Practical Introduction to Non-linear Finite Element Analysis

Course Lecture Notes
by
Prof. Adib Becker

Contents

Practical Introduction to Non-linear Finite Element Analysis

Part 1

Introduction to Non-linear FE Analysis

Lecture Outline

1.1 Introduction

- In linear analysis, the behaviour of the structure is assumed to be **completely reversible**, i.e. the body returns to its original undeformed state upon the removal of the applied loads, and solutions for various load cases can be superimposed.

- **Examples of non-linear applications** include
  - Elastoplasticity of metals
  - Contact problems
  - Metal forming
  - Creep behaviour
  - Buckling
  – Impact probelms

- The application of the FE method to non-linear problems usually requires the use of **small load increments** and/or an **iterative procedure**.

- Solving non-linear problems using FE techniques requires **much more effort** than linear problems, with many additional issues that need to be addressed, such as convergence, automatic load incrementation, accuracy, etc.

1.2 Comparison of Linear and Non-linear Problems

Feature	Linear problems	Non-linear problems
Load-displacement relationship	Displacements are assumed to be linearly dependent on the applied loads	The load-displacement relationships are usually non-linear.
Stress-strain relationship	A linear relationship is assumed between stress and strain.	In problems involving material non-linearity, the stress-strain relationship is a non-linear function of stress, strain and/or time.

Feature	Linear problems	Non-linear problems
Magnitude of displacement	Changes in geometry due to displacement are assumed to be small and hence ignored, and the original (undeformed) state is always used as the reference state.	Displacements may not be small, hence an updated reference state may be needed.
Material properties	Linear elastic material properties are usually easy to obtain.	Non-linear material properties are difficult to obtain and may require additional experimental testing.
Units	Units of material properties and applied loads have to be consistent, and can be scaled.	Units of material properties and applied loads are very important, and cannot be scaled.

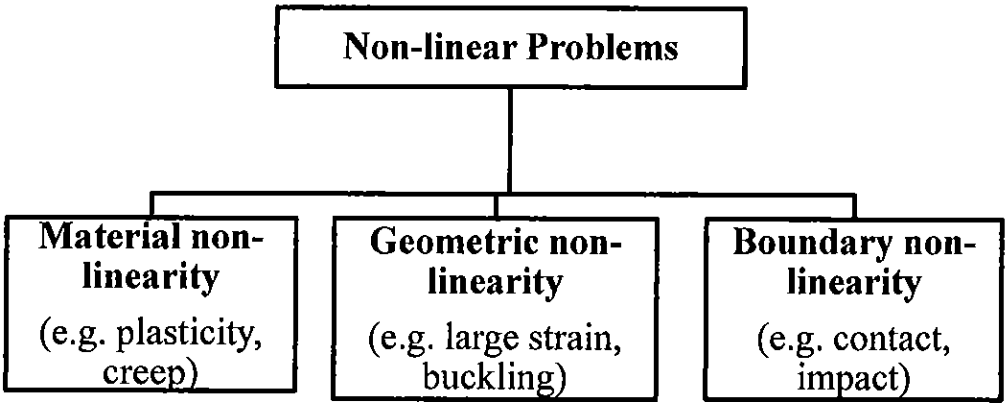
(Comparison of Linear and Non-linear Problems/ Continued)

Feature	Linear problems	Non-linear problems
Reversibility	The behaviour of the structure is completely reversible upon removal of the external loads.	Upon removal of the external loads, the final state may be different from the initial state.
Superposition	Solutions for various load cases can be linearly superimposed.	Solutions from several load cases cannot be superimposed.
Loading sequence	Loading sequence is not important, and the final state is unaffected by the load history.	The behaviour of the structure may depend on the load history, hence the load may have to be applied sequentially.

Feature	Linear problems	Non-linear problems
Iterations and increments	The load is applied in one go with no iterations.	The load is often divided into small increments with iterations performed to ensure that equilibrium is satisfied at every load increment.
Computation time	Computation time is relatively small in comparison to non-linear problems.	Due to the many solution steps required for load incrementation and iterations, computation time is high, particularly if a high degree of accuracy is sought.

Feature	Linear problems	Non-linear problems
Robustness of solutions	A solution can easily be obtained with no interaction from the user.	In difficult non-linear problems, the FE code may fail to converge without some interaction from the user.
Initial state of stress/strain	The initial state of stress and/or strain is unimportant.	The initial state of stress and/or strain is usually required for material non-linearity problems.

1.3 Classifications of Non-linear Problems



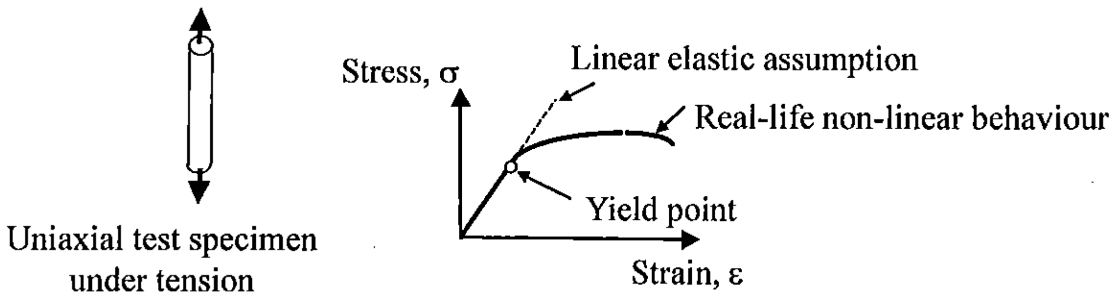
(a) Material non-linearity

Here the **stress-strain constitutive relationships are non-linear**. Such material non-linearities are usually further classified into three categories:

(i) **Time-independent behaviour** such the elasto-plastic behaviour in metals in which the material is loaded past the yield point.

(ii) **Time-dependent behaviour** such as creep of metals at high temperatures in which a power law stress-strain relationship is often used and the effect of variation of stress/strain with time is studied.

(iii) **Viscoelastic/viscoplastic** behaviour in which both the effects of plasticity and creep are exhibited. Here the stress is dependent on the strain rate.



(b) Geometric non-linearity

- Geometric non-linearity occurs when the **changes in the geometry** of a structure (due to its displacement under load) are taken into account in analysing its behaviour.

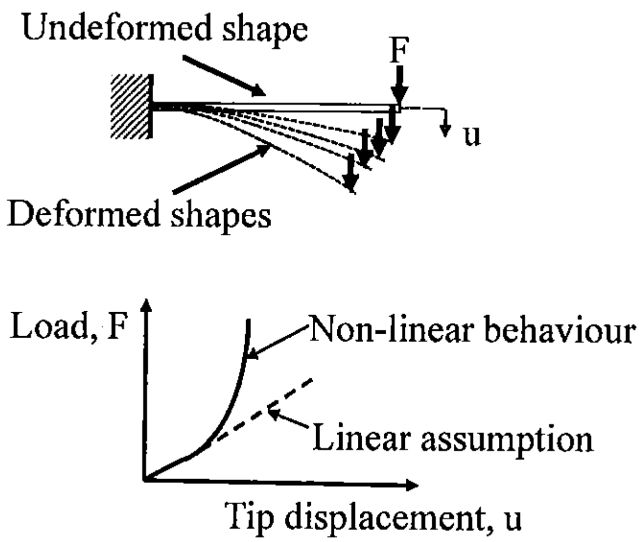
- Examples of geometric non-linearities

  - **Small strains and small rotations** : the deformation of shallow shells and arches subjected to lateral loading and elastic buckling of struts

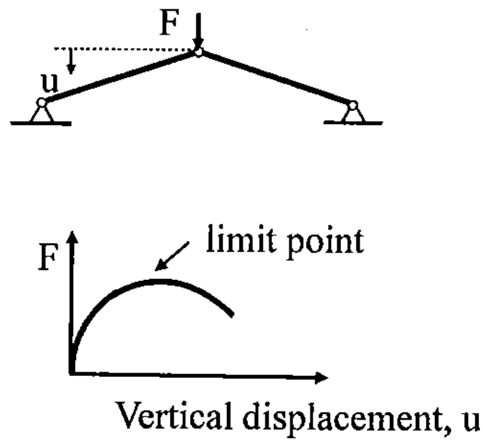
  - **Small strains with large rotations** : the displacement of a fishing rod under the weight of a heavy fish

  - **Large displacements, large rotations and large strains** (and also possible changes in the loading and boundary conditions as the structure deforms.): metal forming

Examples of Geometric non-linearity



GNL behaviour in a cantilever beam



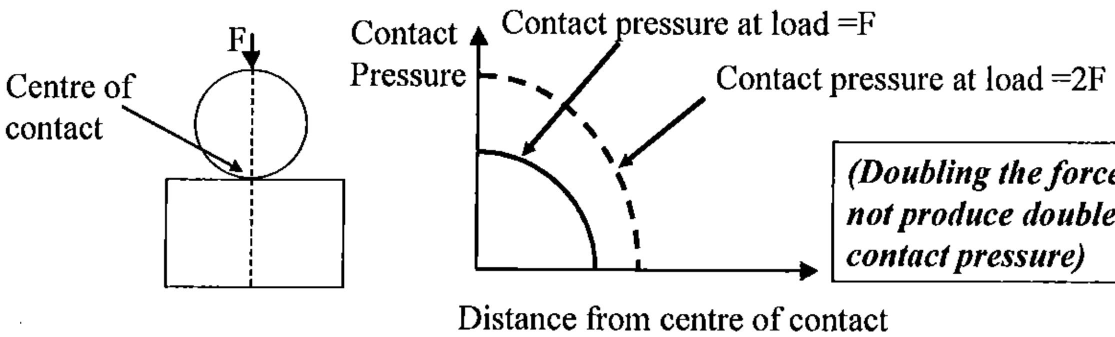
GNL snap-through problem

GNL small strains with large rotations

Examples of Geometric non-linearity

(c) Boundary non-linearity (Contact)

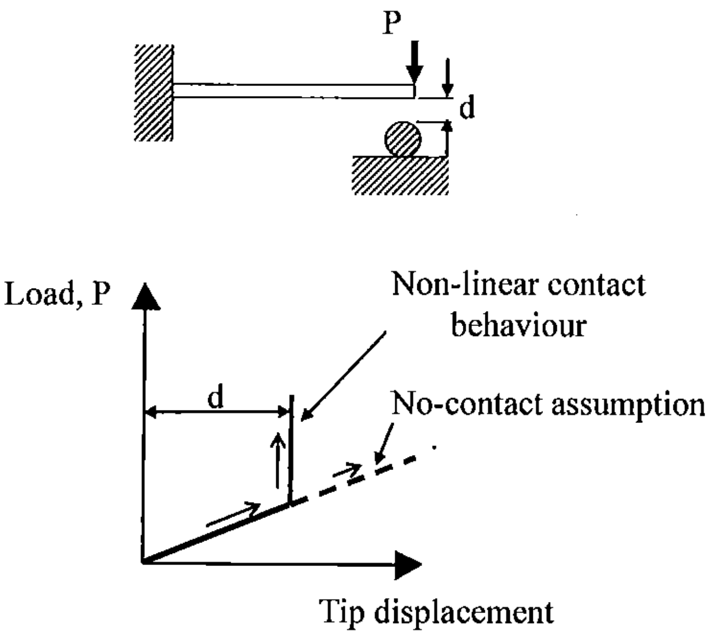
- Boundary non-linearity occurs in **most contact problems**, in which two surfaces come into or out of contact, and the displacements and stresses of the contacting bodies are not linearly dependent on the applied loads.
- If the **effect of friction** is included in the analysis, then a stick-slip behaviour may occur in the contact area which adds a further non-linear complexity that is normally dependent on the loading history



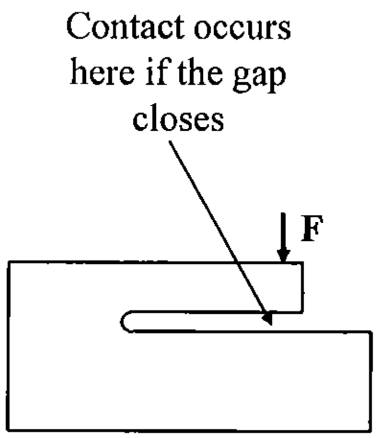
(Doubling the force does not produce double the contact pressure)

Example of contact of a roller on a flat plane

Examples of Boundary non-linearity (Contact)



Cantilever in contact with a rigid support



Self contact

1.4 Non-linear FE Procedures

The Newton-Raphson Iterative Method

- Several well established numerical techniques already exist for using iterations to solve non-linear equations of the following form:

$$f(x) = 0$$

where the function $f(x)$ is a non-linear function of x.

- In the **"Newton"** (also known as **"Newton-Raphson"**) method **a trial solution,** x_i, is first used which is reasonably close to the exact solution.

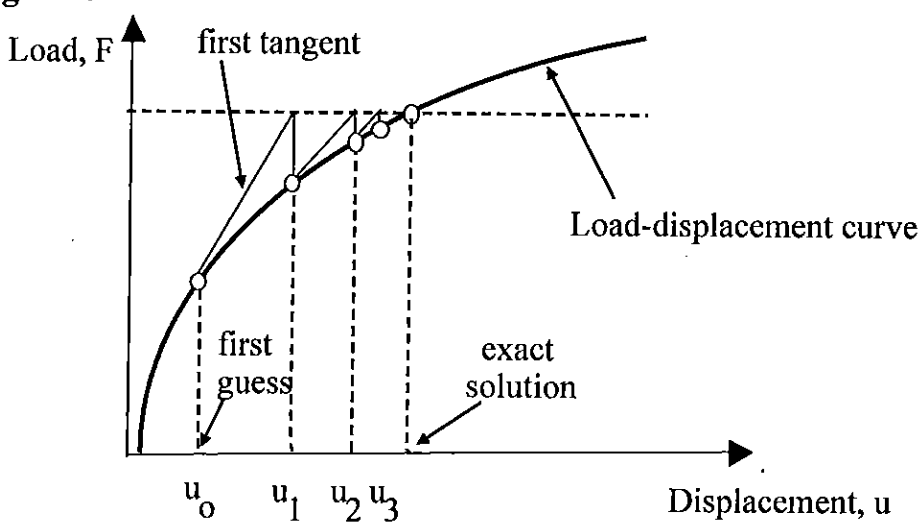
- The **next trial solution,** x_{i+1}, is then estimated using the slope (dy/dx) of the curve at the point x_i, as follows:

$$x_{i+1} = x_i - \frac{f(x_i)}{(dy/dx)_i}$$

Newton-Raphson Iterations

- **Iterations are performed** until the exact solution can be found to a specified degree of accuracy or tolerance, i.e. when the right hand side of the non-linear equation is very close to zero.

- The non-linear curve is effectively approximated by **a series of suitable tangents.**



Schematic representation of the Newton-Raphson method

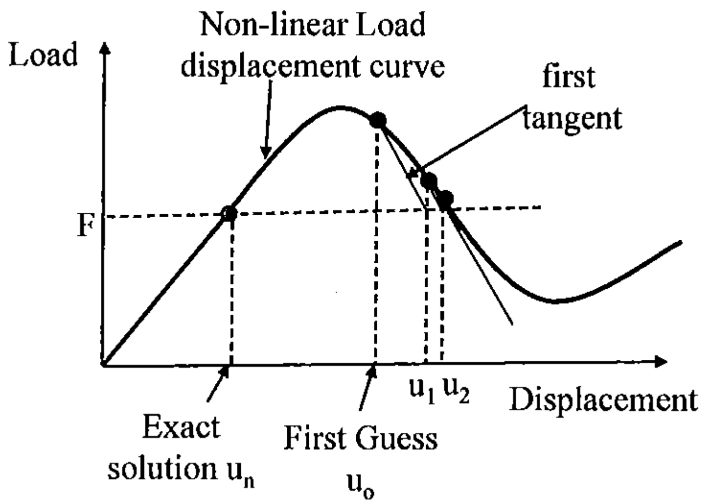
Limitations of the Newton-Raphson Method

For **convergence**, two conditions must be met:

(i) The initial guess is not very far from the exact solution, and

(ii) The slope of the non-linear load-displacement curve does not change its sign.

If the above conditions are not met, the solution may not converge.

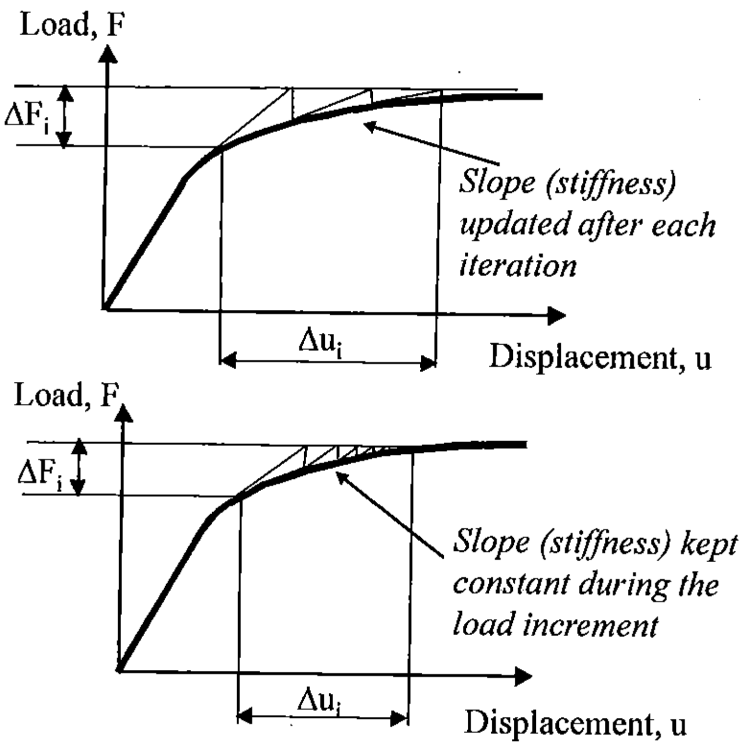


Example of non-convergence of the Newton-Raphson method

The Modified Newton-Raphson Method

- **Keeping the slope constant** for successive iterations is usually referred to as the **"Modified Newton-Raphson"** method.
- Computing the slope of the force-displacement curve (i.e. the stiffness) is an expensive, particularly in very large meshes.
- In large mesh problems, keeping the slope constant **may require more iterations** for convergence, but can be computationally less expensive than updating the slope.

Schematic representation of the modified Newton-Raphson method



(a) Standard Newton-Raphson method

(b) Modified Newton-Raphson method

Newton Method Example

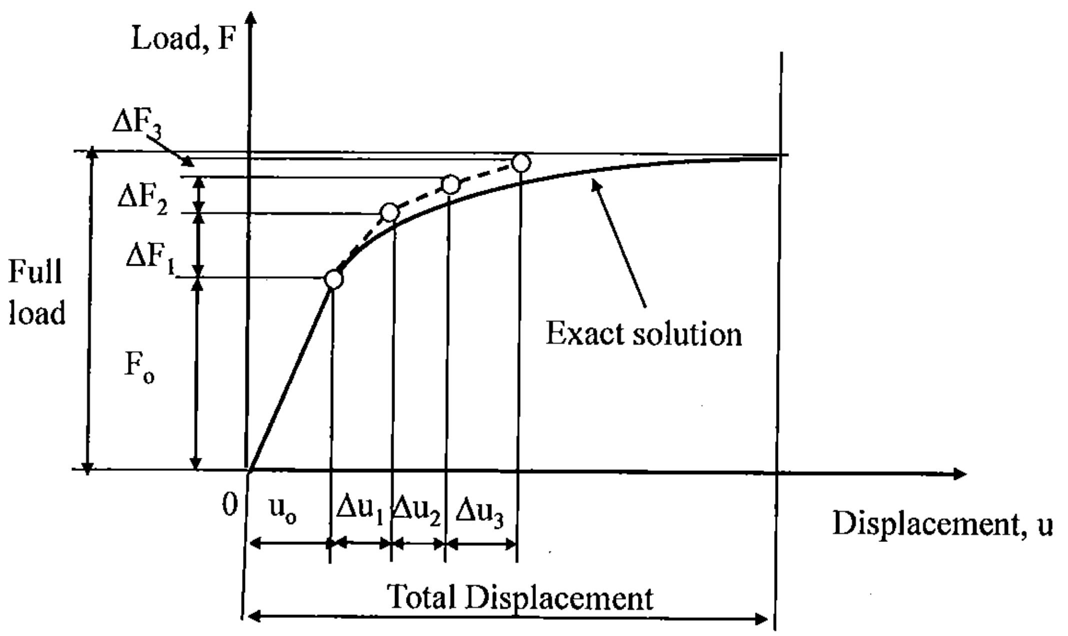
Consider the solution to the following non-linear equation:	$x^3 = 12\ ;\ \ i.e.\ f(x) = x^3 - 12$
The slope, i.e. the differential of f(x), is given by:	$\dfrac{d\,f(x)}{d\,x} = 3x^2$
Starting from an initial "guess" of x_o=1.0, the first approximation is obtained as:	$x_1 = 1.0 - \dfrac{(1.0)^3 - 12.0}{3\,(1.0)^2} = 4.6667$
The second approximation, x_2, is similarly obtained from x_1, as follows:	$x_2 = 4.6667 - \dfrac{(4.6667)^3 - 12.0}{3\,(4.6667)^2} = 3.2948$
If this procedure is continued, convergence (to 4 decimal points) will be achieved.	$x_3 = 2.5650$ $x_4 = 2.3179$ $x_5 = 2.2898$ $\mathbf{x_6 = 2.2894}$ (which is the exact solution).

- Note that a quicker convergence would be obtained if the initial guess of x, i.e. x_0, had been closer to the exact solution.

Load Incrementation Procedure

- The total applied load is divided into **small increments** and each increment is applied individually.
- The incremental procedure often **tends to drift away** from the equilibrium path.
- In order to satisfy equilibrium, **iterations should be performed** within each load increment to ensure that the equilibrium is satisfied, i.e. any out-of-balance forces must remain below a specified tolerance.
- The procedure **terminates when the final load is reached and equilibrium is satisfied**. The total deformation of the body is then calculated as the sum of the deformations associated with each load increment.
- The stiffness matrix calculated for each increment is often referred to as the **"tangent stiffness matrix"**, i.e. the tangent of the load displacement curve.

Schematic Representation of Load Incrementation



Iterative Procedure

- It is important to perform a check, within each load increment, to ensure that the solution is acceptable, i.e. **equilibrium is satisfied,** before proceeding to the next load increment.
- This is usually done by computing the **'residual' or 'out-of-balance' forces** within the structure, and reducing them to a negligible value.

Outline of the Iterative Procedure

Step (1) Using the strain-displacement and stress-strain relationships, **calculate the internal forces,** $[\Delta F_{internal}]$, resisting the load increment.

Step (2) Determine the **residual force vector,** $[\Delta F_{residual}]$, as the difference between the applied external load increment, $[\Delta F_{external}]$, and the resisting internal force, $[\Delta F_{internal}]$, i.e.

$$[\Delta F_{residual}] = [\Delta F_{external}] - [\Delta F_{internal}]$$

Step (3) Check **convergence.**

Step 3(a) Convergence achieved

Convergence is achieved if the nodal values of $[\Delta F_{residual}]$ are negligible, i.e. below a certain tolerance, defined either by the user or automatically by the FE code.

Typically, the **"norm"** of the vector is used to check convergence, which is defined as the sum of the square-roots of the nodal values, as follows:

$$Norm\left[\Delta F_{residual}\right] = \sqrt{\sum_{i=1}^{N}\left(\left[\Delta F_{residual}\right]_i\right)^2}$$

where subscript i represents the nodal value and N is the total number of nodes.

Step 3(b) No Convergence

- If convergence is not achieved, i.e. the ratio $[\Delta F_{residual}]$ to $[\Delta F_{external}]$ is not smaller than a specified tolerance, then a correction to the displacement vector is necessary.

- The residual force $[\Delta F_{residual}]$ can be used to obtain a correction to the displacement, as follows:

$$\left[\Delta u_{correction}\right] = \left[K\right]^{-1}\left[\Delta F_{residual}\right]$$

- Using this correction, obtain a new (improved) value of the displacement vector and **go back to step (3).**

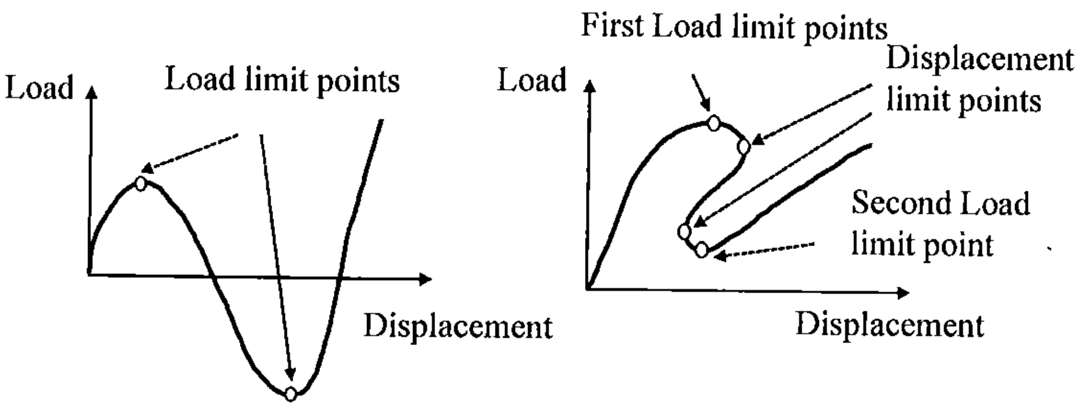
- **Repeat this correction procedure** until the displacement corrections or the residual force vector $[\Delta F_{residual}]$ become negligible, i.e. below a given tolerance.

Step (4) **Store all the values** of displacements and forces at all the nodal points and proceed to the next load increment.

Using Displacement Control and/or Load Control

- In the load incrementation procedure, the solution is driven by "**load control**", i.e. the structure is only allowed to deform by a single load increment at a time.
- This does not work when the load-displacement tangent becomes horizontal, because the load must decrease in order to satisfy equilibrium. In such problems, '**displacement control**' may be used where the displacement of a specified node is limited to a small displacement.
- In some problems, such as those associated with geometric non-linearity, **snap-back and snap-through problems**, a mixture of load control and displacement control may be the best approach.

Examples of Using Displacement and load controls



(Displacement control can be used here)

(a) Snap-through behaviour

(A mixture of displacement and load controls can be used here)

(b) Snap-back behaviour

Using Displacement control and/or load control

1.5 Difficulties in Modelling Non-linear Problems

Questions to consider when analysing Non-linear Analysis

Material Properties?	• How accurate are the material properties? • Would the solution be sensitive to the accuracy of the material data? • If so, is it is worth performing experimental tests to obtain a reliable set of material data?
Initial Conditions?	• Is there an initial state of stress/strain, e.g. locked-in residual stresses ? • If so, how can they be accounted for?
Uniqueness of Solution?	• Is there a possibility of more than one solution existing, e.g. an unstable path in buckling problems or more than one buckling mode? • If so, will it be necessary to introduce some kind of an artificial imperfection to arrive at only one stable solution?

(Questions to consider/Continued)

Load History?	• Will the structural behaviour be influenced by the load history? • Will the material behaviour be the same in loading and unloading?
Load Direction after Deformation?	• In large displacement or large rotation problems, will the direction of the applied loads remain in the same initial direction, or will they follow the deformed shape?
Non-linear Assumptions ?	• Are the assumptions used in the non-linear model adequate for the anticipated accuracy of the solution? • Should a larger factor of safety be used to compensate for any inaccuracies or uncertainties in the solutions?

