$\quad\quad\quad \Delta\boldsymbol{u} = \Delta\boldsymbol{u} + \delta\boldsymbol{u}$	
$\quad\quad$ `end`	
$\quad$ `end`	*end iterations*
$\quad \boldsymbol{u}_{k+1} = \boldsymbol{u}_k + \Delta\boldsymbol{u}$	*update*
$\quad \boldsymbol{f}_{k+1} = \boldsymbol{f}_k + \Delta\boldsymbol{f}_k$	
`end`	*end load step*

Table 3: Modified Newton-Raphson method.

References

[1] W. F. Chen and D. J. Han. *Plasticity for Structural Engineers*. Springer Verlag, New York, (1988).

[2] O. C. Zienkiewicz, S. Valliappan, and I. P. King. Elasto-plastic solutions of engineering problems; initial stress finite element approach. *International Journal for Numerical Methods on Engineering*, **1**, 75–100, (1969).

[3] O. C. Zienkiewicz and R. L. Taylor. *The Finite Element Method*. Butterworth–Heinemann, Oxford, (2000).

[4] R. D. Krieg and D. B. Krieg. Accuracies of numerical solution methods for the elastic-perfectly plastic model. *ASME Journal of Pressure Vessel Technology*, **99**, 510–515, (1977).

[5] S. G. Nash and A. Sofer. *Linear and Nonlinear Programming*. McGraw-Hill, New York, N.Y., (1996).

[6] S. Krenk. An orthogonal residual procedure for nonlinear finite element equations. *Computers & Structures*, **38**, 823–839, (1995).

[7] S. Krenk. *Non-Linear Analysis with Finite Elements*. Aalborg University, Aalborg, Denmark, (1993).

[8] S. Krenk. *Non-Linear Analysis with Finite Elements*. Technical University of Denmark, Lyngby, Denmark, (1997).

[9] M. A. Crisfield. *Non-linear Finite Analysis of Solids and Structures*. John Wiley & Sons, Inc., New York, (1997).