Difficulties experienced by the FE user

Obtaining material properties	• Material properties are often difficult to obtain. It may even become necessary, to commission a series of experimental tests to obtain a reliable set of accurate material data. •The FE analyst must be able to interpret the data from the experimental tests and correctly feed them into the FE code.
Deciding the direction of the loads	• In non-linear problems where the displacements can be large, a decision has to be made by the user whether to keep the applied loads in the same initial direction (conservative loads), or to allow them to follow the deformed shape (follower loads),
Limited experience of the Analyst	• Most analyses of real-life non-linear problems heavily rely on experience and past knowledge. This makes it difficult for an inexperienced analyst to attempt to analyse non-linear problems for the first time.
Checking the accuracy of solutions	• In the absence of other alternative solutions, it is difficult to judge the accuracy of the final solutions.

Difficulties Encountered by the FE Code

Size of load increment	• It is difficult for the FE code to decide in advance, for all non-linear problems, what size of load increment and how many increments should be used. • FE codes often automate the load incrementation procedure, starting from a very small increment, e.g. 1% of the applied load, and then continually adjusting it as required.
Convergence Criteria	• The default automatic convergence criterion used in the FE code may not be suitable for all types of non-linear problems.
Changing sign of the load-displacement slope	• If the slope of the non-linear curve changes its sign during the loading, e.g. in a work softening situation, the standard incremental-iterative procedures may fail.
Non-unique Solutions	• In some problems, e.g.. buckling, the converged solutions may not be unique, i.e. there may be more than one solution that satisfies equilibrium.
Complex User Input	• In complex non-linear material behaviour, it may become necessary to specify the material law through "subroutines".

1.6 Using Non-linear FE Benchmarks

- The benchmark should be devised to **verify the reliability, robustness and accuracy** of the FE code.
- The non-linear problem must have a **reliable reference solution**; ideally a closed form analytical solution or alternatively a reliable numerical solution.
- Data input needed to define the benchmark should be kept to a minimum so that lengthy data generation is avoided.
- Ideally the non-linear benchmark should have **some educational merit**, in order to provide teaching material on particular aspects of non-linear behaviour through a case study.
- Whenever possible, the benchmark should reflect **real-life non-linear applications.**

1.7 Summary of Key Points

- The FE strategy in obtaining solutions to non-linear problems is to reduce the non-linear loading history of the structure to a **sequence of linear or weakly non-linear increments.**
- The applied load is usually applied in **small load increments** starting from an initial value and an **iterative procedure** is used to ensure that equilibrium is satisfied at each load increment.
- In non-linear applications, more care must be taken to use **accurate material properties** and a **consistent set of units** for all variables, since the solutions cannot be linearly scaled.
- There are **difficulties that may be encountered** in analysing and modelling non-linear problems using the FE technique. The user should use engineering judgement to check that the FE solutions are reliable.
- In practice, when faced with a non-linear problem for the first time, the user should run a **non-linear benchmark test.**

Part 2

Plasticity
(Time-independent Material Non-linearity)

Lecture Outline

2.1 Introduction

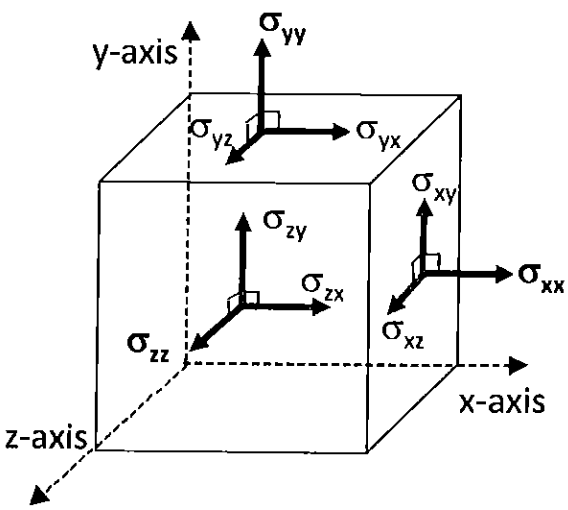
- Elasto-plastic non-linear material behaviour is discussed in which the effect of time is ignored.
- Introduction to elasto-plasticity theory is presented covering the fundamentals of elasto-plastic models.
- A general overview of the FE numerical procedures is presented, in order to appreciate the difficulties that may be encountered in the FE analysis of plasticity problems.

2.2 Stress-strain Relationships

- In a 3D Cartesian axes system, there are 6 components of stress:
  (a) Three **direct stresses (σ_{xx} , σ_{yy} , σ_{zz})**
  (b) Three **shear stresses (σ_{xy} , σ_{xz} , σ_{yz})**
- A stress matrix or a stress vector can be conveniently expressed as follows:

where superscript T indicates the transpose of the matrix or vector.

$$[\sigma]=\begin{bmatrix}\sigma_{xx} & \sigma_{yy} & \sigma_{zz} & \sigma_{xy} & \sigma_{xz} & \sigma_{yz}\end{bmatrix}^T$$



3D Cartesian stresses

Constitutive Equations (Hooke's Law)

- The 3D elastic stress-strain relationships (Hooke's law) can be written as follows:

$$\varepsilon_{xx} = \frac{1}{E}\left[\sigma_{xx} - \nu\left(\sigma_{yy} + \sigma_{zz}\right)\right]$$
$$\varepsilon_{yy} = \frac{1}{E}\left[\sigma_{yy} - \nu\left(\sigma_{xx} + \sigma_{zz}\right)\right]$$
$$\varepsilon_{zz} = \frac{1}{E}\left[\sigma_{zz} - \nu\left(\sigma_{xx} + \sigma_{yy}\right)\right]$$
$$\varepsilon_{xy} = \frac{1}{2\mu}\sigma_{xy}$$
$$\varepsilon_{xz} = \frac{1}{2\mu}\sigma_{xz}$$
$$\varepsilon_{yz} = \frac{1}{2\mu}\sigma_{yz}$$

  where E is Young's modulus, ν is Poisson's ratio and μ is the shear modulus defined as follows:

$$\mu = \frac{E}{2(1+\nu)}$$

- Hooke's law can be rearranged such that the stresses are on the left hand side, resulting in the following matrix expression, often referred to as the **'material constitutive equation'**:

$$[\sigma] = [D]\,[\varepsilon]$$

  where $[D]$ is called the elastic property matrix.

Definition of Strain

The 3D **direct (non-shear) strains** are related to the displacements in the following relationships:

$$\varepsilon_{xx} = \frac{\partial u_x}{\partial x}$$
$$\varepsilon_{yy} = \frac{\partial u_y}{\partial y}$$
$$\varepsilon_{zz} = \frac{\partial u_z}{\partial z}$$

The 3D **shear strains** are defined as follows:

$$\varepsilon_{xy} = \frac{1}{2}\left(\frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x}\right)$$
$$\varepsilon_{xz} = \frac{1}{2}\left(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x}\right)$$
$$\varepsilon_{yz} = \frac{1}{2}\left(\frac{\partial u_y}{\partial z} + \frac{\partial u_z}{\partial y}\right)$$

Mathematical and Engineering Shear Strains

- Two definitions of shear strain can be used:
  (a) **'Mathematical shear strain'** : where the 1/2 factor is used
  (b) **'Engineering shear strain'** : where the 1/2 factor is not used

- The 1/2 factor in the definition of the shear strains is mainly used for the convenience of use in tensor notation.

- If stress and strain vectors, rather than matrices, are used the 1/2 factor should not be used in the strain equations.

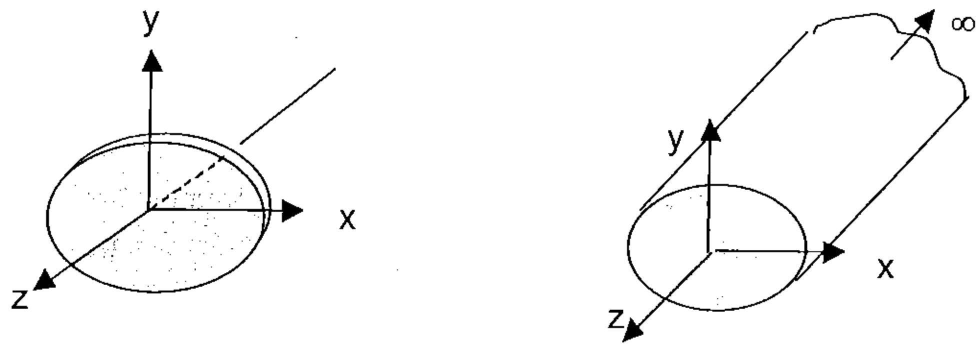
Two-dimensional Assumptions

There is no such problem as a truly 2D one; all 2D solutions are approximations of 3D solutions, and **the z direction is not ignored**.

(a) **<u>2D Plane stress</u>** is used to define 'thin' geometries in the z-direction where the stress across the thickness is neglected (i.e. $\sigma_{zz} = 0$).

(b) **<u>2D Plane strain</u>** is used to define very 'thick' geometries in the z-direction where the strain across the thickness is neglected, but the stress there is non-zero (i.e. $\varepsilon_{zz} = 0$ but $\sigma_{zz} \neq 0$). In other words, the x-y section is remote from the ends where $z = \pm \infty$.

(c) <u>2D 'generalised' plane strain</u>
A special case of plane strain arises when thermal strains are present, where the cross-section is prevented from distorting out of plane, but is allowed to have thermal strains in the z-direction. In such cases, the strain the z-direction, i.e. ε_{zz}, is constant.



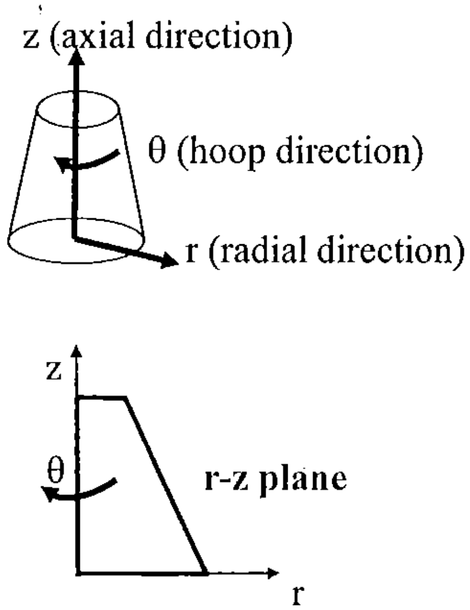
(a) 2D Plane stress
(σ_{zz}=0)

(b) 2D Plane strain
(ε_{zz}=0)

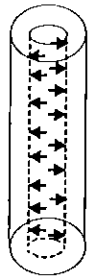
Two-dimensional Assumptions

Axisymmetric Problems

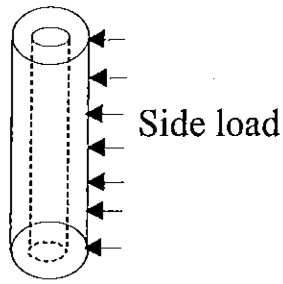
- Axisymmetric problems may also be viewed as two-dimensional, but the x and y directions are replaced by the radial (r) and axial (z) direction.
- Axisymmetric geometries, or bodies of revolution, are formed by rotating a two-dimensional flat plane through 360° about the z-axis.
- For an axisymmetrtic assumption to be valid, the geometry and all the variables must be axisymmetric. Therefore, all loads and boundary conditions must be applied across the whole circumference.



Axisymmetric Geometry



Axisymmetric Problem
(Pressurised cylinder)



NOT an Axisymmetric Problem
(Cylinder under side loading)

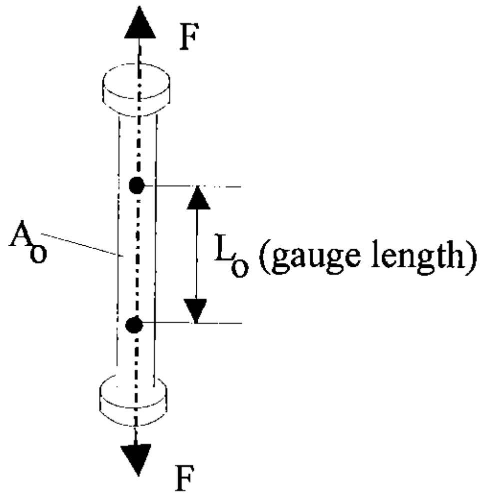
Axisymmetric Problems

Principal Stresses

- In any 3D stress state, three planes exist on which only the normal component of the stress remains, i.e. the shear component of the stress is zero.
- These planes are called the ***principal planes*** and the stresses acting on them are called the ***principal stresses*** (σ_1, σ_2 , σ_3).
- The usual convention is $\sigma_1 > \sigma_2 > \sigma_3$.
- The strains on the principal planes are referred to as the ***principal strains*** (ε_1, ε_2 , ε_3).
- The magnitudes of the principal stresses and strains are independent of the coordinate system used.

Uniaxial Testing

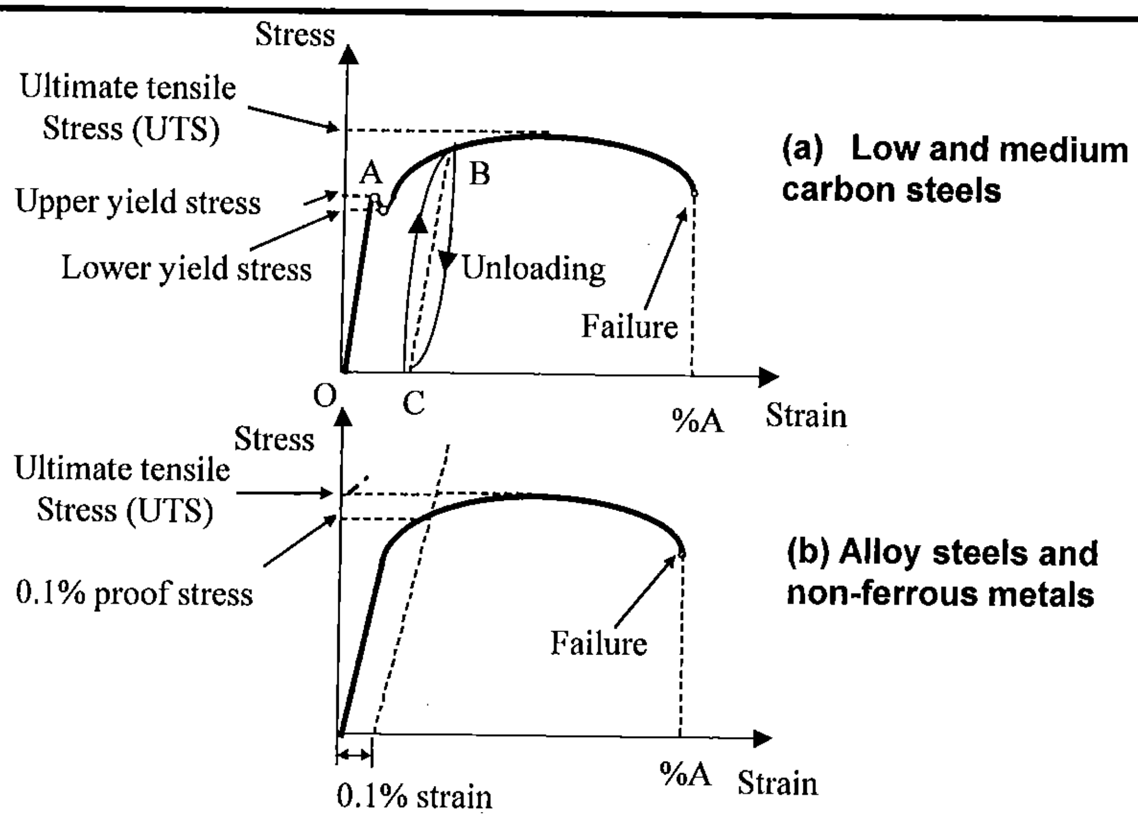
A test specimen in the shape of a cylindrical bar is subjected to a uniaxial tension F.



Uniaxial test specimen

Uniaxial Stress-Strain Curves

- Uniaxial tests are usually performed under **displacement control.**
- If the test is conducted under **load control**, i.e. the load is gradually increased, the softening part of the stress-strain curve will not be featured.
- The maximum stress reached in the test is called the **ultimate tensile stress (UTS).**
- The strain reached when the specimen fails (%A), is called the **percent elongation**, which is measure of the **ductility** of the material.
- Note: the load rate or the **strain rate** influences the stress-strain curve.



Typical uniaxial stress-strain curves

Nominal (Engineering) Stress and Strain

- For simplicity, the **'nominal' stress** is defined with respect to the constant original cross-sectional area, A_o as follows:

$$\sigma_O = \frac{F}{A_O}$$

- The **'nominal' or 'engineering' strain** is defined as the ratio of the change in length over a given gauge length to the original length L_o, as follows:

$$\varepsilon_O = \frac{L - L_O}{L_O} = \frac{L}{L_O} - 1$$

- More accurate definitions of stress and strain, called **'true stress'** and **'true strain'**, can be used in which the current cross-sectional area and current length are used.

True Stress and True Strain

- Assuming that the load remains constant on the test-piece, the stress will continue to rise, not fall, as **necking** develops before final fracture.
- In order to obtain a better understanding of plastic flow behaviour, **'true' stresses and true strains** are defined.
- To take account of the change in the dimensions of the test-piece, the ***true stress***, σ_{true} is defined as the force over the current (instantaneous) area, A, as follows:

$$\sigma_{true} = \frac{F}{A}$$

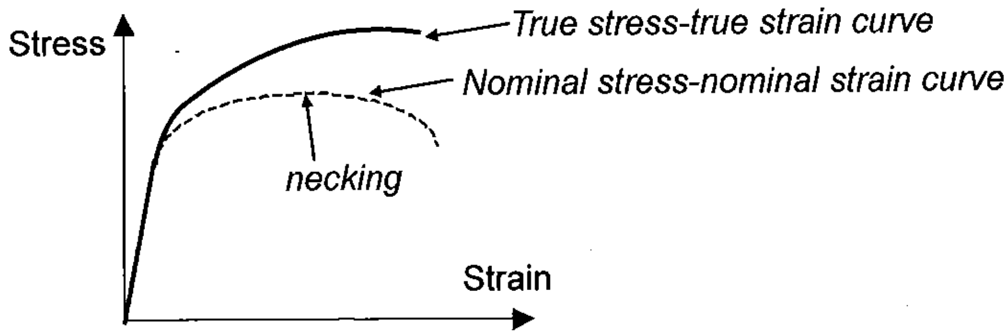
where L is the current (instantaneous) length.

- The ***true strain*** (also called natural or logarithmic strain), ε_{true}, is defined as follows:

$$\varepsilon_{true} = \int_{L}^{L_O} \frac{dL}{L}$$

True-stress-True strain curves

- In a true-stress vs. true strain plot, the curve continues rising beyond the point where necking appears. The curve clearly indicates a **'strain hardening'** effect, i.e. the material becomes harder as the strain is increased.
- It should be noted that the true stress-strain curve is strictly only valid up to the onset of necking, since the **formation of neck gives rise to a complex state of stress** which is no longer uniaxial, i.e. the stress is not simply the force divided by the cross-sectional area.



Nominal and true stress and strain curves

Yield Stress and Proof Stress

- **Yield Stress**
  In the mathematical theory of plasticity, it is assumed, for convenience, that the stress value which separates the stress-strain curve into an elastic portion and a plastic portion, is called the ***yield stress***, σ_{ys}. In some materials, σ_{ys} can be easily identified from the stress strain curves.

- **Low and medium carbon steels**
  Two yield points can be identified, an *upper yield stress* and a *lower yield stress*. However, often only a single value of the yield stress, usually the lower yield stress, is quoted.

- **Alloy steels and non-ferrous metals**
  The transition from linear to non-linear stress is a gradual process with no clearly identifiable yield stress. In such materials, a stress quantity called the ***proof stress*** at a given plastic strain is defined. Typically, 0.1% and 0.2% proof stresses are used for engineering problems.

Post Yield Behaviour

- Beyond the yield point deformations are permanent and unloading follows a path which is approximately parallel to the linear elastic part of the stress-strain curve.

- In real life, a slight rounding of the curve occurs with a small ***hysteresis loop*** caused by the loss in energy during the cycle of unloading and reloading. Such effects are usually ignored for the sake of simplicity.

- The total strain, ε_{total}, in the material at point B may be regarded as composed of two components; an **elastic component** $\varepsilon_{elastic}$ which is recoverable upon the removal of the load, and a **plastic strain,** $\varepsilon_{plastic}$, which is permanent and cannot be recovered if the load is removed, i.e.

$$\varepsilon_{total} = \varepsilon_{elastic} + \varepsilon_{plastic}$$

Derivation of True stress/strain

Using the **incompressibility (constant volume)** assumption in the plastic flow of metals, i.e. $AL = A_o L_o$, the following relationship between the nominal and current dimensions is obtained:

$$A = \frac{A_O L_O}{L} = \frac{A_O}{1+\varepsilon_O}$$

Using this value of A, the true stress can be expressed as follows:

$$\sigma_{true} = \frac{F\,(1+\varepsilon_O)}{A_O} = \sigma_O\,(1+\varepsilon_O)$$

Integrating the expression for the true strain in equation gives:

$$\varepsilon_{true} = \ln\left(\frac{L}{L_O}\right) = \ln\,(1+\varepsilon_O)$$

which is the total (logarithmic) strain between the original and current limits of length.Therefore, the **true plastic strain component** can be obtained by subtracting the elastic strain from the total strain, as follows:

$$\varepsilon_{plastic} = \varepsilon_{true} - \varepsilon_{elastic} = \ln\,(1+\varepsilon_O) - \frac{\sigma_O}{E}(1+\varepsilon_O)$$

2.3 Review of Elasto-plasticity Theory

In order to model elasto-plastic behaviour, a number of analytical relationships must be defined, as follows:

(i) **A yield criterion**
This defines how the multi-axial behaviour of the material is related to the uniaxial behaviour.

(ii) **A yield function (or yield surface)**
This defines when initial yielding occurs.

(iii) **A hardening model**
This defines what happens after initial yield and to accommodate the possible change in the yield surface.

(iv) **A plastic flow rule**
This defines the constitutive relationship between the plastic strain increment and the stress increment.

Elasto-plastic Material Models

- To simplify the mathematical modelling of plasticity, the material behaviour in the plastic range can be idealised into two types:

  (a) Perfectly plastic (also called non-work hardening), where there is no further increase in yield stress after initial yielding.

  (b) Strain hardening where the stress after initial yielding increases with continuing plastic strains. For materials obeying the von Mises criterion, strain hardening is equivalent to *work hardening*.

- In all elasto-plastic models, the total strain is divided into two parts; an elastic strain and a plastic strain as follows:

$$\varepsilon = \varepsilon^e + \varepsilon^p$$

where the superscripts e and p refer to the elastic and plastic components, respectively.

(Elasto-plastic Material Models/ Continued)

- A **hardening parameter or a plastic modulus, H**, is usually used to describe the relationship between the stress rate and the plastic strain rate, as follows:

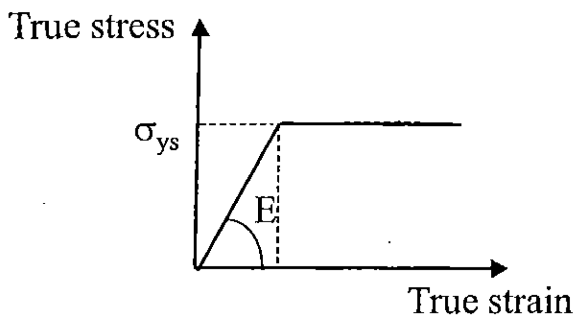
$$H = \frac{d\sigma}{d\,\varepsilon^p}$$

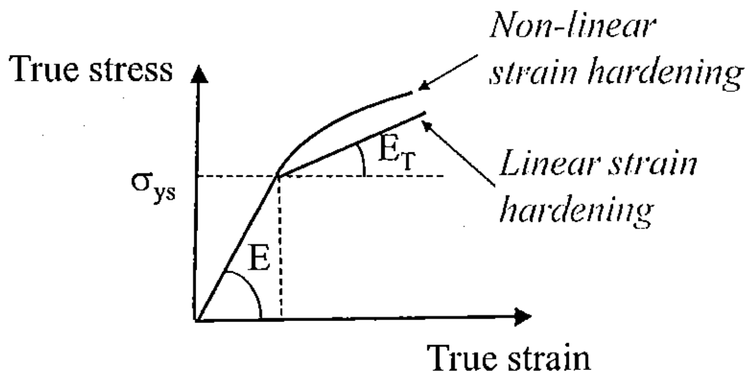
- It should be emphasised that the hardening parameter H is not the same as the tangent modulus, E_T, which is the slope of the stress-strain curve in the linear strain hardening model after the yield stress is exceeded, defined as follows:

$$E_T = \frac{d\sigma}{d\varepsilon^e + d\,\varepsilon^p}$$

- A relationship between H and E_T can be obtained, as follows:

$$\frac{1}{E_T} = \frac{1}{E} + \frac{1}{H}$$





(a) Elastic perfectly plastic model **(b) Elastic strain hardening model**

Simple elasto-plastic material models

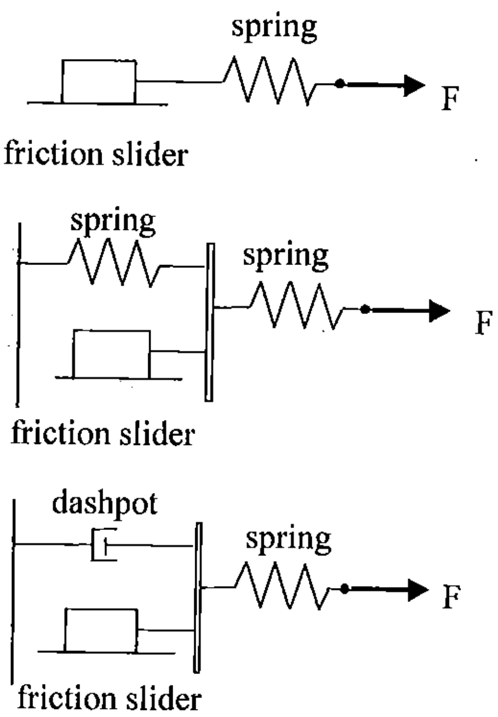
Rheological models

Rheological models use three basic components:

(i) **Elastic spring** (called the *Hookean* spring) in which the force is linearly proportional to the displacement.

(ii) **Viscous dashpot** or damper (called the *Newtonian* model) in which the force is linearly proportional to the velocity of the displacement, i.e. displacement rate. This model is usually used to represent time-dependent behaviour such as creep, visco-elasticity or visco-plasticity.

(iii) **Friction slider** (called the *St. Venant* model) in which no motion takes place until the force reaches a limiting value after which displacement takes place with no further increase in the force.



(a) Elastic perfectly plastic material

(b) Elastic linear strain hardening material

(c) Visco-plastic material

Rheological models used to describe plasticity

Deviatoric and Hydrostatic Stress

In the analysis of non-linear material behaviour, such as plasticity and creep, it is convenient to decompose the stress tensor into two distinct parts:

(a) A ***hydrostatic stress*** or mean stress, σ_m , which is the average of the direct stresses with no shear component, as follows:

$$\sigma_m = \frac{1}{3}\left(\sigma_{xx} + \sigma_{yy} + \sigma_{zz}\right) = \frac{1}{3}\left(\sigma_1 + \sigma_2 + \sigma_3\right)$$

(b) A ***deviatoric stress***, S_{ij} , which represents the shear component of the stress, i.e. the remainder of the stress after deducting the hydrostatic stress component, defined as follows:

$$S_{xx} = \sigma_{xx} - \sigma_m$$

$$S_{yy} = \sigma_{yy} - \sigma_m$$

$$S_{zz} = \sigma_{zz} - \sigma_m$$

$$S_{xy} = \sigma_{xy}; \quad S_{xz} = \sigma_{xz}; \quad S_{yz} = \sigma_{yz}$$

- The deviatoric stress is responsible for the plastic flow through the shear component, whereas the hydrostatic stress is composed of principal stresses with no shear component and thus is responsible only for the change in volume (but not shape) of an element of the material.

- Similarly, the strains can also be decomposed into **hydrostatic and deviatoric strains.**

Constant Volume and Strain Rate

- Experimental evidence for metals indicates that, for small strains, **plasticity (and creep) are constant volume processes**. Therefore, the principal strain rates can be combined in one expression as follows:

$$\dot{\varepsilon}_1^p + \dot{\varepsilon}_2^p + \dot{\varepsilon}_3^p = 0$$

- The dot displayed on top of the strain is used to indicate **strain rate** or an increment of strain. For a small strain increment, the following expression can be written:

$$[\dot{\varepsilon}] = \frac{d[\varepsilon]}{dt} = \frac{[\Delta\varepsilon]}{\Delta t}$$

where t is the time.

- Since plasticity is assumed to be independent of time, t may be considered a **'pseudo-time'** which refers to the **sequence of the load increments**, i.e. the strain rate is defined as the ratio of the strain increment to the load increment.

Von Mises Yield Criterion

- One widely used criterion of yielding is the ***Von Mises yield criterion*** which relates the complex 3D stresses in plasticity to the uniaxial behaviour by using the concept of the critical value of the shear strain energy stored in the material.

- The stress measure associated with the von Mises yield criterion is the ***Von Mises equivalent or effective stress, σ_{eff}***. This is a scalar measure of the three-dimensional stress state and can be expressed in terms of the principal stresses as follows:

$$\sigma_{eff}=\frac{1}{\sqrt{2}}\left[(\sigma_1-\sigma_2)^2+(\sigma_2-\sigma_3)^2+(\sigma_1-\sigma_3)^2\right]^{1/2}$$

Alternatively, the effective stress can be written in terms of the Cartesian stresses as follows:

$$\sigma_{eff}=\frac{1}{\sqrt{2}}\left[(\sigma_{xx}-\sigma_{yy})^2+(\sigma_{yy}-\sigma_{zz})^2+(\sigma_{xx}-\sigma_{zz})^2+6\left(\sigma_{xy}^2+\sigma_{yz}^2+\sigma_{xz}^2\right)\right]^{1/2}$$

(Von Mises Yield Criterion/ Continued)

- The Von Mises effective stress can also be expressed **in terms of the deviatoric stresses**, in tensor notation, as follows.

$$\sigma_{eff}=\left[\frac{3}{2}S_{ij}\,S_{ij}\right]^{1/2}$$

which clearly demonstrates that the effective stress is only a function of the deviatoric stress and is unaffected by the hydrostatic stress.

- Similarly, an expression for the **equivalent Von Mises or effective plastic strain rate** can be written in terms of the Cartesian plastic strain rates as follows:

$$\dot{\varepsilon}^p_{eff}=\frac{\sqrt{2}}{3}\left[\left(\dot{\varepsilon}^p_{xx}-\dot{\varepsilon}^p_{yy}\right)^2+\left(\dot{\varepsilon}^p_{yy}-\dot{\varepsilon}^p_{zz}\right)^2+\left(\dot{\varepsilon}^p_{xx}-\dot{\varepsilon}^p_{zz}\right)^2+6\left(\dot{\varepsilon}^p_{xy}\right)^2+6\left(\dot{\varepsilon}^p_{yz}\right)^2+6\left(\dot{\varepsilon}^p_{xz}\right)^2\right]^{1/2}$$

Yield Function

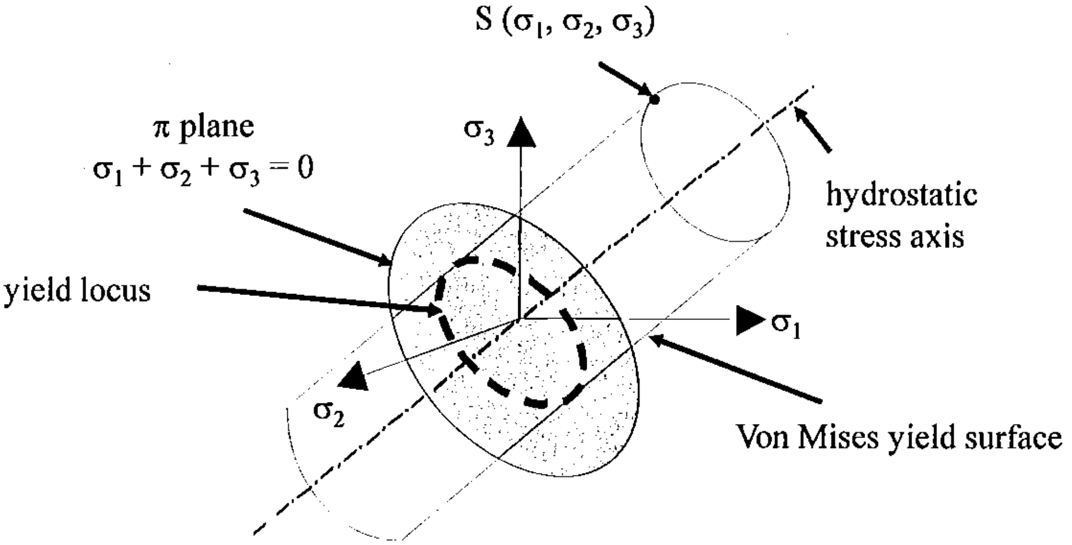
- The von Mises criterion states that yielding in an elasto-plastic material occurs when the effective stress σ_{eff} reaches the yield stress σ_{ys}, and is governed by a scalar function termed the ***yield function, F,*** which is a function of the stress invariants for an isotropic material.

- The yield function is written such that elastic behaviour occurs if $F <$ 0, while plastic yielding occurs if F = 0, and can be expressed as follows:
$$F = \sigma_{eff} - \sigma_{ys} = 0$$

- This yield criterion has been verified by a series of experiments mostly on thin metal tubes under biaxial stress states. It is important to note that, under this criterion, yielding is assumed to be unaffected by the hydrostatic stress.

Geometric representation of the Von Mises yield criterion

In geometric terms, s_m can be represented by a line subtending equal angles with the coordinate axes of s_1, s_2 and s_3, and the deviatoric stress is a vector lying in a plane perpendicular to this line (often called **the *p- plane***).



Geometric representation of Von Mises yield criterion

- In principal stress (σ_1, σ_2, σ_3) space, a **yield surface (F)** is defined by the following condition:

$$F(\sigma_1, \sigma_2, \sigma_3) = 0$$

- The Von Mises yield criterion can be **graphically represented** by:
  - An infinitely long cylinder with its axis subtending equal angles with the coordinate axes of σ_1, σ_2 and σ_3.
  - The axis of the cylinder represents the hydrostatic stress, s_m.
  - Stress points which lie inside the cylindrical yield surface represent elastic stress states.
  - Stress points that lie on the surface represent plastic stress states.
  - **Points are not allowed to lie outside the yield surface.**

The π-plane

- It is possible to represent the yield surface geometrically by using a plane in the (σ_1, σ_2, σ_3) space which passes through the origin and subtends equal angles with the coordinate axes, defined as follows:

$$\sigma_1 + \sigma_2 + \sigma_3 = 0$$

- This plane is called the ***π-plane***. The intersection of the π-plane with the von Mises yield surface appears as a **circle** called the ***yield locus.***

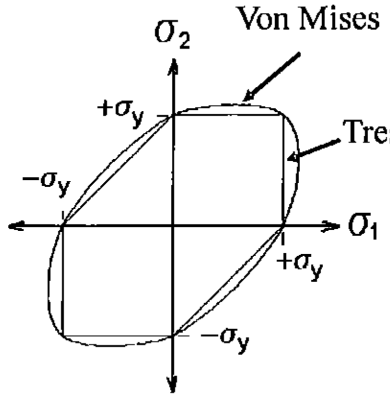
- To obtain a visual impression of a stress state it is possible to project a stress point S(σ_1, σ_2, σ_3) onto the π-plane

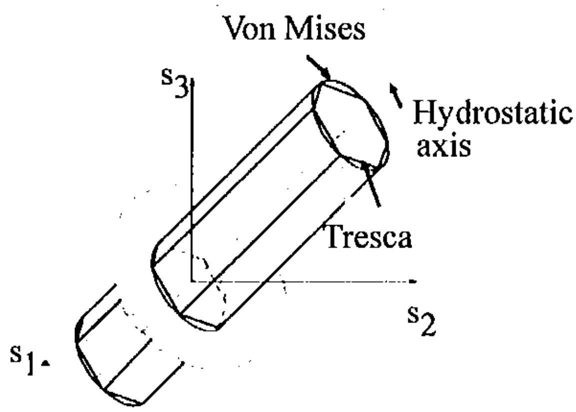
Other Yield Criteria

Tresca Yield Criterion

Tresca yield criterion forms a regular hexagon in the p plane

$$\left(\sigma_{eff}\right)_{Tresca} = Max\ of\ \left|\frac{\sigma_1 - \sigma_2}{2}\right|;\left|\frac{\sigma_1 - \sigma_3}{2}\right|\ or\ \left|\frac{\sigma_2 - \sigma_3}{2}\right|$$





Geometric representation of Tresca yield criterion

Other yield criteria exist. These include:

- **Drucker-Prager yield criterion** which is a generalisation of the von Mises criterion that includes the dependency on the hydrostatic stress
- **Mohr-Coulomb yield criterion** used for rock and concrete, which is dependent on the hydrostatic stress
- **Hill yield criterion** for orthotropic materials.

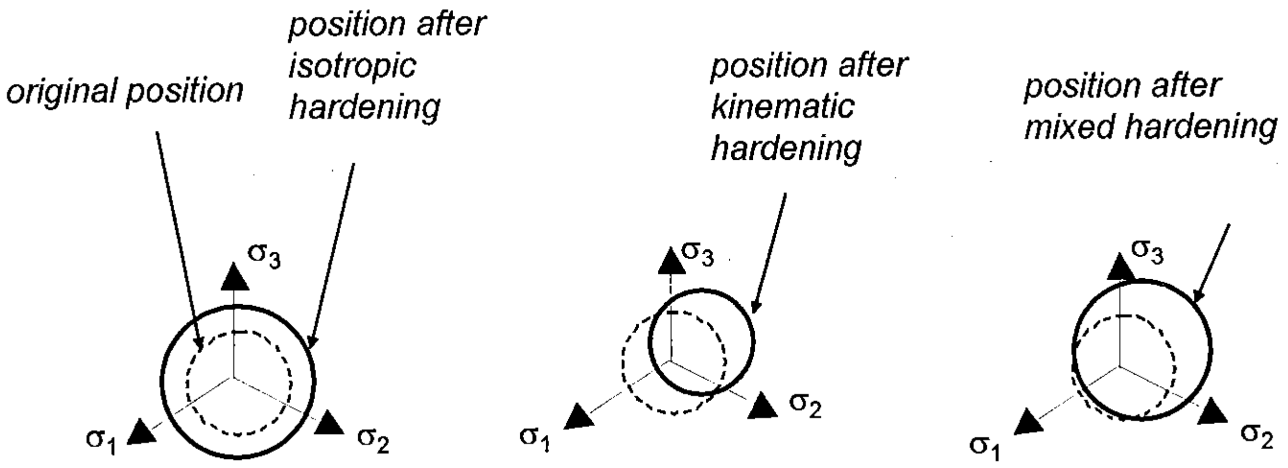
Post-Yield Behaviour

- **Perfectly plastic material,**
  There is no change in the yield surface during plastic deformation. This means that the yield stress σ_{ys} does not increase with increasing plastic strain.

- **Isotropic hardening material**
  The yield surface increases in size with increasing plastic strain but maintains its original shape. On the π-plane, the von Mises yield loci appear as a set of concentric circles.

- **Kinematic hardening material**
  The original yield surface is translated to a new position in the stress space as the plastic strain is increased, with no change in size or shape.

- **Mixed hardening material**
  The original yield surface both expands and translates to a new position with increasing plastic strain.



(a) Isotropic hardening **(b) Kinematic hardening** **(c) Mixed hardening**

View of von Mises yield criterion on the π-plane with hardening

Plastic Flow Behaviour

- When the stress at any point reaches the yield surface, the material starts to undergo plastic deformation. The plastic strain is defined by a ***plastic flow rule*** which can be written for a given direction such as the x-direction as follows:

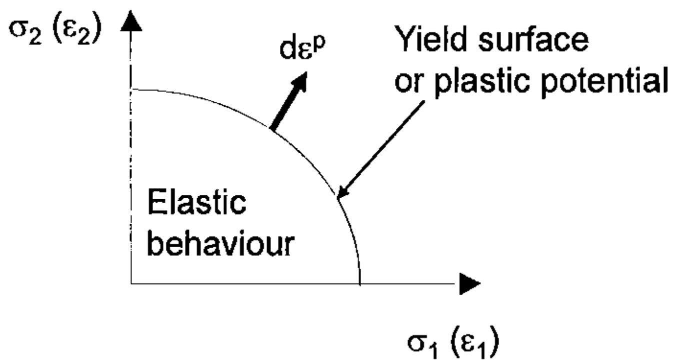
$$d\varepsilon_{xx}^{p} = d\lambda \frac{\partial Q}{\partial \sigma_{xx}}$$

  where Q is called the *plastic potential* function, and $d\lambda$ is a positive scalar factor of proportionality. All other stress and strain components can be expressed similarly.

- This rule is called the ***normality rule*** as it implies that the plastic strain components are in a ratio such that their resultant is in a direction normal to the plastic potential surface.

- If the yield function F and the plastic potential Q are identical then the flow rule is called **an *associated flow rule***, which is valid for most metals.

(Plastic Flow Behaviour/ Continued)



Yield surface and normality rule for a 2D stress space

2.4 Plasticity under Cyclic Loading

- **The Bauschinger effect**
  $(\sigma_{ys})_{\text{compression}}$ is less than $(\sigma_{ys})_{\text{tension}}$
- In **isotropic hardening** it is assumed that during **cyclic loading** in which the load changes from tensile to compressive, the yield point and the effects of work hardening are the same in tension and compression (i.e. no Bauschinger effect),
- In **kinematic hardening** the yield point in compression is lower.
- **Cyclic hardening or softening** is observed in some metals.
- It is often convenient to represent this type of cyclic hardening or softening by a **'backbone' or 'cyclic' stress-strain curve**, which is drawn through the tips of the stable hysteresis loops.

Shakedown and Ratchetting

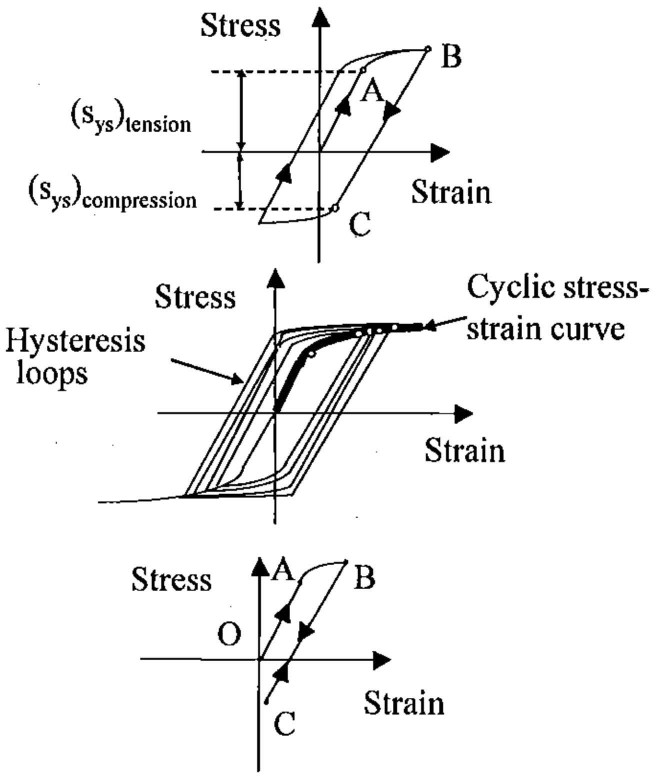
Three important phenomena can be observed in cyclic loading when the load amplitude is kept constant.

(i) <u>**Elastic shakedown**</u> : This occurs when the plastic strain in the cycle is relatively small. The material is 'shaken down' to a stabilising condition, i.e. the structure is assumed to have 'settled down' to an elastic state.

(ii) <u>**Ratchetting**</u> : Depending on the load level, in some loading situations where a constant amplitude of stress is imposed on the material, a stable hysteresis loop may not be reached. Instead, plastic strains keep on accumulating incrementally with each cycle, leading to eventual failure. This mechanism is called 'ratchetting', also known as 'cyclic creep', and can occur due to a cyclic thermal loading under a constant mechanical load.

(iii) <u>**Alternating plasticity**</u>: Here the behaviour settles down to a state where the plastic strains in each cycle are equal and opposite, and there is a progressive increase in total strain.

Plasticity under cyclic and reversed loading



(a) Hysteresis loop

(b) Cyclic stress-strain curve

(c) Elastic Shakedown

Plasticity under cyclic and reversed loading

2.5 Finite Element Treatment of Plasticity

Assumptions Used in the Analysis of Elasto-Plasticity

Before describing the FE treatment of elasto-plasticity, it is worth summarising the assumptions made for modelling elasto-plastic flow, as follows:

(i) **Elasto-plastic flow is path-dependent,** i.e. dependent on the sequence of loading, but is independent of time.

(ii) **All variables can be expressed as 'rates',** simply to indicate the sequence of events, rather than the change with time. Time is referred to as 'pseudo-time'.

(iii) **The external loads can be applied in small increments**, and the deformations and stresses resulting from these increments can be accumulated to give the final state of stresses and strains.

Load and Strain Increments in Plasticity

- The total strain **rate** or increment can be divided into two parts; an elastic recoverable part and a plastic irrecoverable part as follows:

$$\left\lfloor \dot{\varepsilon}^{total} \right\rfloor = \left\lfloor \dot{\varepsilon}^{e} \right\rfloor + \left\lfloor \dot{\varepsilon}^{p} \right\rfloor$$

- Note that the stress rate or increment is not divided into elastic and plastic parts.

- The rate forms of the equations of plastic flow can be considered as "incremental" stress-strain relationships.

- To follow the plastic flow path, these **increments must be integrated** in order to determine the final plastic deformation of the body. In practical problems, these integrations are too complex to be performed analytically, and a numerical scheme must be devised.

Pseudo-time in FE Plasticity Analysis

- The loading path is divided into small **pseudo-time increments** over which the non-linear elasto-plastic constitutive relationships can be assumed to be linear.

- Since plasticity is assumed to be independent of time, the pseudo-time increment is regarded as a **load increment.**

- The final state of stress and strain in the body can be obtained by the **accumulation of all increments** over all the pseudo-time steps.

Outline of FE Plasticity Algorithm

- In elasto-plastic applications, **both load incrementation and iterations** are necessary to arrive at the final solution.

- Within **each load increment**, the plasticity relationships which depend on the current state of loading are satisfied and the variables updated ready for the next load increment.

- **Iterations are required within each load increment** in order to ensure that the dual requirements of satisfying the constitutive plasticity equations and the overall structural equilibrium are satisfied.

- The iterations are terminated when the solution has converged according to a suitable convergence criterion. The number of iterations depends on the size of the load increment, and whether the stiffness matrix has been updated to reflect the current stress-strain state.

Typical FE Algorithms For Modelling Elasto-plasticity

(i) Apply the load **assuming elastic behaviour** everywhere. The overall stiffness matrix is obtained, and the following equations solved to obtain the displacement vector:

$$[K_e][u]=[F]$$

where $[K_e]$ is the elastic stiffness matrix, and $[F]$ is the full load vector.

(ii) **Calculate the effective stress** at all Gauss (integration) points, and check the maximum value against the yield stress, σ_{ys}.

- If it does not exceed σ_{ys}, then there is no need for a plasticity analysis.

- If it does, scale down the magnitudes of all nodal displacements such that the node or Gauss point with the highest effective stress is just yielding. The **scaling factor** is used to determine the fraction of the applied loads that causes initial yielding.

(iii) Divide the remainder of the applied load into **small increments**, either as specified by the user, or according to a suitable automatic scheme.

(iv) **Apply one load increment**, and re-solve the equations to obtain the new displacements corresponding to this load vector.

Use either the initial (elastic) stiffness matrix $[K_e]$, or the tangent stiffness matrix, $[K_{ep}]$, which is updated to contain the current state of plasticity, as follows:

$$\lfloor K_{ep} \rfloor [\Delta u] = [\Delta F]$$

Many FE codes use the initial stiffness matrix and then update it after a few load increments or iterations to reflect the current state of plasticity.

(v) **Perform iterations** to ensure that the solution is acceptable, i.e. it satisfies both equilibrium conditions and the plasticity material laws.

From the computed $[\Delta u]$, the total and plastic strain increments can be calculated, and then the corresponding stress.

An **out-of-balance or residual force vector, [*R*],** is calculated by integrating the stresses over all the elements and subtracting the internal forces from the external forces.

If convergence is not achieved, i.e. $[R]$ is not smaller that a specified tolerance, an iteration is commenced by solving a new set of equations, as follows:

$$\lfloor K_{ep} \rfloor [\Delta u_{correction}] = [R]$$

where $[\Delta u_{correction}]$ is a correction to the displacement vector in order to balance the residual force vector.

(vi) Using this correction, **the displacement vector is improved** and the iteration repeated to obtain a new residual force vector. The displacement corrections and the residual forces should be getting smaller with each iteration until convergence is achieved.

(vii) Store the computed increments of displacements, strains and stresses at each node, and **update the existing values.**

(viii) **Apply the next load increment** and perform iterations as necessary. Terminate the calculations when the final load is reached.

Notes on the FE Plasticity Algorithms

- **Accelerated Increments:**
  The above simplified procedure can be refined and accelerated by modifying the size of the load increment according to the amount of plastic strain generated in the increment.
- **Automated Procedures:**
  It is possible to automate the load incrementation procedure and the convergence criterion in most applications, for example by ensuring that the norm of the residual force vector is within 1% of the applied load.

- **Visco-plastic Behaviour:**
  In more complex analyses of material non-linearities, for example when plasticity is combined with creep, it is important to monitor the solution with respect to 'equivalent time' in plasticity and 'actual time' in creep. More robust algorithms are necessary to increment the time as well as the load in such applications.

2.6 Plasticity Examples

Example 1: Fundamental 2D Plasticity

Example 2: Kinematic Hardening Plasticity

Example 3: Multi-Axial Pressurised Cylinder Plasticity

Example 4: Cyclic Plasticity

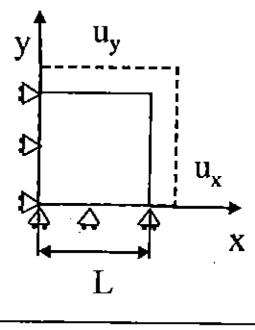
Example 5: Rigid Punch Plasticity

2.6.1 Isotropic Hardening Example

Physical Attributes

A simple 2D geometry subjected to prescribed displacements in order to demonstrate the following:

- Biaxial yielding
- Perfect plasticity
- Isotropic hardening
- Plastic flow
- Unloading and subsequent reloading
- Residual stresses after unloading
- Prescribed displacements at all nodes (displacement-control)



Problem Definition

A 2D square plate is stretched in the x- and y-directions and then returned to its original shape.
The prescribed displacements are applied in 8 steps, as follows

Step 1: Stretching in the x-direction until the plate just yields

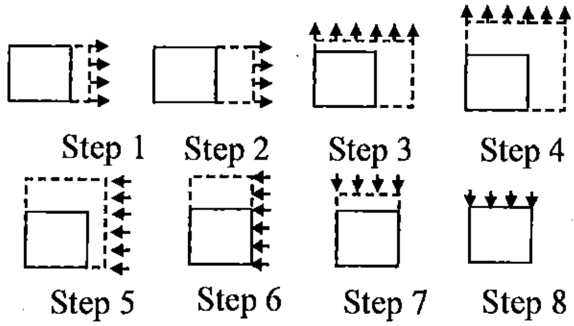
Step 2: Further stretching in the x-direction causing plastic flow, i.e. post-yield behaviour.

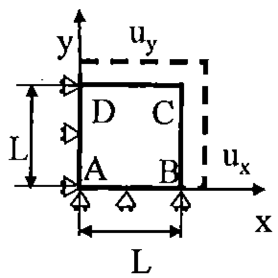
Steps 3 & 4: Stretching in the y-direction in two steps.

Steps 5 &6: Compression in the x-direction in two steps.

Steps 7 &8: Compression in the y-direction in two steps.

At the end of the final load step, the plate is returned to its original dimensions.



<table>
<tr><td>Geometry</td><td>2D Plane strain
L = 1.0 mm</td></tr>
<tr><td>Material Properties</td><td>E = 250.0 x 10^3 N/mm^2, v = 0.25
Plasticity model :
(a) Perfect plasticity, σ_{ys} = 5.0 N/mm^2
(b) Isotropic hardening, E_T = 50.0 x 10^3 N/mm^2</td></tr>
<tr><td>Boundary Conditions</td><td>$u_x = 0$ on line AD, $u_y = 0$ on line AB
$u_x = u_x$ on line BC, $u_y = d_y$ on line CD
u_x and u_y are applied in 8 steps as follows: (R=2.5 x 10^{-5})
<table>
<tr><th>Step</th><th>Disp. Change</th><th>u_x</th><th>u_y</th></tr>
<tr><td>Step 1</td><td>$\Delta u_x = R$</td><td>R</td><td>0.0</td></tr>
<tr><td>Step 2</td><td>$\Delta u_x = R$</td><td>2R</td><td>0.0</td></tr>
<tr><td>Step 3</td><td>$\Delta u_y = R$</td><td>2R</td><td>R</td></tr>
<tr><td>Step 4</td><td>$\Delta u_y = R$</td><td>2R</td><td>2R</td></tr>
<tr><td>Step 5</td><td>$\Delta u_x = -R$</td><td>R</td><td>2R</td></tr>
<tr><td>Step 6</td><td>$\Delta u_x = -R$</td><td>0.0</td><td>2R</td></tr>
<tr><td>Step 7</td><td>$\Delta u_y = -R$</td><td>0.0</td><td>R</td></tr>
<tr><td>Step 8</td><td>$\Delta u_y = -R$</td><td>0.0</td><td>0.0</td></tr>
</table>

</td></tr>
<tr><td>Loading</td><td>No applied forces</td></tr>
</table>

2D Projection of the Von Mises Yield Surface

In order to follow the stress paths graphically on the yield surface, a special projection can be used. The yield function can be expressed as a function of the principal stresses, σ_1, σ_2, σ_3 as follows:

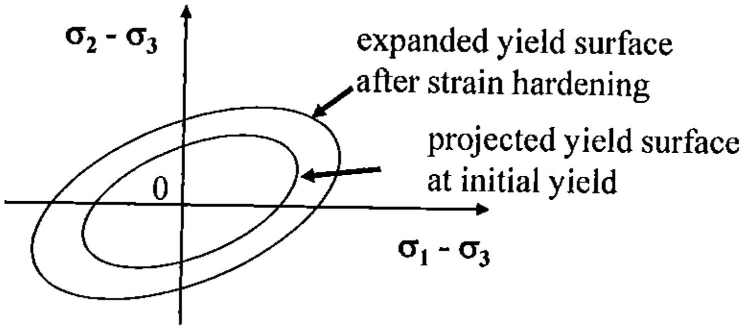
$$F = 0 = \frac{1}{2}\left[(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_1 - \sigma_3)^2\right] - (\sigma_{ys})^2$$

Using the following substitution:

$$x = \sigma_1 - \sigma_3 \;\; ; \;\; y = \sigma_2 - \sigma_3$$

the yield function can be expressed by the following ellipse:

$$\frac{x^2}{\sigma_{ys}^2} - \frac{xy}{\sigma_{ys}^2} + \frac{y^2}{\sigma_{ys}^2} = 1$$



Notes on the FE Model

- **Elements**

This problem can be modelled by a single or any number of 2D plane strain elements.

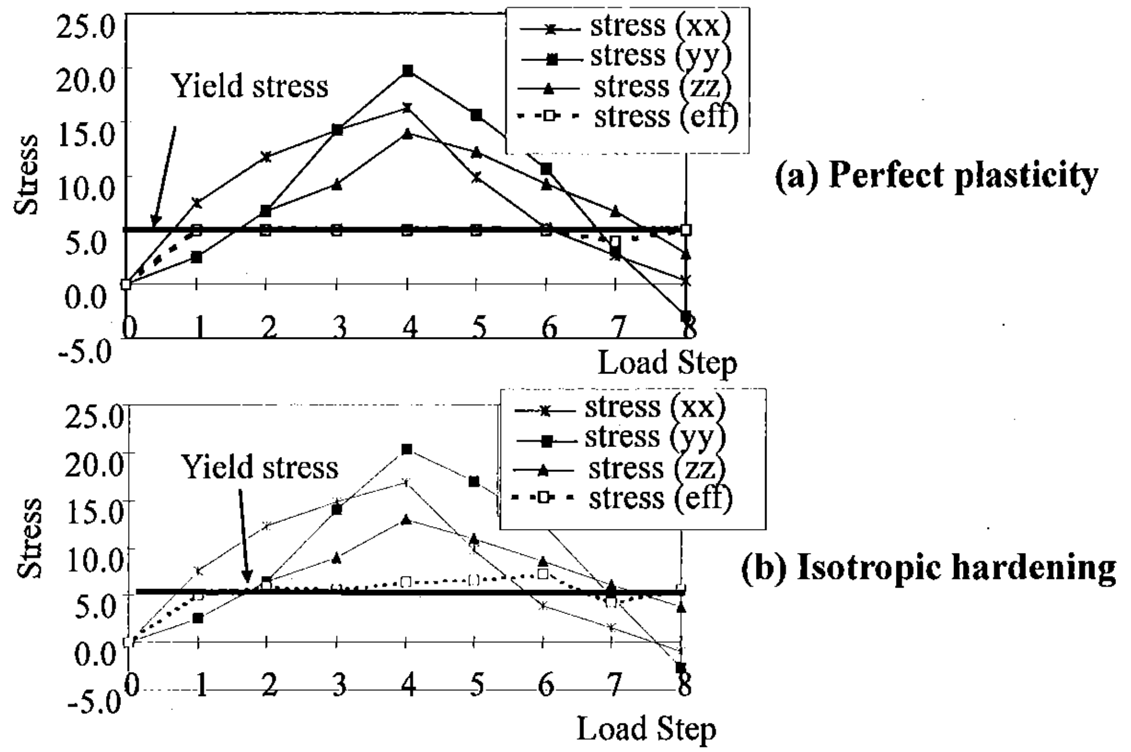
- **Residual stresses**

After the final load step (step 8), although the geometry is returned back to its original shape and is under no external loads, residual stresses are present for both perfect plasticity and isotropic hardening cases.

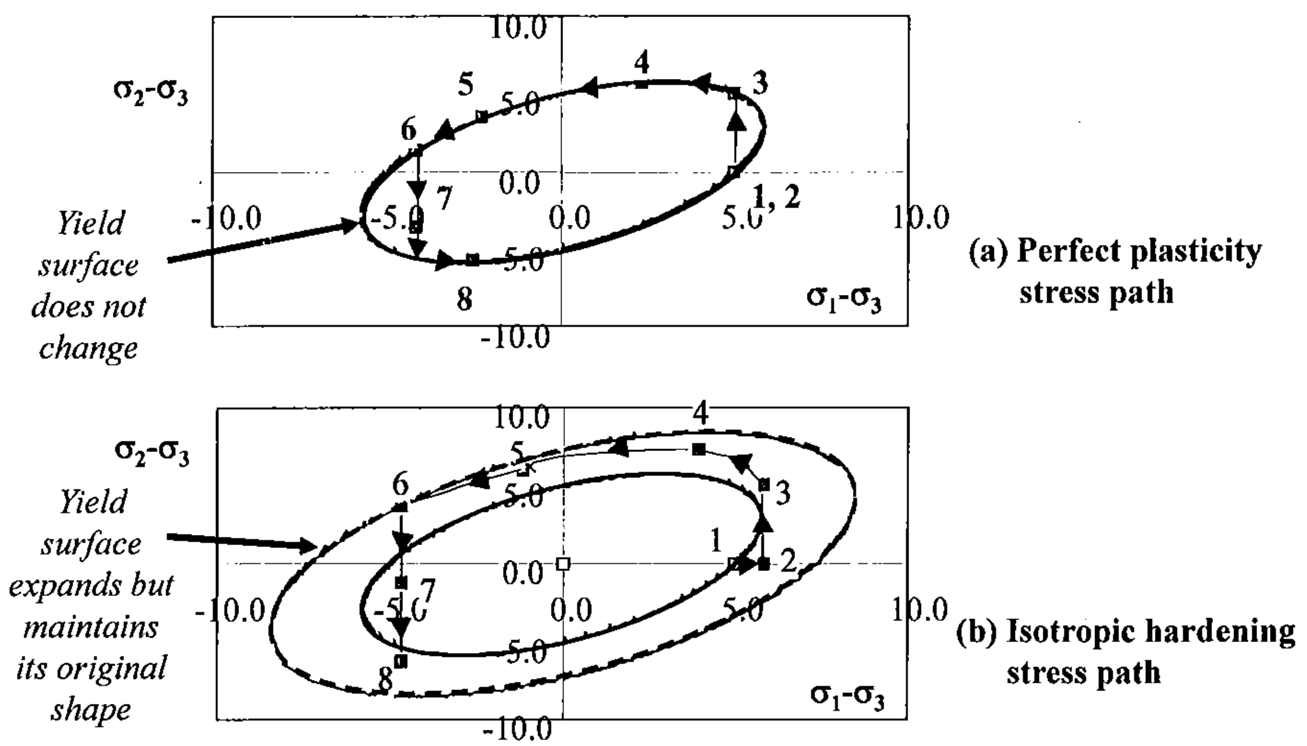
- **Yield surface**

In the perfect plasticity case, the effective stress never exceeds the value of the yield stress, σ_{ys} (here 5.0 N/mm^2), and the yield surface remains in the same position

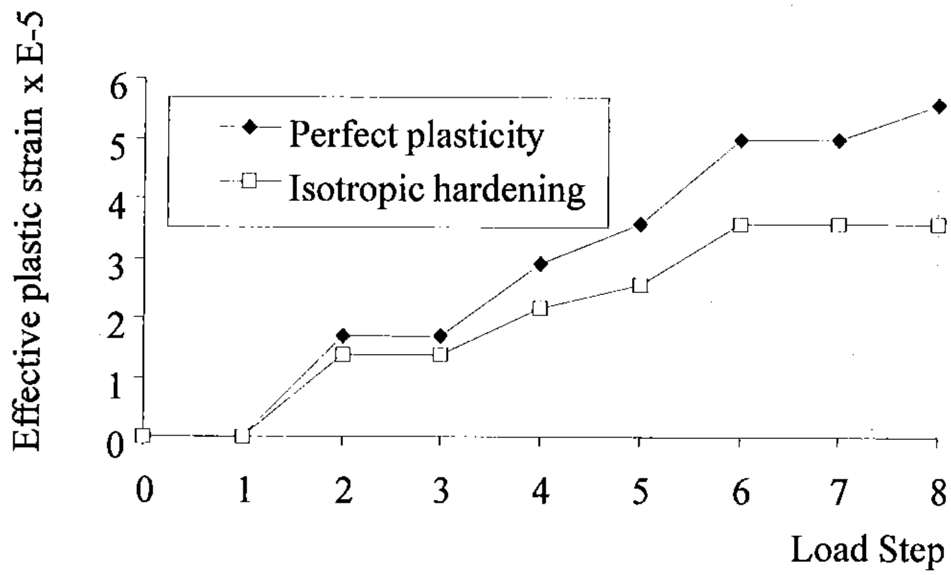
In the isotropic hardening case, the yield stress value is clearly exceeded and the yield surface expands but retains its shape in the isotropic hardening case.



Stress components in the isotropic hardening example



Yield surfaces in the isotropic hardening example



Effective plastic development in the isotropic hardening example

Example 2: Kinematic Hardening Example

Physical Attributes

This example is chosen to demonstrate the differences between kinematic and isotropic hardening, for a full cycle of uniaxial straining involving tensile and compressive loads.

The main attributes of this problem are:

- Biaxial yielding
- Isotropic hardening
- Kinematic hardening
- Plastic flow
- Unloading and subsequent reloading
- Prescribed displacements at all nodes (displacement-control)

- **Modelling Kinematic hardening**
  The kinematic hardening is implemented using an **"overlay model"**. This enables the simulation of kinematic hardening behaviour by overlaying completely elastic elements with elastic perfectly plastic elements. The elements are connected at the nodal points and hence experience identical strain fields.

- **Boundary conditions:**

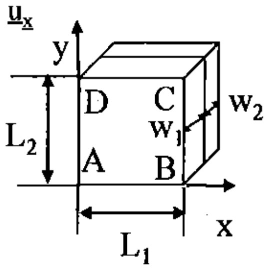
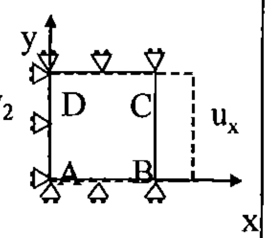
Step 1: Stretching in the x-direction until the plate just yields

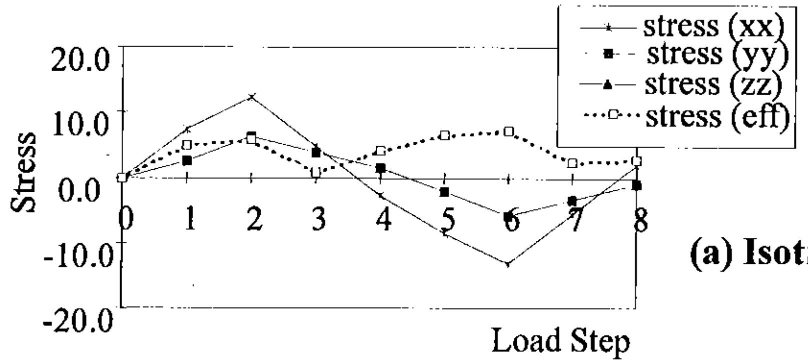
Step 2: Further stretching in the x-direction causing plastic flow, i.e. post-yield behaviour.

Steps 3&4: The plate is compressed to return to its first yield position and then to its original dimensions, resulting in elastic unloading.

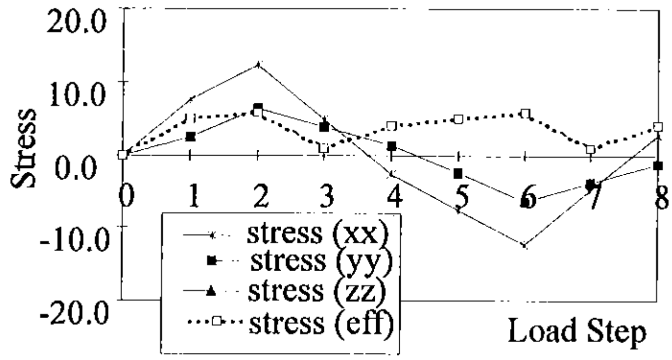
Steps 5&6: Compression is continued further in two steps leading to reverse plastic flow and reverse hardening.

Steps 7&8: The element is stretched in two steps back to its original dimensions.

Geometry	2D plane strain overlay model, $L_1 = L_2 = 100$ mm
Material Properties	$E = 250.0 \times 10^3$ N/mm², $v = 0.25$ **Element 1 : Perfect plasticity; $\sigma_{ys} = 1000.0$ N/mm² ; $w_1 = 0.1724$** **Element 2 : Perfect plasticity; $\sigma_{ys} = 5.0$ N/mm² ; $w_2 = 0.8276$**
Boundary Conditions	$u_x = 0$ on line AD, $u_y = 0$ on line AB $u_y = 0$ on line CD, $u_x = d_x$ on line BC Displacements prescribed in 8 increments ($R=2.5 \times 10^{-5}$) Step — Disp. change Step 1 — $\Delta u_x = R$ — R Step 2 — $\Delta u_x = R$ — 2R Step 3 — $\Delta u_x = -R$ — R Step 4 — $\Delta u_x = -R$ — 0.0 Step 5 — $\Delta u_x = -R$ — -R Step 6 — $\Delta u_x = -R$ — -2R Step 7 — $\Delta u_x = R$ — -R Step 8 — $\Delta u_x = R$ — 0.0  
	No applied forces

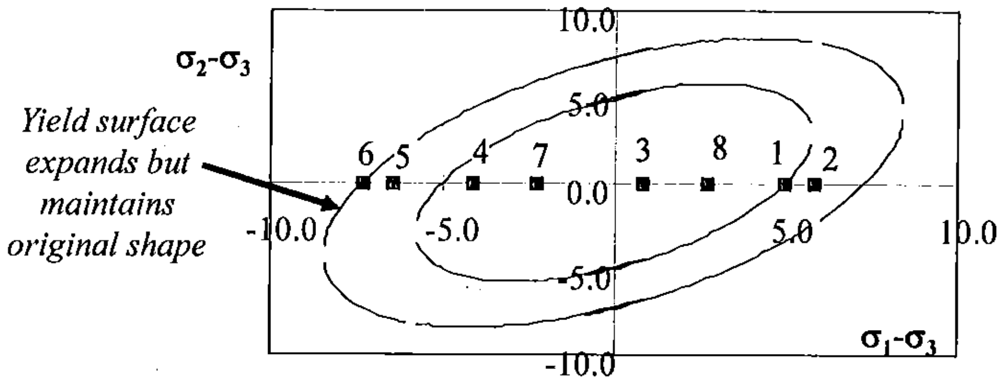


(a) Isotropic hardening

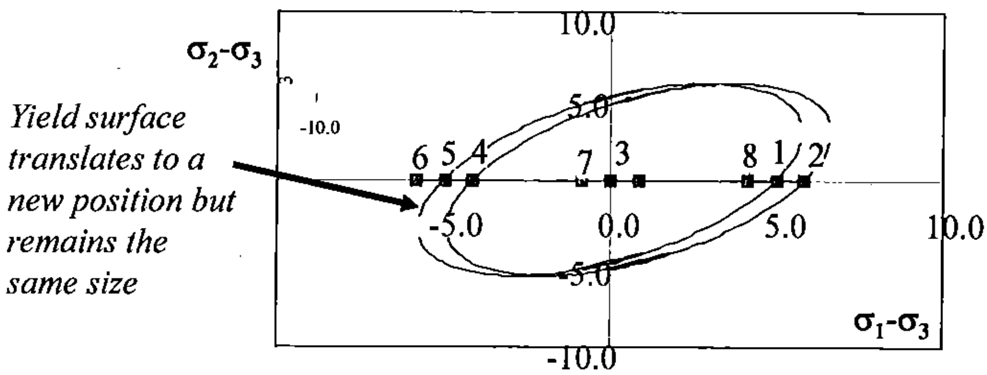


(b) Kinematic hardening

Stress components in the kinematic hardening example



(a) Isotropic hardening stress path



(b) Kinematic hardening stress path

Yield surfaces in the kinematic hardening example

Example 3: Plasticity in a Pressurised Cylinder

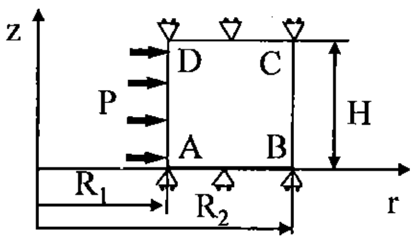
Physical Attributes

This problem is chosen to demonstrate a multi-axial stress situation in which there is a stress gradient over the elements and the plastic zone grows with the increase in load.

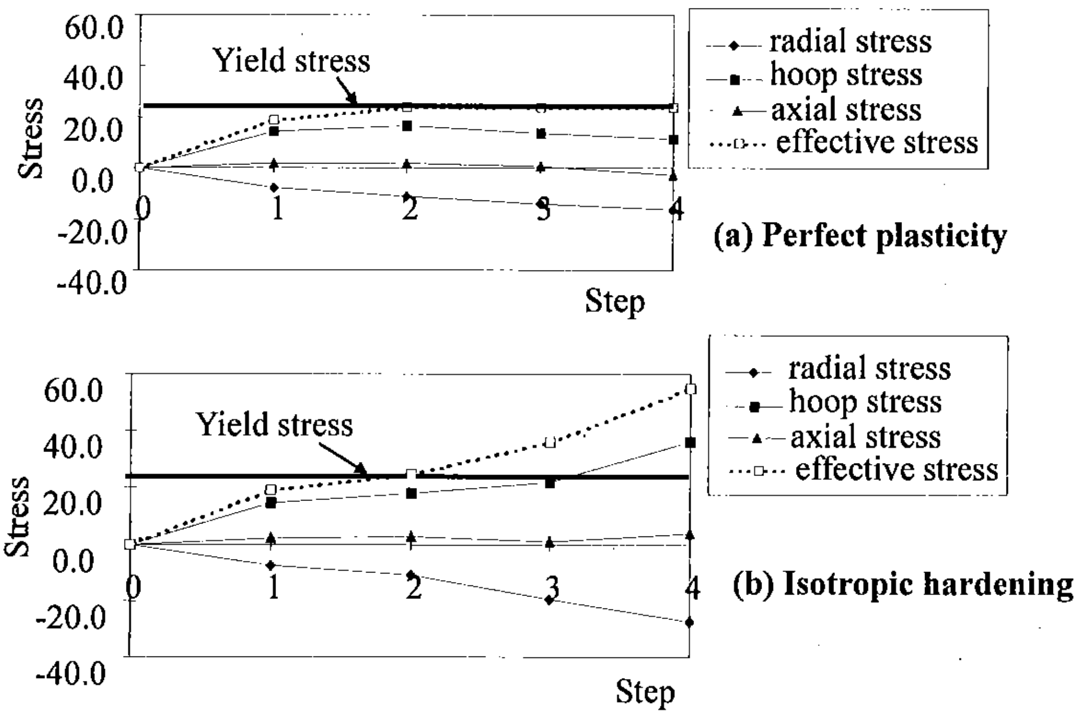
The main attributes of this problem are:

- Axisymmetric yielding
- Isotropic hardening
- Increasing plastic zone with the applied load
- Stress gradients

(Pressurised Cylinder Example/ Continued)

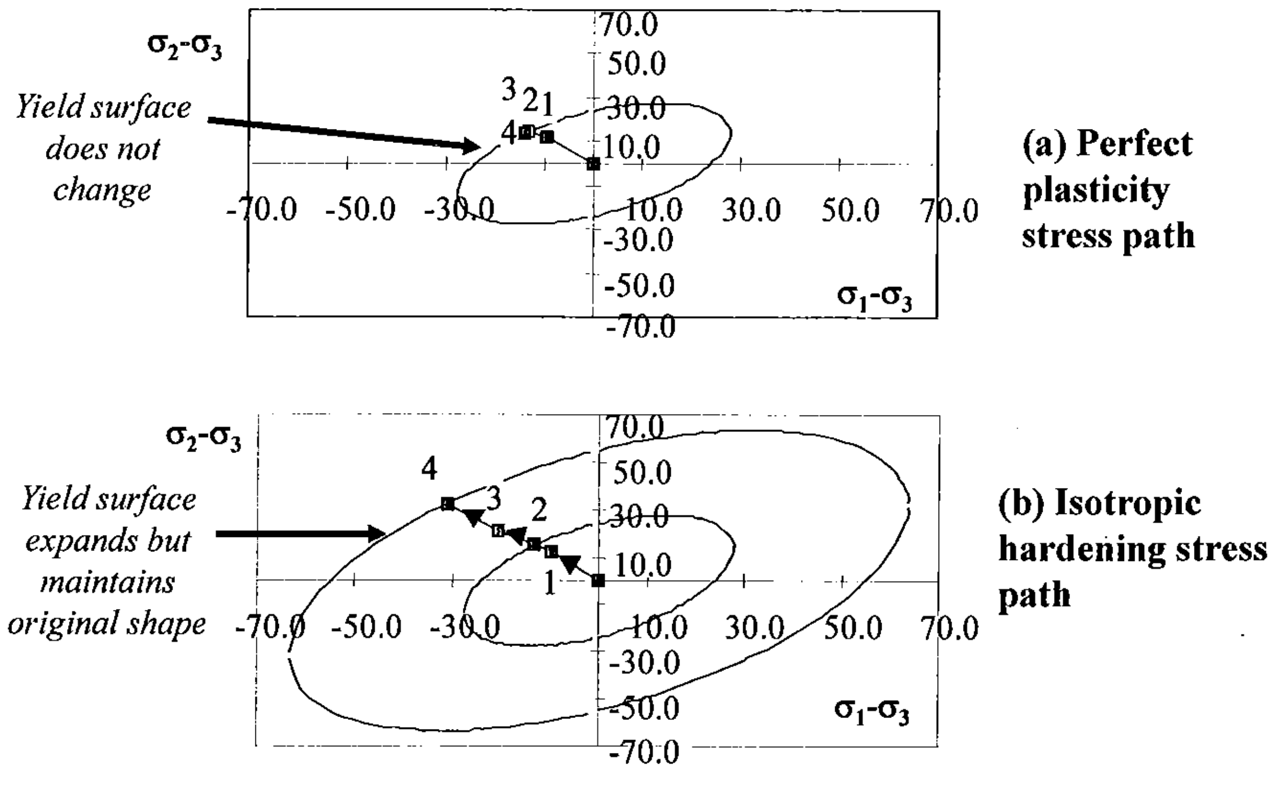
Geometry	Axisymmetric $R_1 = 100$ mm $R_2 = 200$ mm $H = 100$ mm 
Material Properties	$E = 21.0 \times 10^3$ N/mm^2 $\nu = 0.3$ Plasticity model : **(a) Perfect plasticity; $\sigma_{ys} = 24.0$ N/mm^2** **(b) Isotropic hardening; $E_T = 4.2 \times 10^3$ N/mm^2**
Boundary Conditions	$u_z = 0$ on line AB $u_z = 0$ on line DC
Loading	Internal pressure P at the bore (line AD) applied in 4 steps, as follows: (a) Perfect plasticity : P = 10.0, 14.0, 16.6, 19.2 N/mm^2 (b) Isotropic hardening : P = 10.0, 14.0, 24.0, 34.0 N/mm^2

(Pressurised Cylinder Example/ Continued)



Stress components in the pressurised cylinder example

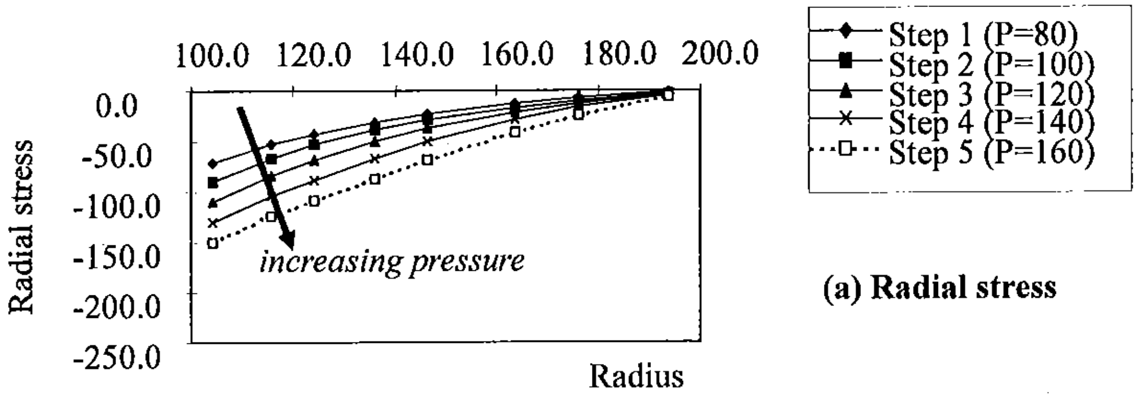
(Pressurised Cylinder Example/ Continued)



Yield surfaces in the pressurised cylinder example

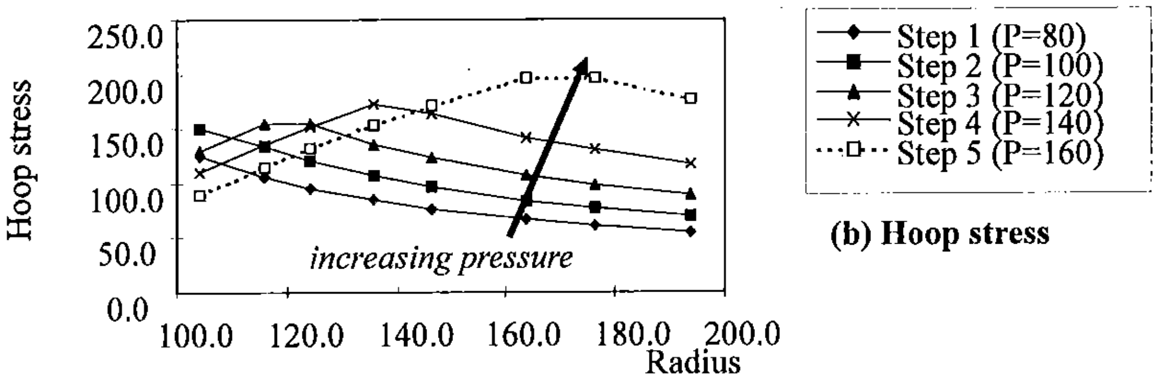
Pressurised Cylinder Example with Increasing Pressure

Perfect plasticity
Yield stress = 208 MPa
Pressure = 160 MPa (applied in 6 steps)

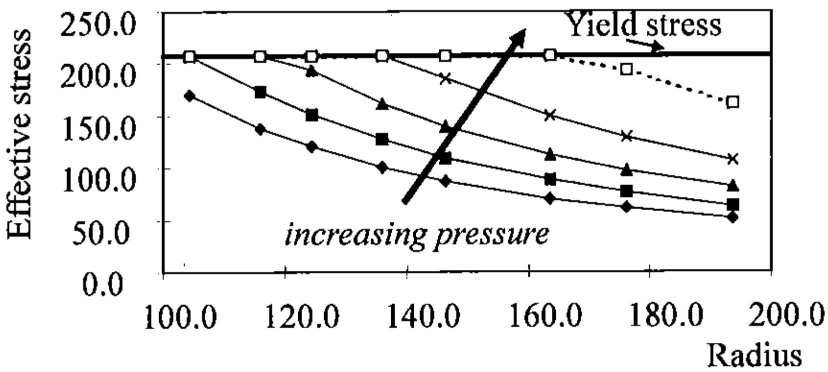


(a) Radial stress

Pressurised Cylinder with Increasing Pressure



(b) Hoop stress



(c) Effective stress

Pressurised Cylinder with Increasing Pressure

Example 4: Cyclic Plasticity

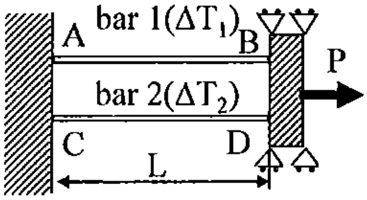
Physical Attributes

This problem demonstrates the effect of cyclic loading on the plastic behaviour in a simple two-bar assembly. The cyclic loading is applied through cycles of thermal heating and cooling.

The main attributes of this problem are:

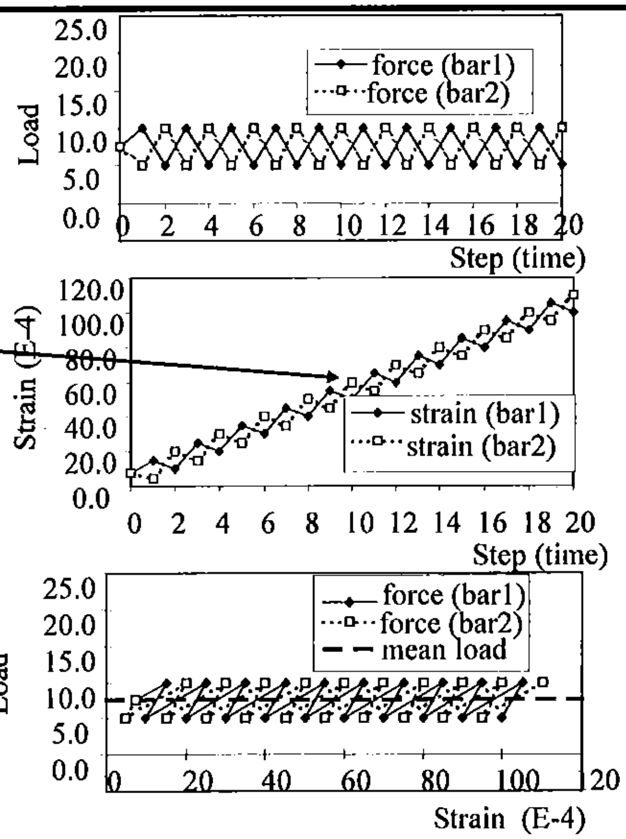
- Cyclic thermal loading
- Ratchetting
- Shakedown
- Alternating plasticity

(Cyclic Plasticity Example/ Continued)

Geometry	2D bars, L = 100 mm Cross-sectional area = 1.0 mm² P = 15.0 N Cyclic temperatures ΔT_1 and ΔT_2 
Material Properties	E = 10.0 x 10^3 N/mm², ν = 0.0 Coefficient of thermal expansion $\alpha = 10^{-5}$ /°C Plasticity model : (a) **Perfect plasticity;** σ_{ys} = 10.0 N/mm² (b) **Kinematic hardening**; E_T = 1.0 x 10^3 N/mm²
Boundary Conditions	$u_x = 0$, $u_y = 0$ at points A and C $u_y = 0$ at points B and D u_x of point B = u_x of point D (multi-point constraint)
Loading	Constant axial force P applied to either point B or D Cyclic temperatures applied in 20 triangular half-cycles: (i) $\Delta T_1 = \pm 100$, $\Delta T_2 = 0$ (ii) $\Delta T_1 = \pm 300$, $\Delta T_2 = 0$ Initial temperatures = 0 °C

(Cyclic Plasticity Example/ Continued)

No limit to plastic strain accumulation



(a) Load vs. Step

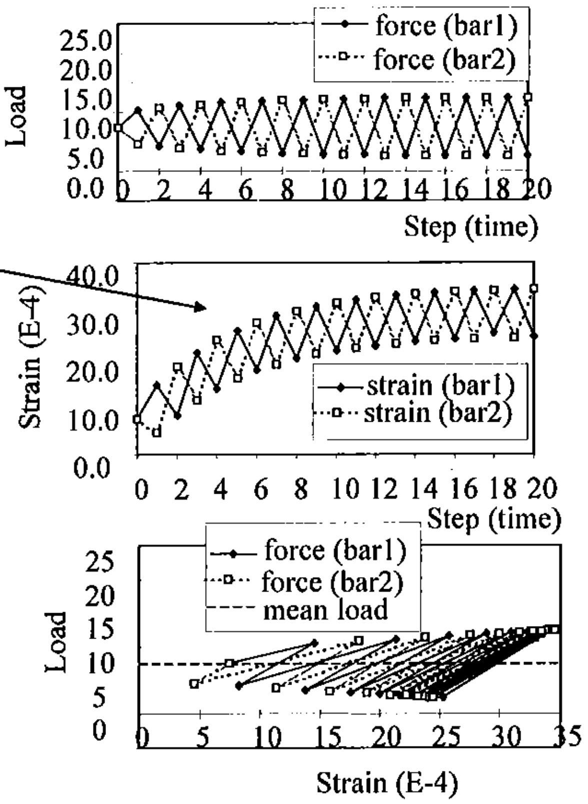
(b) Strain vs. Step

(c) Load vs. Strain

Ratchetting behavior (Perfect plasticity)

(Cyclic Plasticity Example/ Continued)

Shakedown occurs where $\alpha\Delta T$ is less than or equal to twice the plastic strain.



(a) Load vs. Step

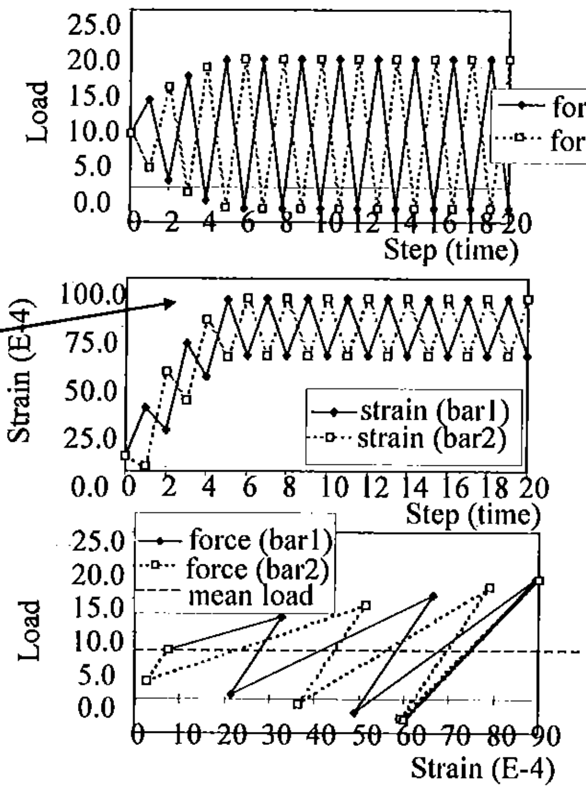
(b) Strain vs. Step

(c) Load vs. Strain

Shakedown behaviour (Kinematic hardening) $\Delta T = \pm 100$ °C

(Cyclic Plasticity Example/ Continued)

Increasing the temperature to $\Delta T = \pm 300\ ^oC$ causes alternating plasticity. Each bar experiences alternating compressive and tensile yield with no progressive increase in the plastic strains.



(a) Load vs. Step

(b) Strain vs. Step

(c) Load vs. Strain

Alternating plasticity behavior (Kinematic hardening) $\Delta T = \pm 300\ ^oC$

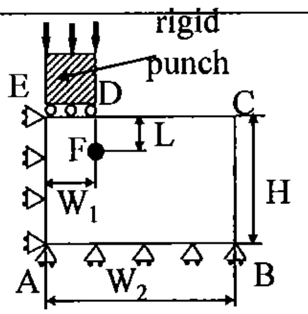
Example 5: Rigid Punch Plasticity

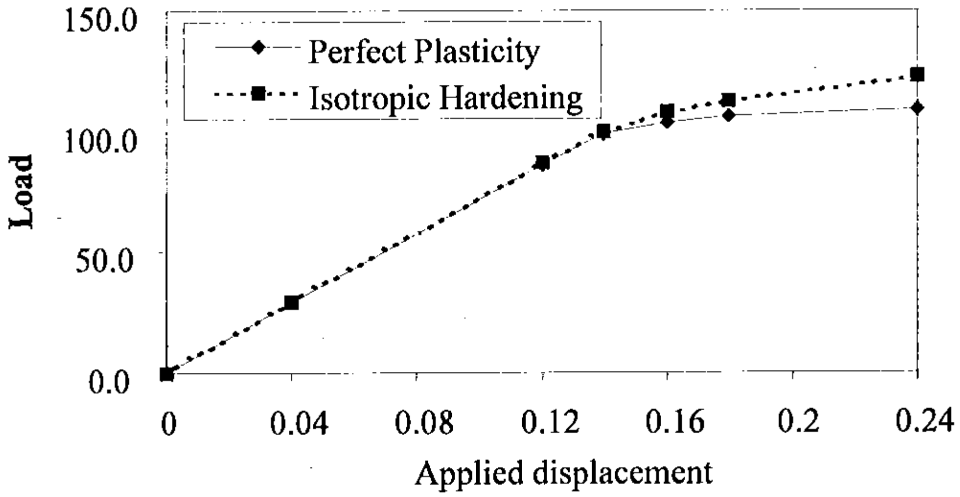
Physical Attributes

This problem is chosen to demonstrate the effects of stress concentration, rapidly changing stresses and stress redistribution on the nucleation and spread of plasticity zones in a continuum.

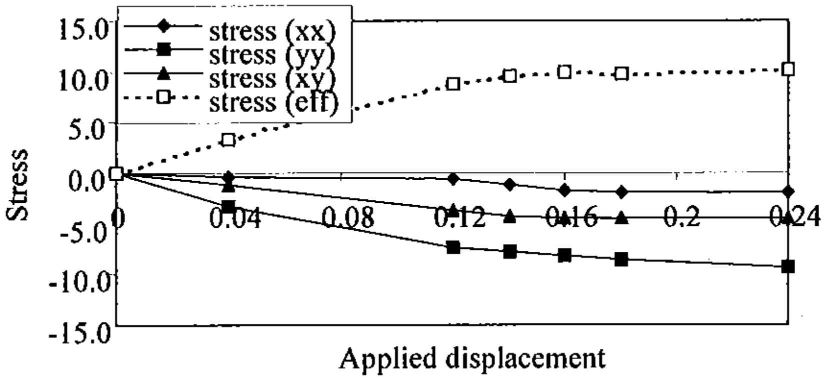
The main attributes of this problem are:

- Stress concentration
- Stress and strain redistribution in a continuum
- Increasing plastic zone with the applied load

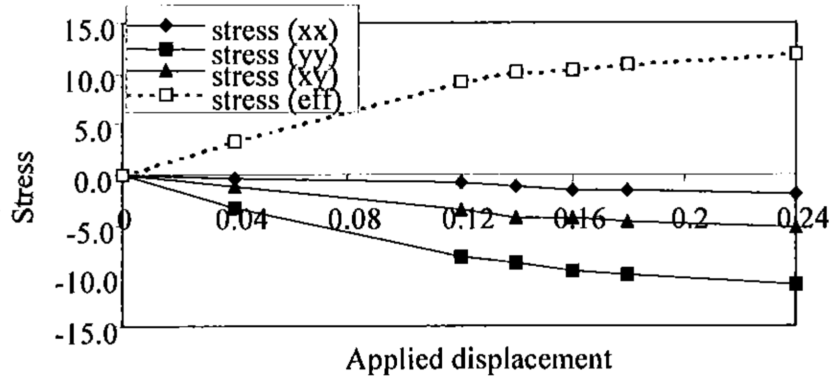
Geometry	2D plane strain $W_1 = 80$ mm $W_2 = 200$ mm $H = 160$ mm $L = 20$ mm	
Material Properties	$E = 1.0 \times 10^3$ N/mm², $\nu = 0.3$ Plasticity model : **(a) Perfect plasticity; $\sigma_{ys} = 1.0$ N/mm²** **(b) Isotropic hardening; $E_T = 0.1 \times 10^3$ N/mm²**	
Boundary Conditions	$u_x = 0$ on line AE $u_y = 0$ on line AB $u_y = \delta$ on line DE, applied in 6 steps δ applied in 6 steps, as follows: d = 0.04, 0.12, 0.14, 0.16, 0.18, 0.24 mm	
Loading	No applied forces	



Load vs. punch displacement for the rigid punch plasticity example

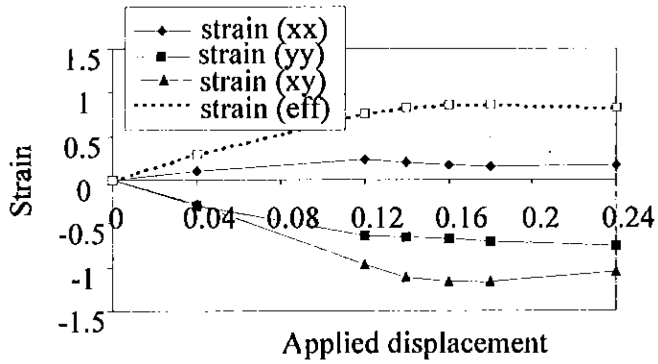


(a) Perfect plasticity

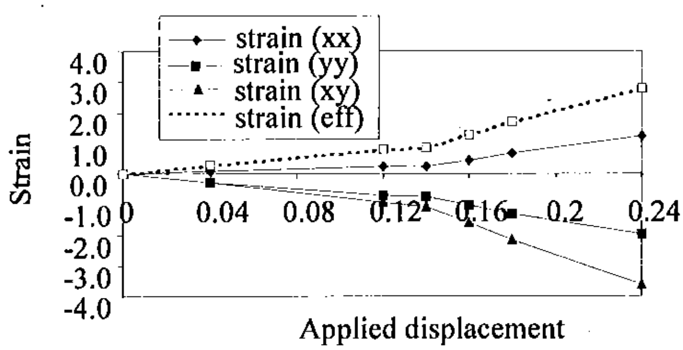


(b) Isotropic hardening

Stress vs. punch displacement (at point F) for the rigid punch plasticity example



(a) Perfect plasticity



(b) Isotropic hardening

Strain vs. punch displacement (at point F) for the rigid punch plasticity example

2.7 Summary of Key Points

- Plastic behaviour is usually assumed to be **independent of time.** However, in real life, strain rates can influence plasticity.
- **True stresses and strains** should be used for plastic stress-strain behaviour.
- Stresses can be decomposed into **hydrosatic stresses and deviatorc stresses.** Only the deviatoric stresses are responsible for plastic behaviour through the shear component.
- Since plasticity is assumed to be independent of time, time may be considered a **'pseudo-time'** which refers to the sequence of the load increments.
- A **yield surface** must be defined. For a **perfectly plastic material,** there is no change in the yield surface during plastic deformation.

(Summary/ Continued)

- In post-yield behaviour, the yield surface may be modified. For an **isotropic hardening material,** the yield surface increases in size with increasing plastic strain but maintains its original shape.
- For a **kinematic hardening material,** the original yield surface is translated to a new position in the stress space as the plastic strain is increased, with no change in size or shape.
- **Bauschinger effect** is observed when $(\sigma ys)_{compression}$ is less than $(\sigma ys)_{tension}$
- Under Cyclic and Reversed Loading, three types of behaviours can be observed; **elastic shakedown, ratchetting and alternating plasticity.**
- In FE formulations, the total **strain rate** or increment can be divided into two parts; an elastic recoverable part and a plastic irrecoverable part as follows:

$$\left\lfloor \dot{\varepsilon}^{total} \right\rfloor = \left\lfloor \dot{\varepsilon}^{e} \right\rfloor + \left\lfloor \dot{\varepsilon}^{p} \right\rfloor$$

Part 3

Creep
(Time-dependent Material Non-linearity)

Lecture Outline

3.1 Introduction

- Creep is a time-dependent deformation caused by the thermally activated movement of vacancies and dislocations, under load, usually at high temperatures.
- Creep behaviour in metals usually occurs at T/Tm >0.4, where T and Tm are the absolute operating temperature and absolute melting temperature, respectively.
- The study of creep behaviour is important in high-temperature engineering applications, such as turbine disks, nuclear reactors and aero-engines.

Creep of lead pipes in old buildings
(Reference: Ashby and Jones [1980])

3.2 Uniaxial Fixed-Load Creep Behaviour

In general, the stress-strain relationship in a uniaxial creep test, under a constant load, can be represented by the following equation:

$$\varepsilon^c = f(\sigma, T, t)$$

where
ε^c is the creep strain
σ is the nominal stress
t is the time
T is the temperature.

This equation is often further simplified by separating the effects of stress, temperature and time as follows:

$$\varepsilon^c = f_1(\sigma)\, f_2(T)\, f_3(t)$$

- Many mathematical models have been proposed for the stress dependence of the secondary creep behaviour, for example:

$$\dot{\varepsilon}^{c}_{min} = A\,\sigma^{n} \quad \textit{(Norton)}$$

$$\dot{\varepsilon}^{c}_{min} = B\,\sinh(\beta\sigma) \quad \textit{(Prandtl)}$$

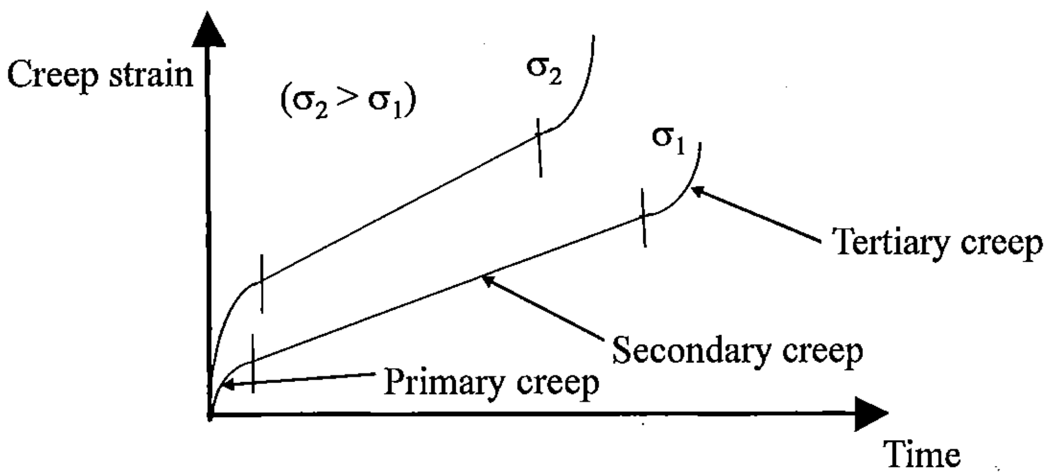
- The dot displayed above any variable indicates the rate of change of the variable with respect to time.

- The time dependence has been represented by several forms, for example:

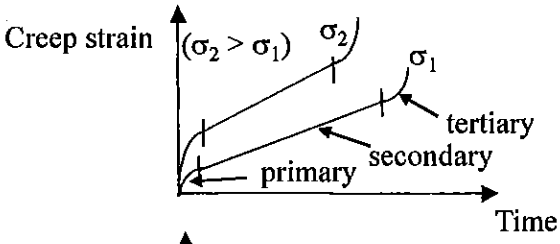
$$f_2(t) = t \quad \textit{(secondary creep)}$$

$$f_2(t) = b\,t^{m} \quad \textit{(Bailey)}$$

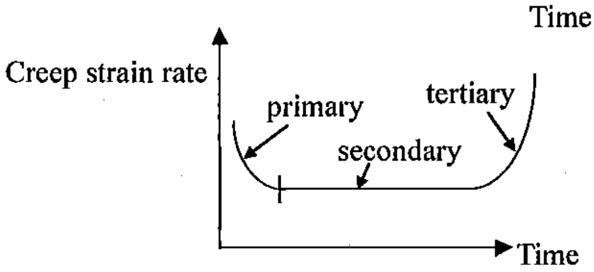
where m (less than 1) is used for the primary creep stage.



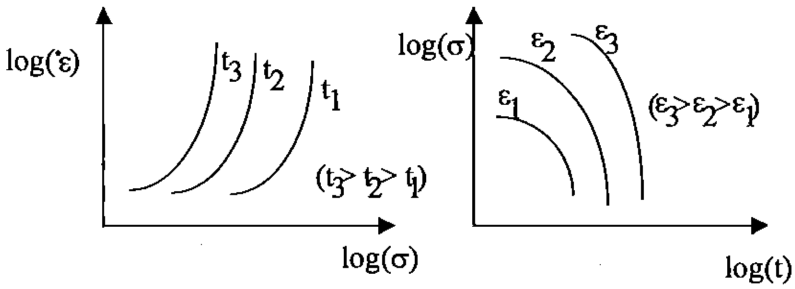
Primary, secondary and tertiary creep stages



(a) Creep strain vs. time



(b) Creep strain rate vs. time



(c) Isochronous and isostrain creep curves

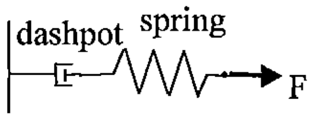
Uniaxial Creep curves

A combination of the Norton-Bailey equations can be used to represent primary and secondary creep in isothermal conditions as follows:

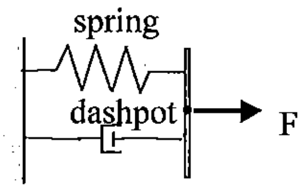
$$\varepsilon^c = A\sigma^n t^m \quad (Norton - Bailey)$$

To represent combined primary and secondary creep behaviour, a combination of the Bailey and secondary creep models can be used as follows:

$$\varepsilon^c = A_1\sigma^{n1} t + A_2\sigma^{n2} t^m$$



(a) Maxwell model



(b) Kelvin model

Rheological Models used to represent creep and viscoelasticity

3.3 Variable-Load Creep Behaviour

(a) Time Hardening

The creep strain rate is assumed to depend on the current stress and the time from the start of the test.

By differentiating the Norton-Bailey equation, the rate of change of the creep strain can be written as follows:

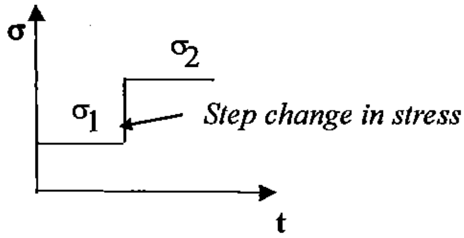
$$\dot{\varepsilon}^c = \frac{d\,\varepsilon^c}{dt} = m\,A\,\sigma^n\,t^{(m-1)}$$

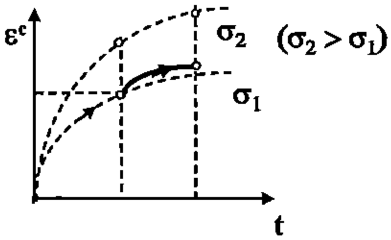
(b) Strain Hardening

The creep strain rate is assumed to depend on the current stress and the accumulated creep strain.

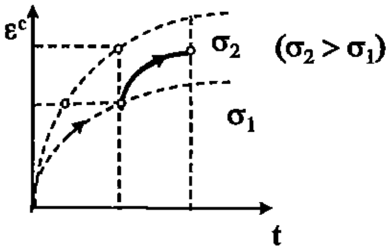
An equation can be derived for the creep strain rate by substituting the time from the Norton-Bailey equation into the time hardening equation, as follows:

$$\dot{\varepsilon}^c = \frac{d\,\varepsilon^c}{dt} = m\,A^{\frac{1}{m}}\,\sigma^{\frac{n}{m}}\left(\varepsilon^c\right)^{\frac{m-1}{m}}$$





(a) Time hardening



(b) Strain hardening

Time hardening and strain hardening Assumptions

<u>Comparison of Time and Strain Hardening</u>

- Time hardening assumption, although physically unrealistic since time is measured from an arbitrary origin, is widely used and can give reasonable results in situations of constant (or nearly constant) stress.

- Strain hardening assumption is considered more accurate and agrees well with experimental tests.

3.4 Stress Relaxation In Creep

Creep stress relaxation occurs when a uniaxial specimen is loaded up to a stress level of σ_o and then held at **constant strain** ε_o (where $\varepsilon_o = \sigma_o/E$). Therefore, the strain at a given time is composed of an elastic component and a creep component, as follows:

$$\varepsilon_o = \varepsilon^e + \varepsilon^c = constant$$

Differentiating this equation with respect to time:

$$\frac{d\,\varepsilon_o}{dt} = 0 = \dot{\varepsilon}^e + \dot{\varepsilon}^c$$

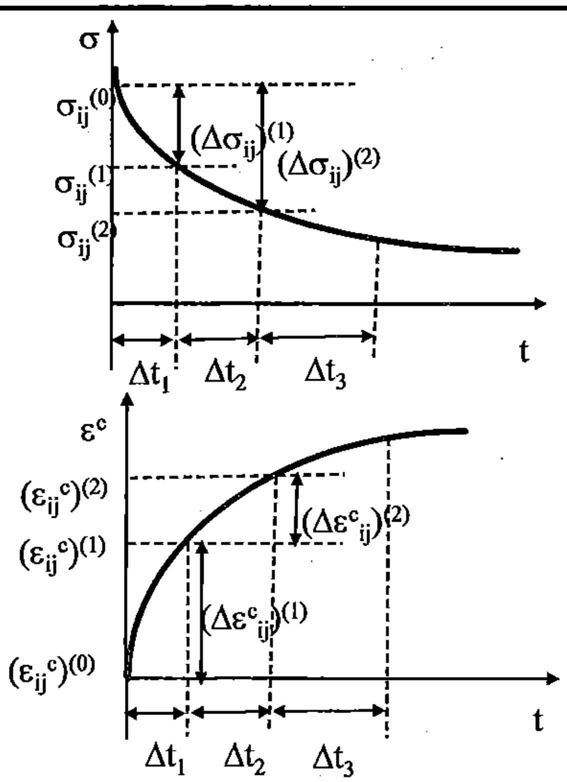
Assuming a uniaxial Norton-Bailey creep equation, and **a time hardening assumption**, this equation becomes:

$$0 = \frac{1}{E}\frac{d\sigma}{dt} + m\,A\,\sigma^n\,t^{(m-1)}$$

This equation can be integrated to give the following expression for the change in stress with time:

$$\sigma = \left[\frac{1}{\sigma_o^{(n-1)}} + (n-1)\,AE\,t^m\right]^{\frac{-1}{(n-1)}}$$

(Stress Relaxation In Creep/ Continued)



Schematic representation of a stress relaxation numerical algorithm

3.5 Multi-Axial Creep Formulations

- Multiaxial formulations of the creep laws are **similar to the multiaxial plasticity formulations**, and are usually based on the von Mises effective stress criterion and the **Prandtl-Reuss flow rule.**

- The **effective (Von Mises) stress** expressions used in plasticity are also applicable in creep.

- Similarly, an expression for the effective creep strain rate can be written in terms of the Cartesian creep strain rates as follows:

$$\dot{\varepsilon}^{c}_{eff} = \frac{\sqrt{2}}{3}\left[\left(\dot{\varepsilon}^{c}_{xx} - \dot{\varepsilon}^{c}_{yy}\right)^2 + \left(\dot{\varepsilon}^{c}_{yy} - \dot{\varepsilon}^{c}_{zz}\right)^2 + \left(\dot{\varepsilon}^{c}_{xx} - \dot{\varepsilon}^{c}_{zz}\right)^2 + 6\left(\dot{\varepsilon}^{c}_{xy}\right)^2 + 6\left(\dot{\varepsilon}^{c}_{yz}\right)^2 + 6\left(\dot{\varepsilon}^{c}_{xz}\right)^2 \right]^{1/2}$$

- As in plasticity, experimental evidence for metals indicates that, for small strains, creep is a **constant volume process** (

(Multi-Axial Creep Formulations/ Continued)

Under multi-axial conditions, the uniaxial Norton-Bailey creep law can be expressed in terms of "effective" quantities, as follows:

$$\varepsilon^{c}_{eff} = A\left(\sigma_{eff}\right)^{n} t^{m}$$ (

A multi-axial expression, based on the time hardening rule and the Norton-Bailey creep law, can be written in tensor notation as follows:

$$\dot{\varepsilon}^{c}_{ij} = \frac{3}{2} m\, A\left(\sigma_{eff}\right)^{(n-1)} S_{ij}\, t^{(m-1)}$$

This expression can be modified for the strain hardening rule as follows:

$$\dot{\varepsilon}^{c}_{ij} = \frac{3}{2} m\, A^{\frac{1}{m}} \left(\sigma_{eff}\right)^{\frac{n-m}{m}} S_{ij} \left(\varepsilon^{c}_{eff}\right)^{\frac{m-1}{m}}$$

This multi-axial creep formulation is reasonable provided that the signs of the significant stress components do not change, and the stress ratios do not exhibit large changes.

3.6 Time Marching Procedures

In creep problems, the strain is a function of time according to the constitutive equations of creep and the solution is marched in time by dividing the total creep time into small time intervals.

The strain rate or the displacement rate is assumed to be a function of the current displacement and time, as follows:

$$\dot{u} = \frac{du}{dt} = f(u, t)$$

where u is the displacement vector.

(i) <u>Explicit Time Marching Scheme</u>
(also called Forward-Difference or Euler Scheme)

This method uses a simple finite difference approximation for the differential of the displacement with respect to time, as follows:

$$\dot{u}_t = \left(\frac{du}{dt}\right)_t = \frac{u_{t+\Delta t} - u_t}{\Delta t}$$

The following "recurrence" formula is used:

$$u_{t+\Delta t} = u_t + \Delta t \left[f(u_t, t) \right]$$

This scheme is known as "forward-difference" because it arrives at the new value of the displacement, $u_{t+\Delta t}$, by moving forward along the tangent at the previous point u_t.

This scheme is also called "explicit" because the new values of the displacement are immediately calculated from the old values (the initial values), i.e. the new prediction of displacement, $u_{t+\Delta t}$, does not appear in the right hand side of the equations.

Advantages and Disadvantages of the Explicit Scheme

Advantage:
It is simple to program and economical.

Disadvantage:
It needs very small time steps to remain stable, i.e. if the time step is large, instability may occur.

(ii) Implicit Time Marching Scheme
(also called the Backward-Difference or Theta Scheme)

This is a more general approach where, to obtain the solution for the displacement at the end of the time step, $u_{t+\Delta t}$, the slope at the initial point, u_t, and the slope at an **intermediate point,** $u_{t+\theta\Delta t}$, are used, as follows:

$$u_{t+\Delta t} = u_t + \Delta t \left[f\left(u_{t+\theta\Delta t}, t+\theta\,\Delta t\right)\right]$$

where

$$u_{t+\theta\Delta t} = \theta\, u_{t+\Delta t} + (1-\theta)\, u_t$$

The parameter θ is effectively a **weighted average** of approximations at the start and end of the time interval.

Possibilities of θ

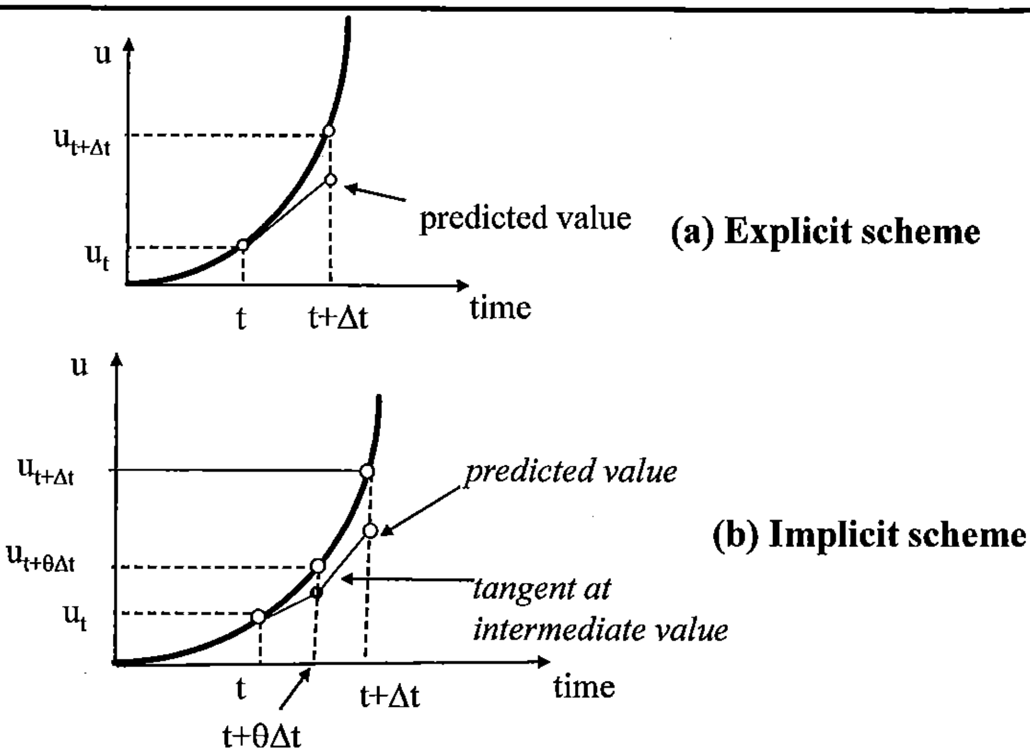
- When $\boldsymbol{\theta} = \mathbf{0}$, i.e. no intermediate point is used, the theta method becomes the explicit method.
- For $\boldsymbol{\theta} > \mathbf{0.5}$, this method becomes *unconditionally stable* for all values of Δt.
- When $\boldsymbol{\theta} = \mathbf{0.5}$, the theta method becomes the Crank-Nickolson method which is an established numerical procedure.
- When $\boldsymbol{\theta} = \mathbf{1}$, the theta method becomes the *implicit* method, i.e.

$$u_{t+\Delta t} = u_t + \Delta t \left[f\left(u_{t+\Delta t}, t + \Delta t\right) \right]$$

Advantages and Disadvantages of the Implicit Scheme

Advantage:
The implicit method is ***unconditionally stable***. This means that this method remains stable even if large time steps are used, although small time steps are more accurate.

Disadvantage:
Each time step requires **more computation time** than the explicit scheme. In FE creep formulations the time step must remain sufficiently small to ensure adequate solution accuracy for the stresses and strains.



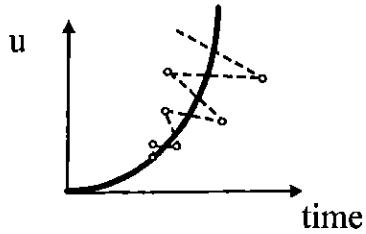
Schematic representation of the explicit and implicit time marching schemes

Convergence and stability

A robust time marching algorithm must ensure that the solution is both convergent and stable.

(i) ***Convergence*** is achieved if the numerical predictions approach the exact solution as the size of the time interval approaches zero.

(ii) ***Stability*** is achieved if the numerical errors made at a given time interval do not cause increasingly larger errors as the solution is marched forward in time. Instead, the numerical errors should damp out with time.



Instability example

3.7 Outline of FE Creep Algorithms

In creep analysis, the total strain increment at any given time consists of:

(i) Elastic strain component [ε^e]

(ii) Creep strain component [ε^c].

(iii) Plastic strain component [ε^p] (If plasticity is included)

$$\lfloor \Delta \varepsilon^{total} \rfloor = \lfloor \Delta \varepsilon^{e} \rfloor + \lfloor \Delta \varepsilon^{c} \rfloor + \lfloor \Delta \varepsilon^{p} \rfloor$$

(FE Creep Algorithms/ Continued)

FE formulations for creep problems based on the initial strain approach usually adopt the following procedure:

(i)Apply the mechanical loads and displacements, and solve the equations to determine the nodal displacements and the elastic (or elastic-plastic) stress distribution at $t = 0$.

(ii)For a small time increment, Δt, assume that the stresses remain constant, and determine the creep strain rate and the creep strain increments (initial strain) using the multiaxial creep law for the material.

(iii) From the creep strain increment, determine a creep load vector and solve the equations to determine a displacement increment Δu, from which an increment in stress $\Delta\sigma$ is determined.

(iv)The resulting $\Delta\sigma$ must be suitably small when compared to the existing stresses at the beginning of the time increment. If not, the time increment is reduced and the procedure is repeated until $\Delta\sigma$ is sufficiently small.

(v)Add the increments $\Delta\sigma$ and Δu to the existing stresses and displacements to obtain the starting values for the next time increment.

(vi) Apply another time increment and repeat the above steps until the final time is reached.

Automatic time steps

- Solution accuracy can be assessed by **limiting the changes in effective stress** in each time interval to an acceptable range.
- If the changes in effective stress are deemed to be too large, the **time increment is reduced** and the calculation repeated.
- If stresses approach a **steady-state situation**, changes of stress during a specified time interval will reduce as time proceeds. Therefore, provided that the stability limit is not exceeded, **time increments may be allowed to increase** as time proceeds, while maintaining the same accuracy of solution.
- **Automatic procedures** for increasing time intervals can be incorporated in FE codes. In the simplest form, the time interval may be simply increased by 1.5 or 2 times the previous successful time interval.

Creep analysis in FE codes

- FE programs usually request **uniaxial creep data**, allowing the multiaxial form of the Norton-Bailey creep law to be implemented.
- Many FE programs also allow **more sophisticated creep laws** to be used.
- Some FE codes allow the user to specify a **user-defined subroutine** in which a more complex creep law can be specified.
- More sophisticated **convergence criteria** are necessary for problems with very non-uniform stresses, or if the stress sensitivity is high, i.e. a large value of n in the creep law.
- Sophisticated procedures must be adopted when tertiary creep or a **continuum damage mechanics** approach is employed.

3.8 Creep Examples

Example 1: Constant-load creep

Example 2: Stress relaxation in creep

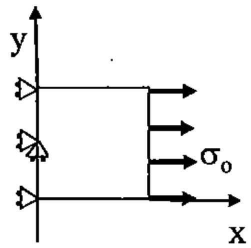
Example 3: Variable-load creep

Example 4: Multi-axial pressurised cylinder creep

Example 1: Constant-Load Creep

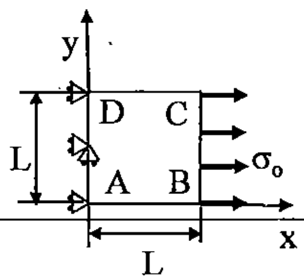
Physical Attributes

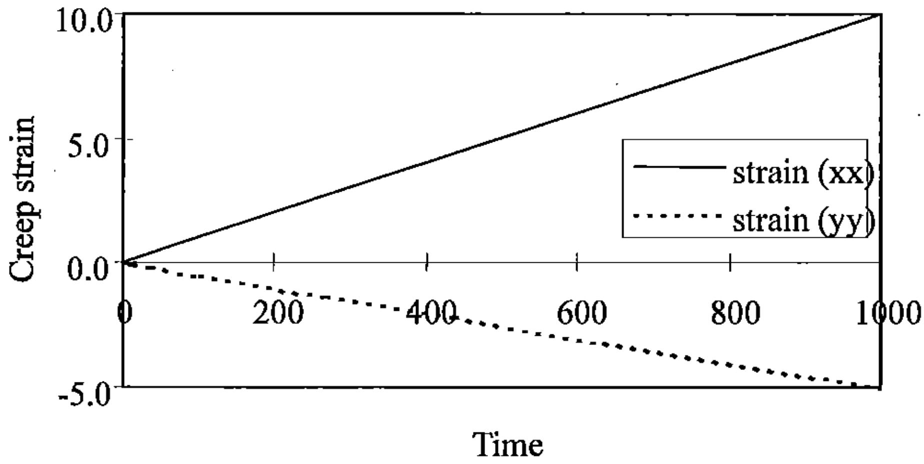
- Uniaxial load situation
- Secondary and primary creep laws
- A combined primary-secondary creep law
- A user-subroutine to specify a non-standard creep law



(Constant-Load Creep Example/ Continued)

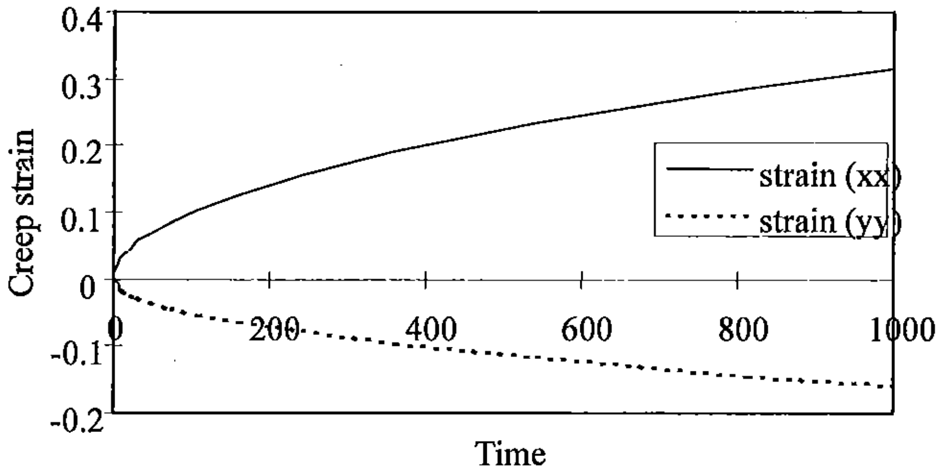
Geometry	2D Plane stress L = 100 mm σ_o = 100 or 200 N/mm^2 Total creep time = 1000 hours
Material Properties	E = 200 x 10^3 N/mm^2, v = 0.3 Creep Law : Case (a) $\varepsilon^c = A\, \sigma^n\, t^m$ A = 3.125 x 10^{-14} (t in hour, σ in N/mm^2) m=1, n = 5 Case (b) As Case (a), except m=0.5 Case (c) $\varepsilon^c = A_1\, \sigma^{n1}\, t + A_2\, \sigma^{n2}\, t^m$ $A_1 = 10^{-16}$, $A_2 = 10^{-14}$ (t in hour, σ in N/mm^2) $n_1 = n_2 = 5$
Boundary Conditions	$u_x = 0$ on line AD $u_y = 0$ on the midpoint of line AD
Loading	Prescribed tensile stress σ_o on line BC σ_o = 200 N/mm^2 (Case a and b) or 100 N/mm^2 (Case c)





(a) Secondary creep

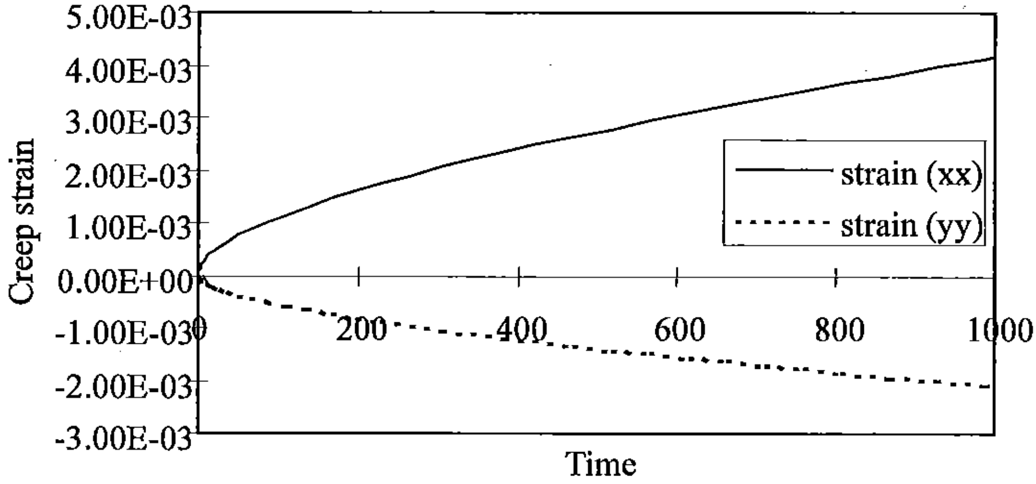
Reference solutions for the constant-load creep problem



(b) Primary creep

Reference solutions for the constant-load creep problem

(Constant-Load Creep Example/ Continued)



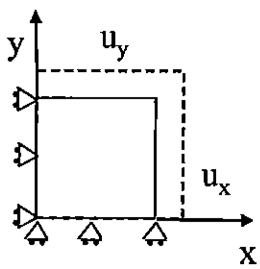
(c) Combined primary and secondary creep

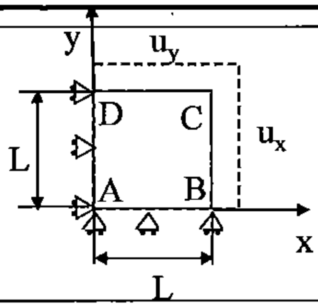
Reference solutions for the constant-load creep problem

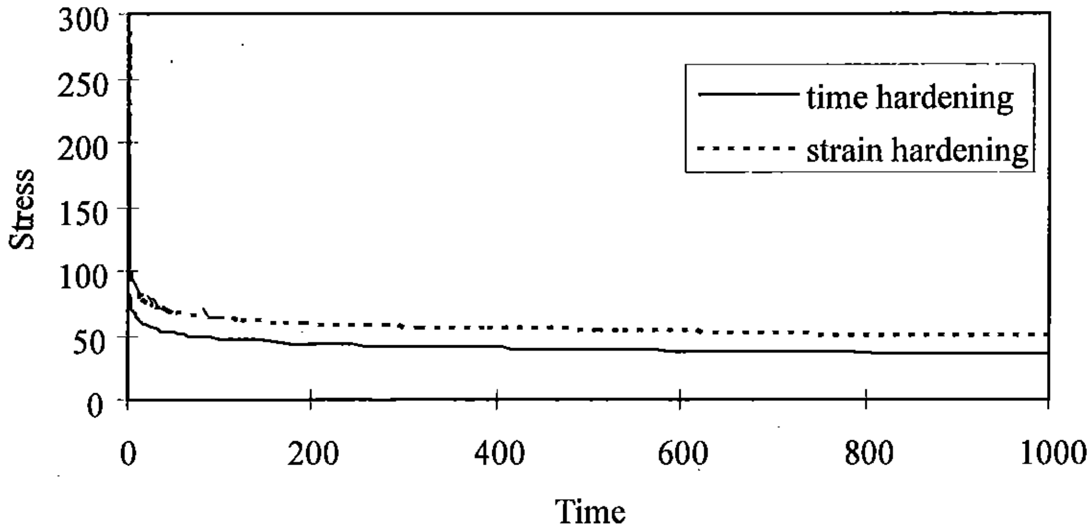
Example 2: Stress Relaxation in Creep

Physical Attributes

- Biaxial and triaxial stress states
- Stress relaxation
- Time and strain hardening
- Steep stress gradient with time



Geometry	L = 100 mm Total creep time = 1000 hours 2D Plane stress $u_x = 0.1$ mm, $u_y = 0.1$ mm 
Material Properties	$E = 200 \times 10^3$ N/mm^2 $v = 0.3$ Creep Law : $\varepsilon^c = A\,\sigma^n\,t^m$ $A = 3.125 \times 10^{-14}$ (t in hour, σ in N/mm^2) $n = 5$ $m = 0.5$
Boundary Conditions	Case (a) and Case (b): $u_x = 0$ on line AD $u_y = 0$ on line AB $u_x = 0.1$ on line BC $u_y = 0.1$
Loading	No applied forces

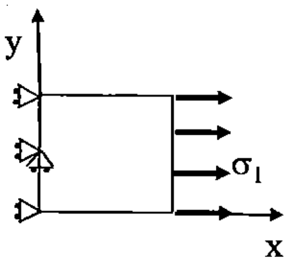


Reference solutions for the stress relaxation creep problem

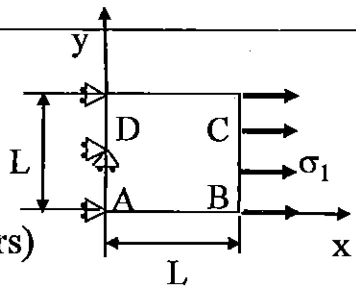
Example 3: Variable-Load Creep

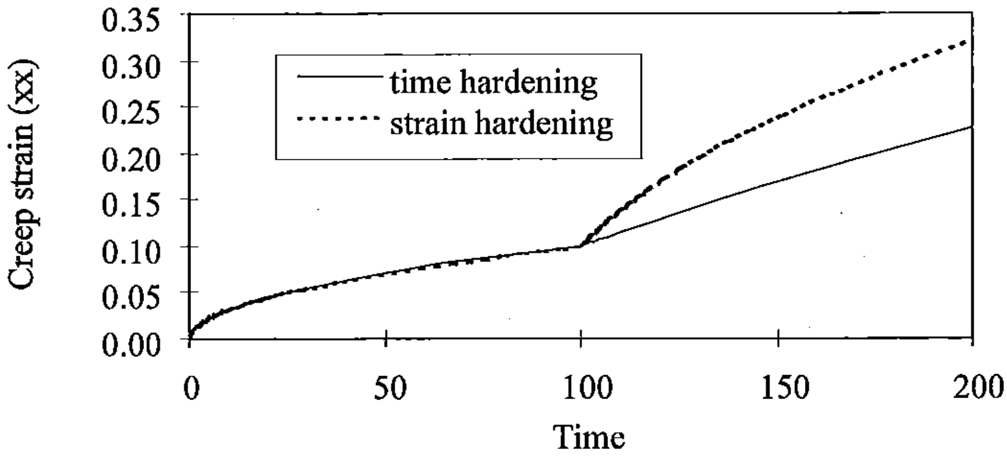
Physical Attributes

- Uniaxial and biaxial stepped loading
- Time hardening
- Strain hardening



(Variable-Load Creep Example/ Continued)

Geometry	2D Plane stress L = 100 mm Total creep time = 200 hours **Uniaxial load** $\sigma_1 = 200$ N/mm^2 (t = 0-100 hours) $\sigma_1 = 250$ N/mm^2 (t >100 hours) 
Material Properties	$E = 200 \times 10^3$ N/mm^2 $\nu = 0.3$ Creep Law : $\varepsilon^c = A\,\sigma^n\,t^m$ $A = 3.125 \times 10^{-14}$ (t in hour, σ in N/mm^2) n = 5 m = 0.5
Boundary Conditions	$u_x = 0$ on line AD $u_y = 0$ on mid-point of AD
Loading	Prescribed tensile stress σ_1 on line BC

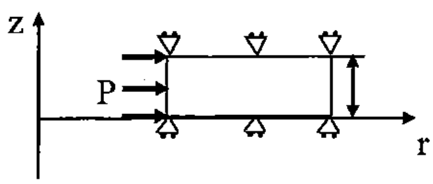


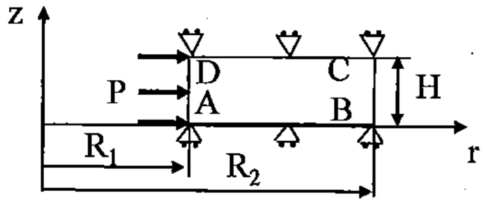
Reference solutions for the variable load creep problem

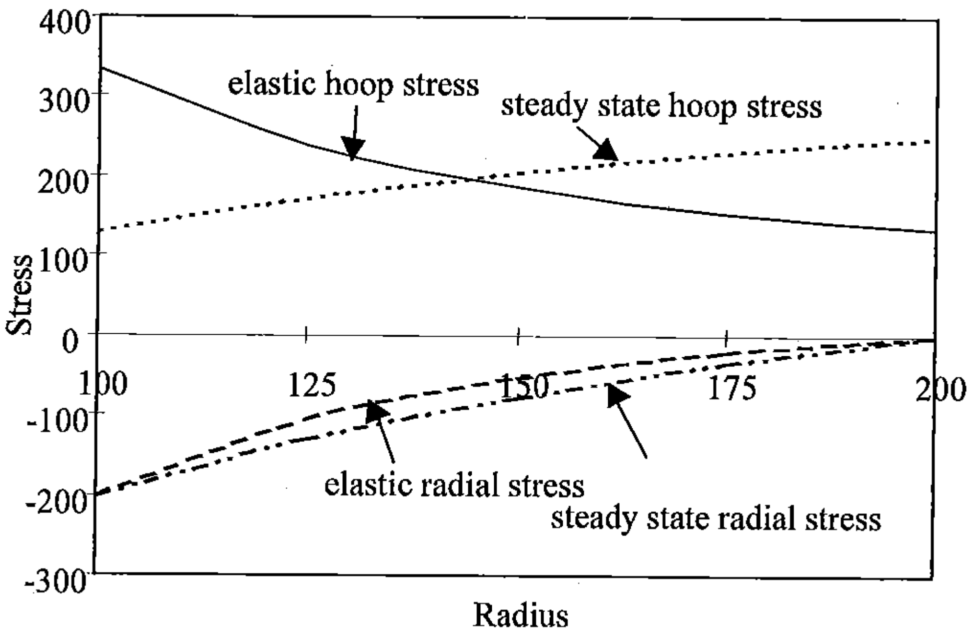
Example 4: Multi-Axial Pressurised Cylinder Creep

Physical Attributes

- Axisymmetric stress situation
- Redistribution of stress with time
- Steady state of stress



Geometry	Axisymmetric $R_1 = 100$ mm $R_2 = 200$ mm $H = 25$ mm $P = 200$ N/mm^2 Total creep time = 100 hours 
Material Properties	$E = 200 \times 10^3$ N/mm^2 $\nu = 0.3$ Creep Law : $\varepsilon^c = A\, \sigma^n\, t^m$ $A = 3.125 \times 10^{-14}$ (t in hour, σ in N/mm^2) $n = 5$ $m = 1$
Boundary Conditions	$u_z = 0$ at line AB $u_z = 0$ at line CD
Loading	Prescribed pressure on the inner surface (line AD)



Reference solutions for the pressurised cylinder creep problem

3.9 Summary of Key Points

- Creep is a time dependent material behaviour, usually described by a stress-strain power law obtained from **uniaxial creep tests** at high temperatures.

- Two approximations can be used for the creep strain rate in variable stress situations; **time hardening** which depends on the current time and **strain hardening** which depends on the accumulated creep strain.

- **Stress relaxation** usually occurs in creep situations where the structure is held at a constant strain.

- **Multi-axial creep formulations** are similar to plasticity formulations, and are based on the von Mises stress criterion and the Prandtl-Reuss flow rule.

(Summary/ Continued)

- **Explicit time-marching schemes** require small time increments to ensure convergence and stability, whereas implicit schemes are unconditionally stable. **Implicit schemes** consume more computational time in each time step.

- **FE codes** usually require a uniaxial creep law, and use automatic procedures for time marching. Complex non-standard creep laws usually require a user-defined subroutine.

- FE codes use both time hardening and strain hardening assumptions. In **variable load problems**, the computed creep strain may be affected by the hardening assumption used.

Part 4

Geometric Non-Linearity

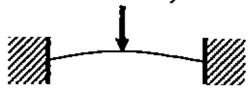
Lecture Outline

4.1 Features of Geometric Non-linearity Problems

Geometric Non-linearity problems

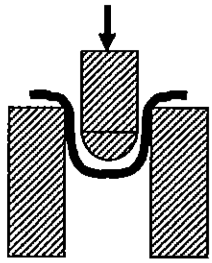
- **Small strain**
  - Example : Shallow arch (Small rotation)
  - Example : Buckling (Large rotation)
- **Large strain**
  - Example : Metal forming

Example : Shallow arch (Small rotation)



Example : Buckling (Large rotation)

Example : Metal forming



Classification of GNL Problems

(i) Small strain GNL problems

These problems are associated with small or large rotations.

- Examples of **small rotation problems** include:
  - shallow struts
  - shells
  - arches deflected by a transverse load
  - shallow spherical caps

- Examples of **large rotation problems** include:
  - fishing rod bent under the weight of a heavy fish
  - stationary helicopter blade hanging by its own weight
  - buckling of an imperfect Euler strut

(Classification of GNL Problems/Continued)

(ii) Large strain GNL problems

- Examples:
  Metal forming and manufacturing processes, such as deep drawing of drink cans, forging, extrusion and rolling.

- With large strains, it is also important to model material non-linearity such as plasticity.

- An exception is **rubber** which can undergo very large strains, of the order of unity, but remains elastic. This type of behaviour is called **'hyperelastic' or 'non-linear elastic' behaviour**. The constitutive equations for rubber can be derived from the expressions for the potential energy density.

Stress and Strain Measures in GNL problems

(i) Strain

- 'Engineering strain' may not be adequate when dealing with GNL problems because it measures the change in length over the original (undeformed) length.

- More suitable strain definitions which take into account the new length are needed, such as:
  'Logarithmic strain'.
  'Green strain'
  'Almansi strain'

(ii) Stress

- 'Engineering' or 'nominal' stress, defined as the force divided by the original undeformed area, may be inappropriate for use in GNL problems in which the cross-sectional area may exhibit large changes.

- Instead, as in material non-linearity, a 'true stress' (also called 'Cauchy stress') can be defined as the force divided by the 'current' cross-sectional area, rather than the original area.

(iii) Stress-strain relationships

- In conventional elasticity equations, stresses are linked to strains through the constitutive law, i.e. Hooke's law.

- In GNL problems, in addition to the constitutive equations, stresses are usually associated with the corresponding strains using the 'virtual work theorem' or 'total potential energy theorem'.

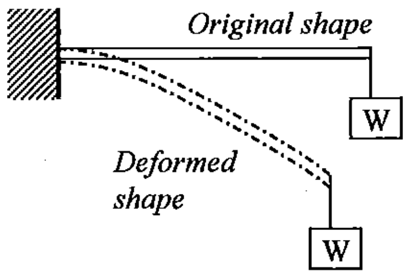
- This is important in large strain problems, where such stresses are called 'work-conjugates' to the corresponding strains.

Conservative and Non-Conservative (Follower) Loads

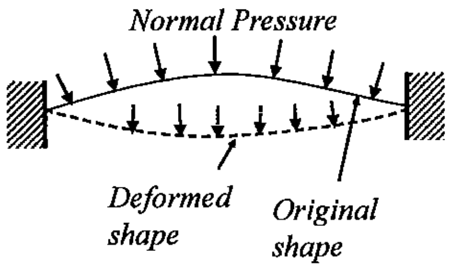
- A **conservative load** is that which always applies in a fixed direction regardless of the deformation of the body.
  Example: is a gravitational load, which always applies vertically.

- A **non-conservative (follower) load** is that which changes its direction during the deformation, i.e. it follows the deformation of the body.

  Example: An internal pressure on a cylinder which is allowed to change its position and direction as the cylinder wall deforms.



(a) Conservative load on a cantilever beam

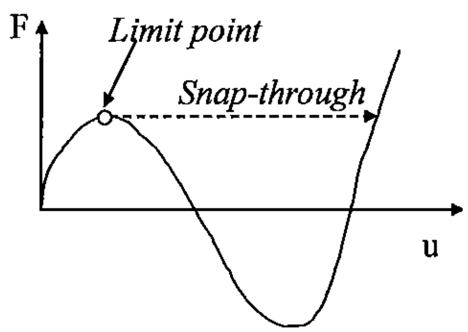


(b) Non- conservative (follower) load on a shell

Snap-Through Problems

- Snap-through behaviour usually occurs when there is a horizontal tangent in the load-displacement curve at the so-called **'limit points'** where the deformed shape suddenly jumps from one position to another, but the load remains constant.

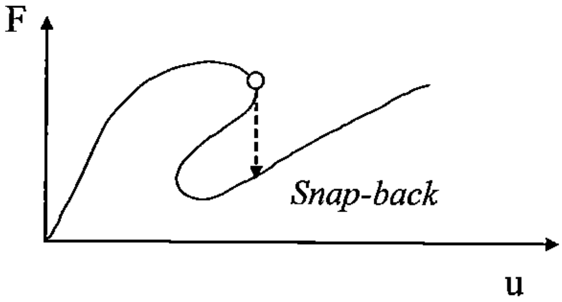
- In such problems, **'displacement control'** may be more appropriate to use than 'load control', i.e. the displacements are prescribed as small increments.



Snap-through behaviour

Snap-Back Problems

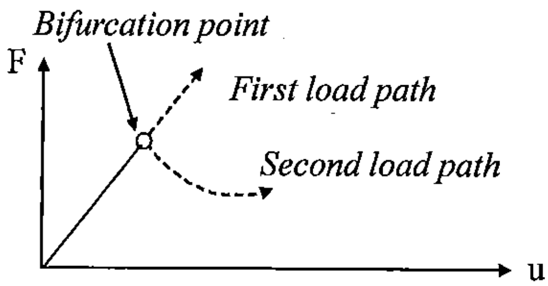
- Snap-back usually occurs when there is a vertical tangent in the load-displacement curve where the load on the body suddenly drops, but the displacement remains constant.
- In such problems, **'load control'**, i.e. the external forces applied as small increments, may be more appropriate than 'displacement-control'.



Snap-back behaviour

Bifurcation and Buckling Problems

- Bifurcation points can occur on the load-displacement curve as one or more solution paths that satisfy equilibrium intersect, i.e. the solution can proceed along more than one path. Only one path is **'stable'**, while all others are **'unstable'**.
- The deformed shape may be influenced by any initial imperfections or eccentricity of the applied loads.
- **Buckling analysis** is a special case of GNL analysis that deals with equations of the type classified in mathematics as **'eigen-value'** problems.



Bifurcation in the load path

GNL Lagrangian Formulations

Lagrangian formulation is the name given to a **'material' description** of the behaviour of a body. It refers to a material particle or fibre and tracks its movement or deformation within the body.

There are two types of Lagrangian formulations widely used in GNL problems:

(i) **Total Lagrangian (TL) formulation** where the equilibrium relationships and integrals are calculated with respect to the initial (undeformed) configuration of the structure, i.e. the displacement of the body is tracked from time **t = 0 to the current time (t + Δt).**

(ii) **Updated Lagrangian (UL) formulation** where the latest (updated) configuration is considered as the initial one, i.e. the displacement of the body is tracked from **time = t to the current time (t + Δt).**

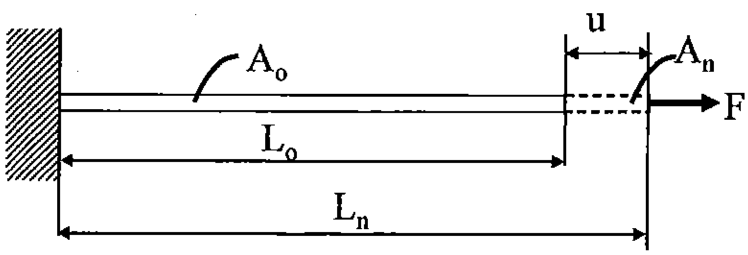
4.2 Uniaxial GNL Strains and Stresses

The principle of virtual work

The work done by the applied forces in deforming the body must be equal to the strain energy stored within the body, i.e. equal to the virtual work of internal forces and stresses.

For a simple one degree of freedom system, the virtual work theorem can be written as follows: $F\,\delta u = \int_V \sigma\,\delta\varepsilon\,dV$

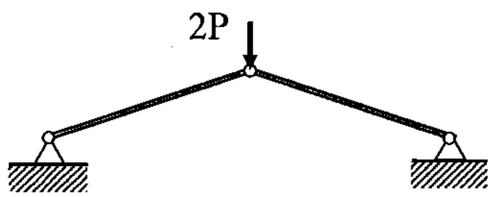
where
F is the applied load,
δu is the virtual displacement
$\delta\varepsilon$ is the virtual strain
V is the original volume.



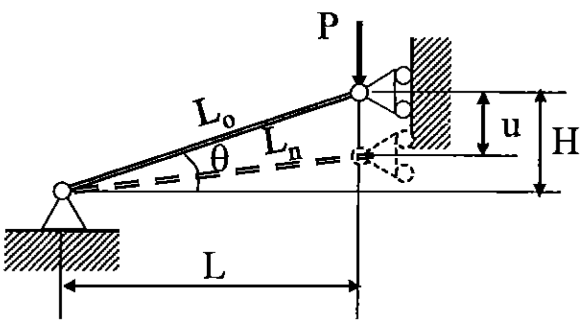
Comparison of Different Strain/Stress Measures in GNL Theory

Strain measure	***Datum used***	***Its work-conjugate stress***	***Relationship to Engineering strain***	***Relationship to Engineering stress***
Engineering strain	L_o	Engineering (nominal) stress, σ_o	$= \varepsilon_o$	$= \sigma_o$
Logarithmic (true) strain	L_n	Logarithmic (Cauchy) stress, σ_L	$\varepsilon_L = log_e(1+\varepsilon_o)$	$\sigma_L = \sigma_o(1+\varepsilon_o)^{2\nu}$
Green strain	L_o	Green (Second Piola-Kirchoff) stress, σ_G	$\varepsilon_G = \varepsilon_o + \frac{1}{2}\varepsilon_o^2$	$\sigma_G = \frac{\sigma_o}{(1+\varepsilon_o)}$
Almansi strain	L_n	Almansi stress, σ_A	$\varepsilon_A = \frac{2\,\varepsilon_o + \varepsilon_o^2}{2(1+\varepsilon_o)^2}$	$\sigma_A = \sigma_o(1+\varepsilon_o)^3$

4.3 A Simple Two-bar GNL Example



(a) Simple two-bar structure

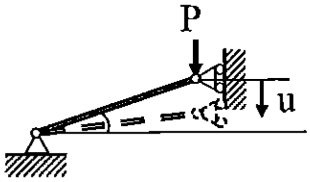


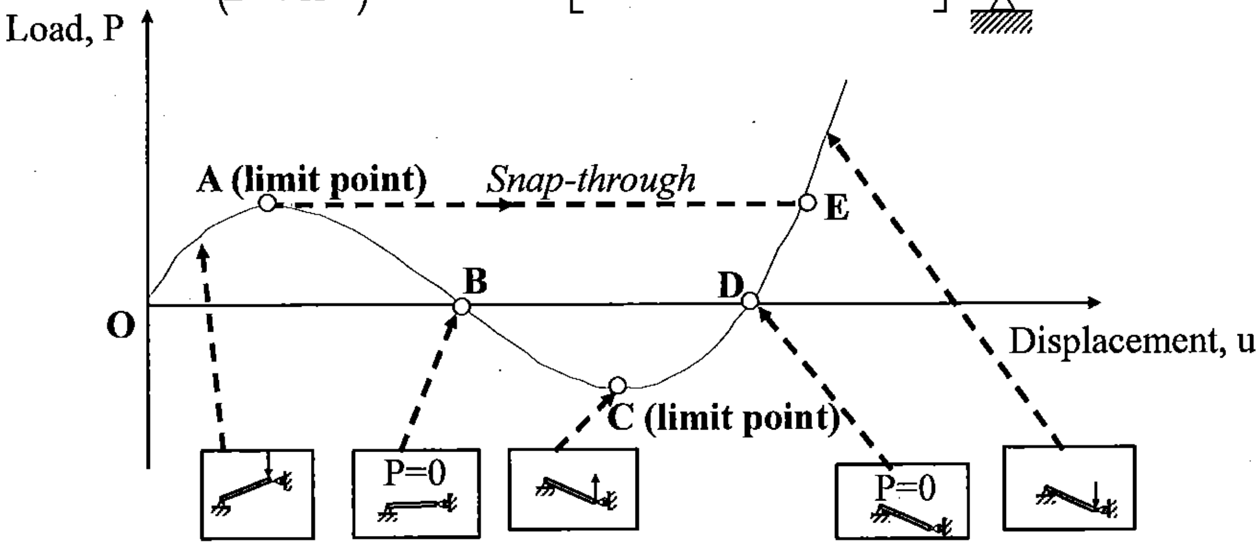
(b) Symmetrical half of the two-bar structure

Analytical Solution

The analytical solution for this problem can be obtained by minimising the total potential energy (TPE) of the system.

$$P = \frac{EA}{\left(L^2+H^2\right)^{1/2}}(H+u)\left[1-\left(\frac{L^2+H^2}{L^2+(H+v)^2}\right)^{1/2}\right]$$





Tangential Stiffness Matrix

(i) Updated Lagrangian Formulation

$$\frac{dP}{du} = K_t = \frac{EA_o(u+H)^2}{L_o{}^3} + \frac{T}{L_o}$$

(ii) Total Lagrangian Formulation

Expanding the above expression results in the following expression

$$K_t = \boxed{\frac{EA_oH^2}{L_o{}^3}} + \boxed{\frac{EA_o\left(u^2+2uH\right)}{L_o{}^3}} + \boxed{\frac{T}{L_o}}$$

$$K_t = \boxed{K_{Linear}} + \boxed{K_{Nonlinear}} + \boxed{K_{Initialstress}}$$

By using the 'updated Lagrangian formulation', the tangent stiffness expressed in terms of the 'current' updated dimension ($u + H$). Therefore, it is possible to eliminate the $K_{Nonlinear}$ term.

4.4 GNL Strains and Stresses in a Continuum

Material and Spatial Formulations

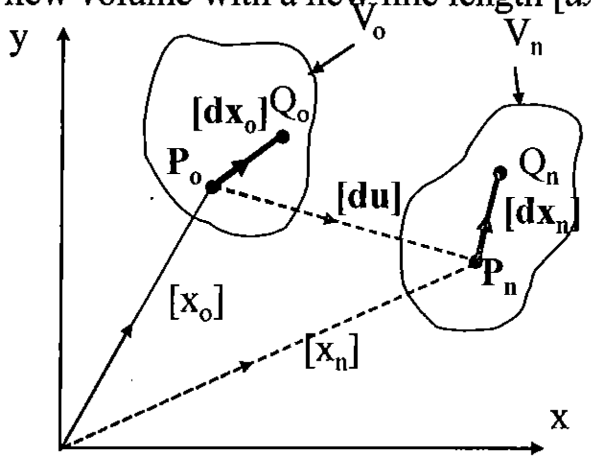
(i) **Lagrangian or 'material' formulation**
Describes the behaviour of a material particle and tracks its behaviour as the body deforms.

(ii) **Eulerian or 'spatial' formulation**
Describes the behaviour of a body referring to a particular 'spatial' position and follows the behaviour of the particles that occupy this location as time progresses.
Eulerian formulation is widely used in fluid mechanics where the motion of material though a volume or fixed space is studied.

Note: One can imagine the Lagrangian formulation as that which tracks the flight of a specific aeroplane (the particle) in space, whereas the Eulerian formulation monitors all aeroplanes that pass over a specific city (the spatial position).

Deformation of a continuum

- P_o with original position $[x_o]$ moves to a new position P_n at $[x_n]$, using the same origin O as datum.
- The displacement vector of point P_o is $[du]$.
- V_o is the original volume containing a line of length $[dx_o]$
  V_n is the new volume with a new line length $[dx_n]$.



Deformation of a continuum

Multi-axial GNL Strain Measures

Green strain: $$[\varepsilon_G] = [\varepsilon_{Linear}] + [\varepsilon_{Nonlinear}]$$

The 3D direct (non-shear) strain component of Green strain can be expressed as follows:

$$(\varepsilon_G)_{11} = \boxed{\frac{\partial u_x}{\partial x}} + \boxed{\frac{1}{2}\left(\left(\frac{\partial u_x}{\partial x}\right)^2 + \left(\frac{\partial u_y}{\partial x}\right)^2 + \left(\frac{\partial u_z}{\partial x}\right)^2\right)}$$

$$(\varepsilon_G)_{22} = \boxed{\frac{\partial u_y}{\partial y}} + \boxed{\frac{1}{2}\left(\left(\frac{\partial u_x}{\partial y}\right)^2 + \left(\frac{\partial u_y}{\partial y}\right)^2 + \left(\frac{\partial u_z}{\partial y}\right)^2\right)}$$

$$(\varepsilon_G)_{33} = \boxed{\frac{\partial u_z}{\partial z}} + \boxed{\frac{1}{2}\left(\left(\frac{\partial u_x}{\partial z}\right)^2 + \left(\frac{\partial u_y}{\partial z}\right)^2 + \left(\frac{\partial u_z}{\partial z}\right)^2\right)}$$

Linear part   Non-linear part

For the **shear components of Green strain**, the following expressions can be derived:

$$(\varepsilon_G)_{12} = \boxed{\frac{1}{2}\left(\frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x}\right)} + \boxed{\frac{1}{2}\left(\frac{\partial u_x}{\partial x}\right)\left(\frac{\partial u_x}{\partial y}\right) + \frac{1}{2}\left(\frac{\partial u_y}{\partial x}\right)\left(\frac{\partial u_y}{\partial y}\right) + \frac{1}{2}\left(\frac{\partial u_z}{\partial x}\right)\left(\frac{\partial u_z}{\partial y}\right)}$$

$$(\varepsilon_G)_{13} = \boxed{\frac{1}{2}\left(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x}\right)} + \boxed{\frac{1}{2}\left(\frac{\partial u_x}{\partial x}\right)\left(\frac{\partial u_x}{\partial z}\right) + \frac{1}{2}\left(\frac{\partial u_y}{\partial x}\right)\left(\frac{\partial u_y}{\partial z}\right) + \frac{1}{2}\left(\frac{\partial u_z}{\partial x}\right)\left(\frac{\partial u_z}{\partial z}\right)}$$

$$(\varepsilon_G)_{23} = \boxed{\frac{1}{2}\left(\frac{\partial u_y}{\partial z} + \frac{\partial u_z}{\partial y}\right)} + \boxed{\frac{1}{2}\left(\frac{\partial u_x}{\partial y}\right)\left(\frac{\partial u_x}{\partial z}\right) + \frac{1}{2}\left(\frac{\partial u_y}{\partial y}\right)\left(\frac{\partial u_y}{\partial z}\right) + \frac{1}{2}\left(\frac{\partial u_z}{\partial y}\right)\left(\frac{\partial u_z}{\partial z}\right)}$$

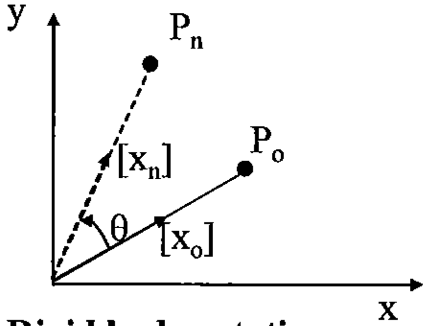
Linear part   Non-linear part

Useful features of Green Strain

- For small strains, the terms with multiples of strain become negligible, and Green strain **converges to the linear strain.**

- Green strain in a continuum is that it is **unaffected by a pure rotation**, i.e. a rigid body rotation with no 'stretching".

- Another useful feature of Green strain is that, provided strains are considered small, the stress associated with Green strain is the same as the Logarithmic (true) stress **measured along the local axes** that rotate with the material as it deforms.



Rigid body rotation

Multi-axial GNL Stress Definitions

Hooke's law:

$$\sigma_{ij} = 2\mu\varepsilon_{ij} + \sum_{k=1}^{3} \frac{2\mu\nu}{1-2\nu}\delta_{ij}\,\varepsilon_{kk}$$

where C_{ijkl} is a fourth-order tensor containing the material properties.

Therefore, the incremental **Green stress-strain** can be written as follows:

$$\sigma_{ij} = C_{ijkl}\,\varepsilon_{kl}$$

In most FE formulations, it is convenient to express the stress-strain relationship (the constitutive equation) as follows:

$$[\delta\sigma_G] = [C][\delta\varepsilon_G]$$

4.5 FE Treatment of Geometric Non-linearity

Total Lagrangian FE Formulation

An increment in Green strain can be written as follows:

$$[\varepsilon_G] = [B_L]\,[\delta u] + [B_{NL}]\,[\delta u]$$

In order to derive an expression for the tangential stiffness matrix, the following **stress-strain constitutive equation** can be used:

$$[\sigma] = [C]\,[\varepsilon]$$

Load and displacement control

- **'Load control'** does not work when a snap-through behaviour is encountered at a limit point, where the load-displacement tangent becomes horizontal, because the load must decrease in order to satisfy equilibrium.

- **'Displacement control'** cannot cope with snap-back behaviour, i.e. when the load-displacement tangent becomes vertical.

- The user of an FE code may be able to **switch from prescribed load increments to prescribed displacement increments** depending on the type of problem encountered.

- Alternatively, the user can **prescribe artificial springs** at a certain node to restrain the structure at or near limit points. The FE code may avoid the solution of the equilibrium equations around the limit point.

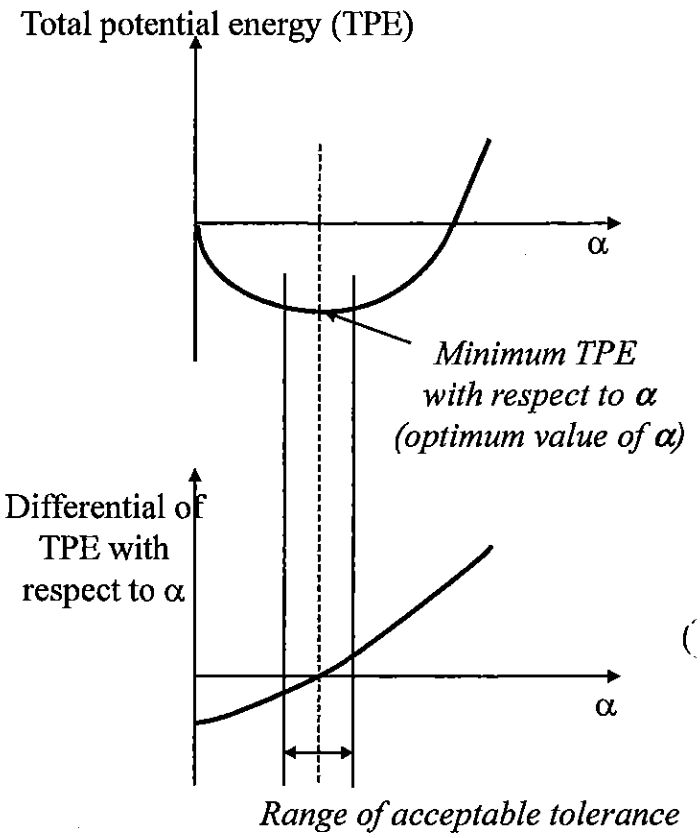
Line Search Technique

In the iteration process, a correction for the displacement vector is obtained from the **residual (out of balance) forces** and added to the old displacements. Instead of simply adding the displacement correction, it can be scaled by **a scalar factor, α**, as follows:

$$[u_{new}] = [u_{old}] + \alpha\,[\delta u_{correction}]$$

The above equation represents a line along which a 'search' is made to obtain the optimum value of α, i.e. the **optimum iterative step length**.

$$\frac{\partial(TPE)}{\partial\alpha} = 0$$



Arc Length Method

- Here, the solution is controlled by an **'arc-length increment'** rather than a load or displacement increment.

- At a given point on the load path, the unbalanced residual load vector can be expressed as follows:

$$[R] = [F_{int}] - \beta\,[F_{ext}] = 0$$

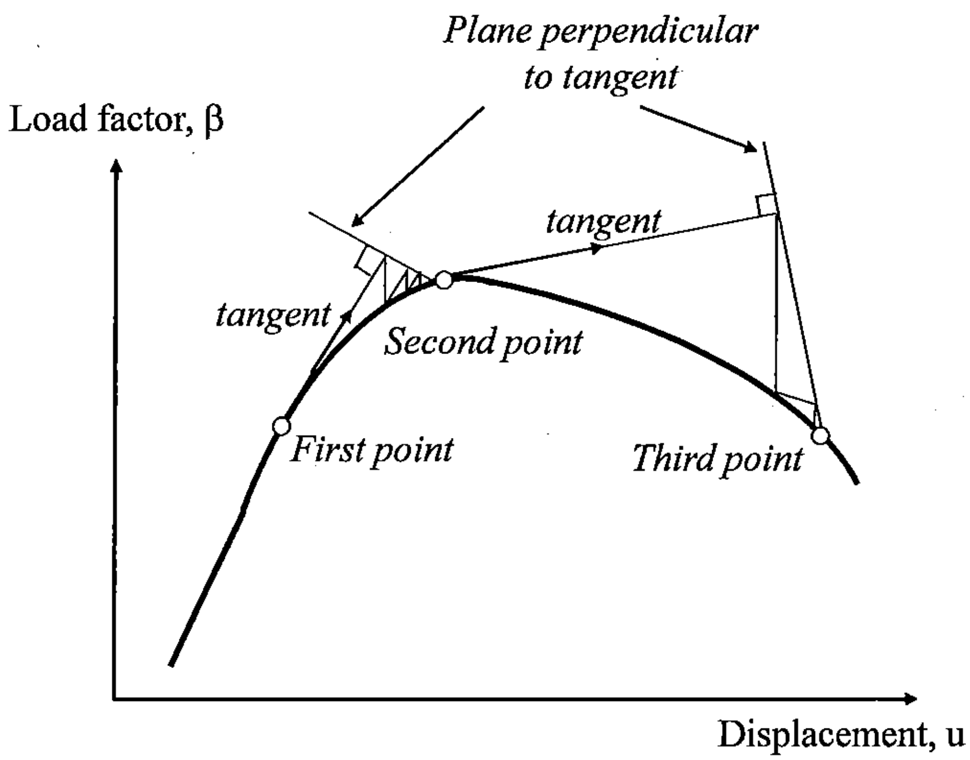
where β is the scalar factor which is independent of the displacements.

- The load multiplier β, which is a constant in the Newton-Raphson method, becomes an additional variable to be computed by the FE solver, i.e. **an extra equation is added** to the solver making it unsymmetric.

Advantages and Disadvantages of the Arc Length Method

- The **advantage** of the extra equation is that the solution matrix never becomes 'singular' even at the limit points. Therefore, the solution matrix is re-assembled with N+1 variables, where N is the total number of the variables (degrees of freedom) of the system.

- The **disadvantage** is that, in some FE formulations, the solution matrix becomes unsymmetric, which may incur an increase in computing time and/or computer storage, particularly for very large problems.



Schematic representation of the arc-length technique

Buckling and Instability Analysis

The buckling problem can be mathematically simplified to solving the FE equations where a small displacement [δu] occurs without any increment of load applied to the structure, i.e. the solution of the following equation:

$$\left(\left[K_{Linear}\right] + \lambda \left[K_{Initial\ stress}\right]\right) \left[\delta u\right] = 0$$

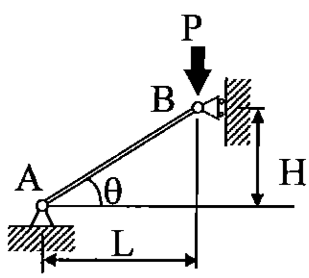
The above equation is the classical **'eigen-value' problem** where:

- λ is the **'eigen-value'** , i.e. the multiple of the load
- [δu] is the **'eigen-vector'**, i.e. the displacement vector describing the bucking mode.

The solution of eigen-value problems is well established in engineering. In non-linear buckling analysis, an eigen-value analysis is used to obtain the buckling modes.

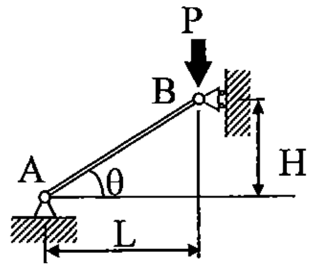
4.6 GNL Examples

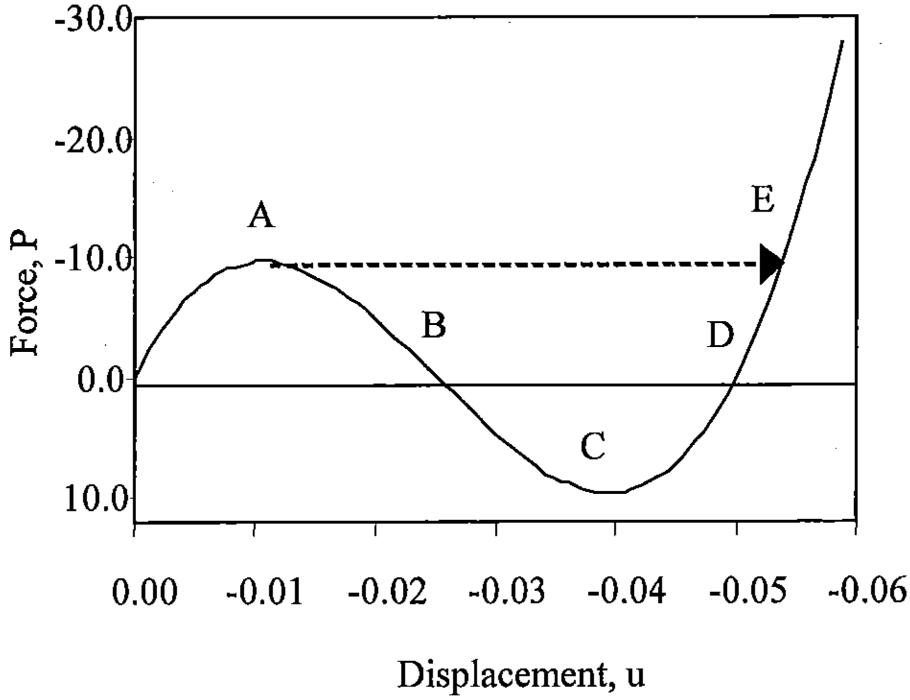
GNL Example 1: Snap-Through Behaviour



Physical Attributes:

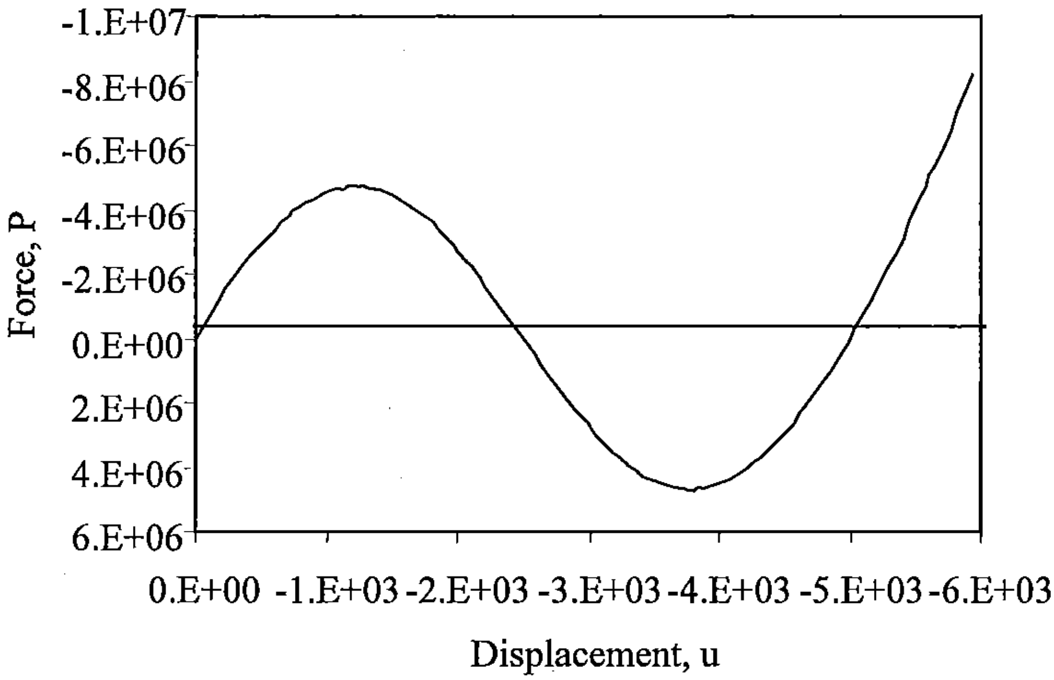
- Limit points in the load path
- Snap-through behaviour
- Displacement control
- Load control with arc-length

Geometry	2D strut L = 2500 mm Cross-sectional area = 100 mm^2 **Case a : Shallow strut**, H = 25 mm **Case b : Deep strut**, H = 2500 mm 
Material Properties	Linear elastic with geometric non-linearity E = 500 x 10^3 N/mm^2 ν= 0.5
Boundary Conditions	u_x = 0 and u_y = 0 at point A (pin-joint to a rigid surface) u_x = 0 at point B (vertical slider)
Loading	P = 11.4 N (Case a - Shallow strut) P = 6 x 10^6 N (Case b - Deep strut)



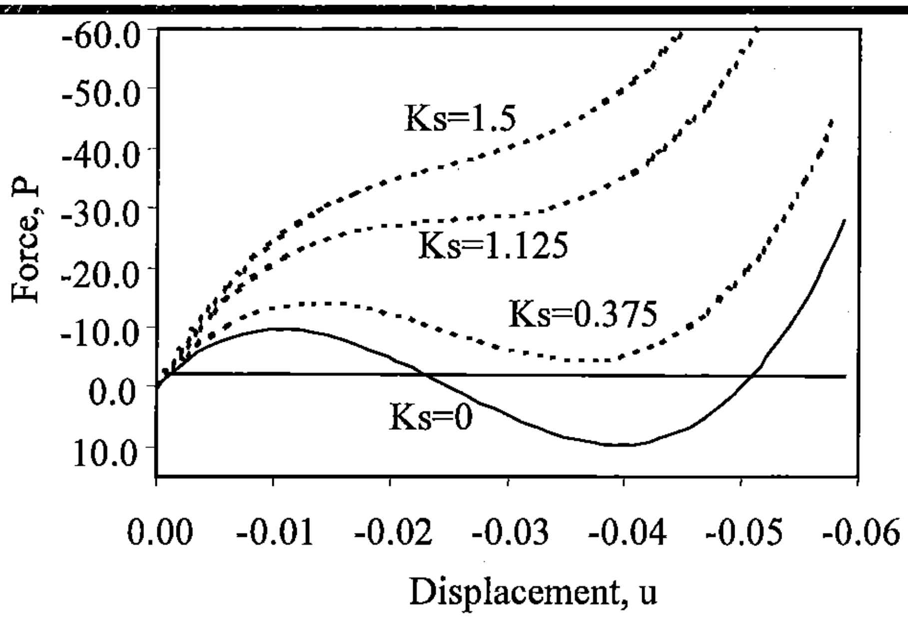
(a) Shallow truss

Reference solutions for the snap-through problem



(b) Deep truss

Reference solutions for the snap-through problem



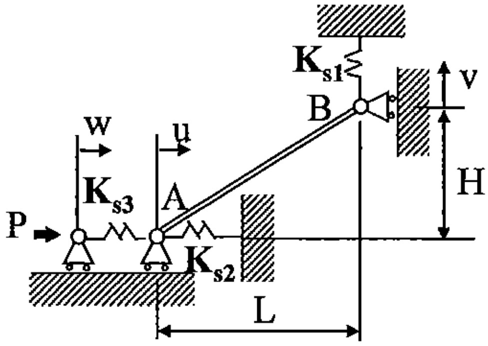
(c) Shallow truss with a vertical spring K_s at the loaded end

Reference solutions for the snap-through problem

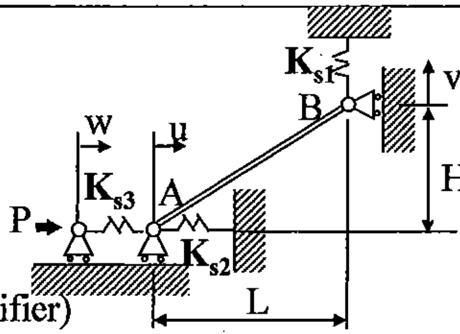
GNL Example 2: Snap-Back bahaviour

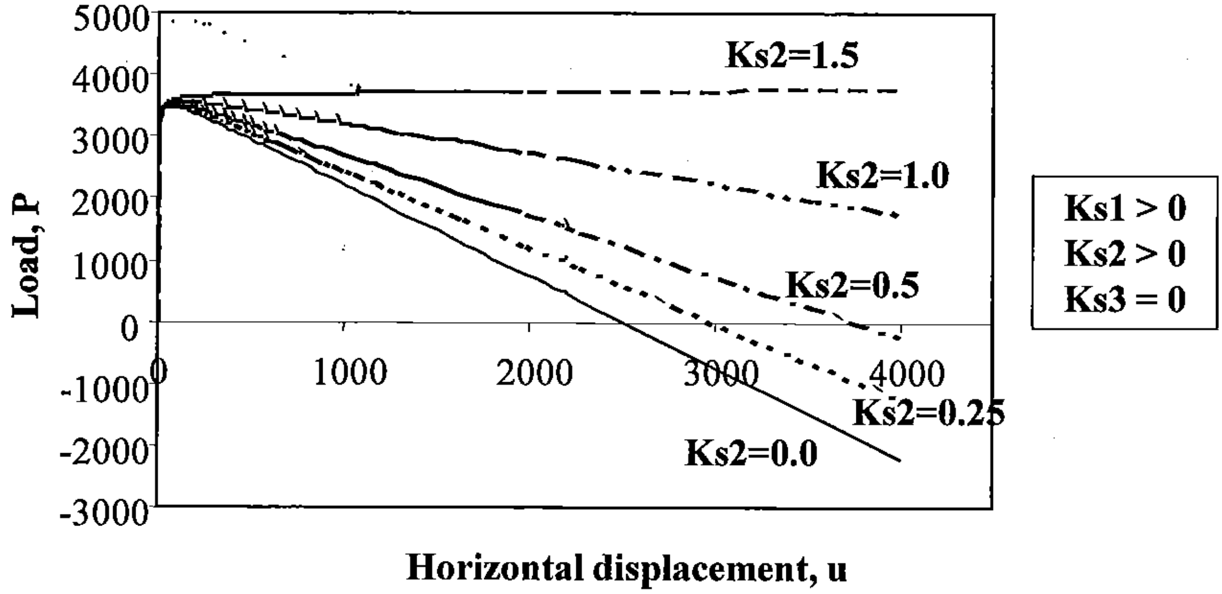
Physical Attributes

- Limit points in the load path
- Load modifier
- Snap back behaviour



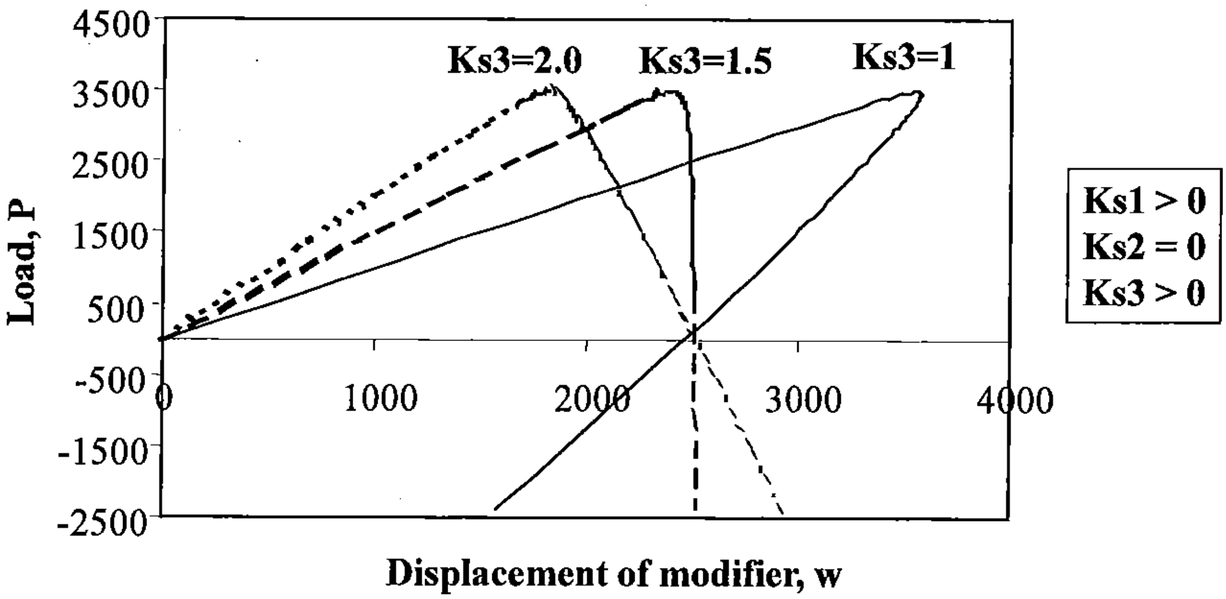
(Snap-back Behaviour/ Continued)

Geometry	2D bar $L_1 = 2500$ mm $L_2 = 25$ mm $A = 100$ mm^2 $K_{s1} = 1.5$ N/mm $K_{s2} = 0.25$ N/mm Case a : $K_{s3} = 0$ (no load modifier) Case b : $K_{s3} = 1.0$ N/mm 
Material Properties	Linear elastic with geometric non-linearity $E = 500 \times 10^3$ N/mm^2, $v = 0.0$
Boundary Conditions	$u_y = 0$ at point A (horizontal slider) $u_x = 0$ at point B (vertical slider) Vertical spring of stiffness K_{s1} attached to point B Horizontal spring of stiffness K_{s2} attached to point A
Loading	Horizontal point force P applied to a point A through a horizontal spring element of stiffness K_{s3} $P = 4.2 \times 10^3$ N , applied in 6 increments



(a) No load modifier

Reference solutions for the snap-back problem



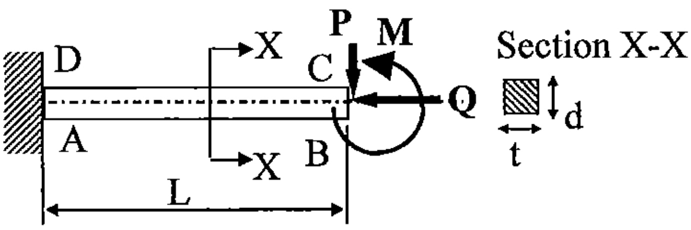
(b) With load modifier

Reference solutions for the snap-back problem

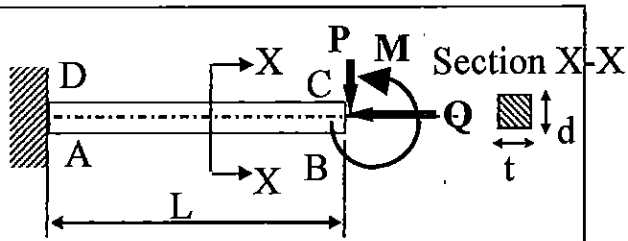
GNL Example 3: Straight Cantilever

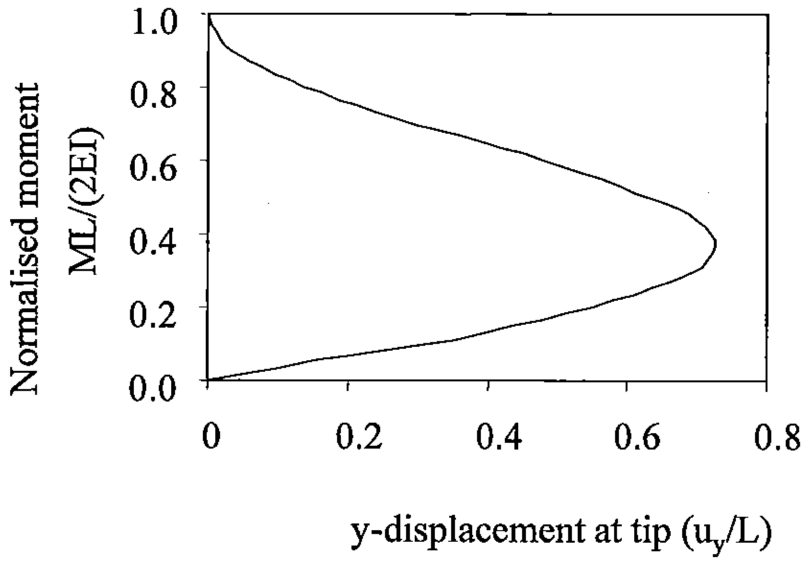
Physical Attributes

- Large displacements and large rotations
- Combined bending and membrane action
- Stiffening structure



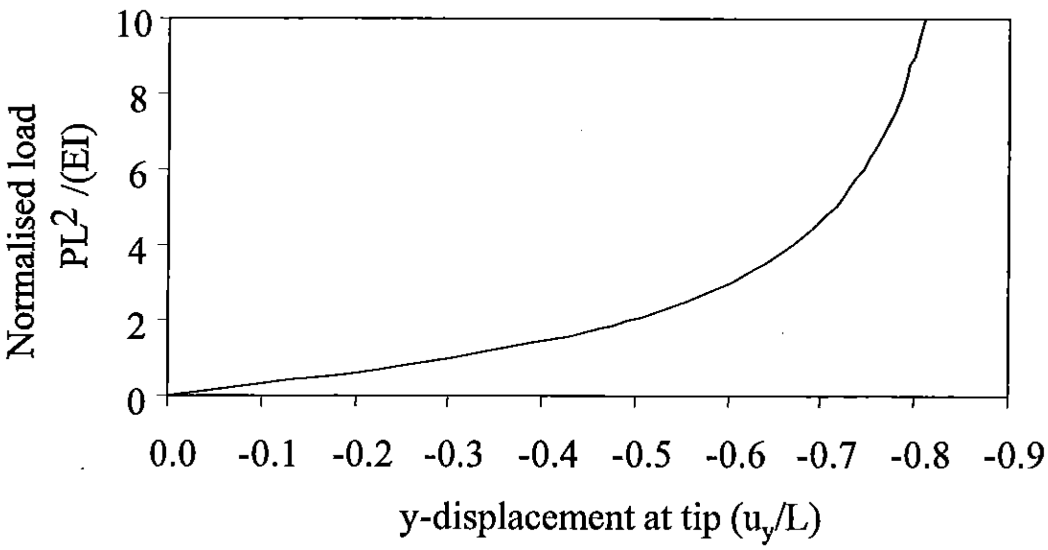
(Straight Cantilever Example/ Continued)

Geometry	2D beam/continuum L = 3.2 m d = 0.1 m t = 0.1 m  Case a : P = 0, Q = 0, M = 3.4361 x 10^6 Nm Case b : P = 1.709 x 10^6 N, Q = 0, M = 0 Case c : P = 3.844 x 10^3 N, Q = 3.844 x 10^6 N, M = 0
Material Properties	Linear elastic with geometric non-linearity E = 210 x 10^9 N/m^2, v = 0.0
Boundary Conditions	Built-in conditions on line AD
Loading	P, Q and M are applied at mid-point of BC
FE Model	Element Type : 3-node thin-beam, 2-node thick-beam, or 8-node quadratic 2D plane stress continuum



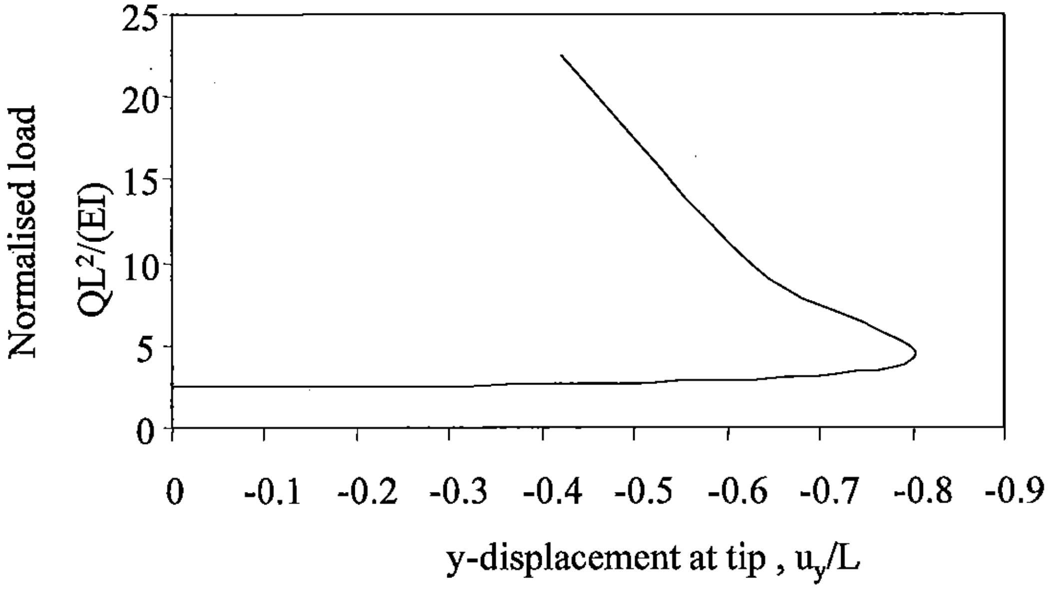
(a) Concentrated moment only
(P = 0, Q = 0, M = 3.4361 x 10^6 Nm)

Reference solutions for the straight cantilever problem



(b) Transverse Load P only
(P = 1.709 x 10^6 N, Q = 0, M = 0)

Reference solutions for the straight cantilever problem



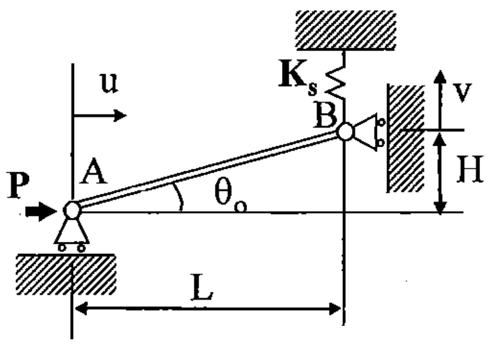
(c) With *axial and small transverse loads*
$P = 3.844 \times 10^3$ N, $Q = 3.844 \times 10^6$ N, $M = 0$

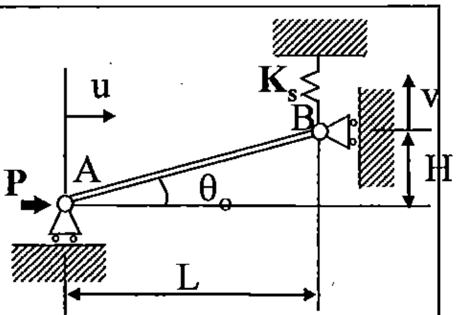
Reference solutions for the straight cantilever problem

GNL Example 4: Bifurcation and Buckling

Physical Attributes

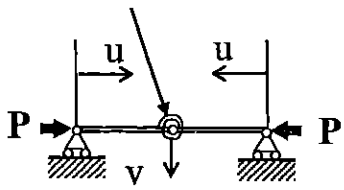
- Perfect buckling with bifurcation point
- Imperfect buckling with no bifurcation
- Stable and unstable load paths
- Non-uniqueness of the equilibrium solution
- Critical buckling load
- Possibility of singular stiffness matrix



Geometry	2D strut L = 2500 mm Cross-sectional area = 100 mm² K_s = 1.5 N/mm Case a : $\theta_o = 0$ (perfect buckling) H = 0 Case b : $\theta_o > 0$ (imperfect buckling), H = 25 mm 
Material Properties	Linear elastic with geometric non-linearity E = 500 x 10^3 N/mm² v = 0.0
Boundary Conditions	$u_y = 0$ at point A (horizontal slider) $u_x = 0$ at point B (vertical slider) Vertical spring of stiffness K_s attached to point B
Loading	Horizontal point force P applied at point A in 6 increments P = 3.9 x 10^3 N

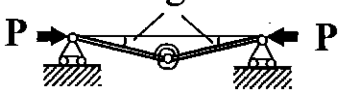
(a) Stable Bifurcation Example

Rotational spring

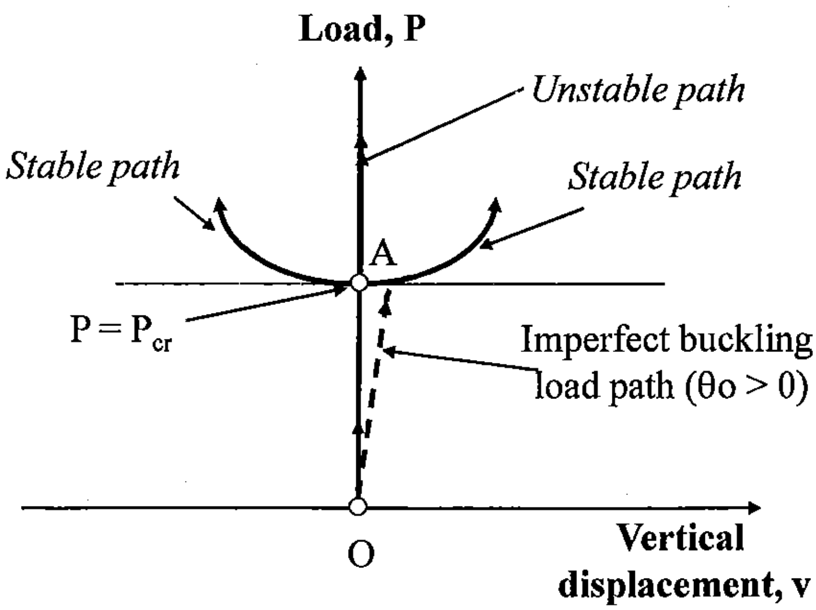


Perfect buckling

Angle θo

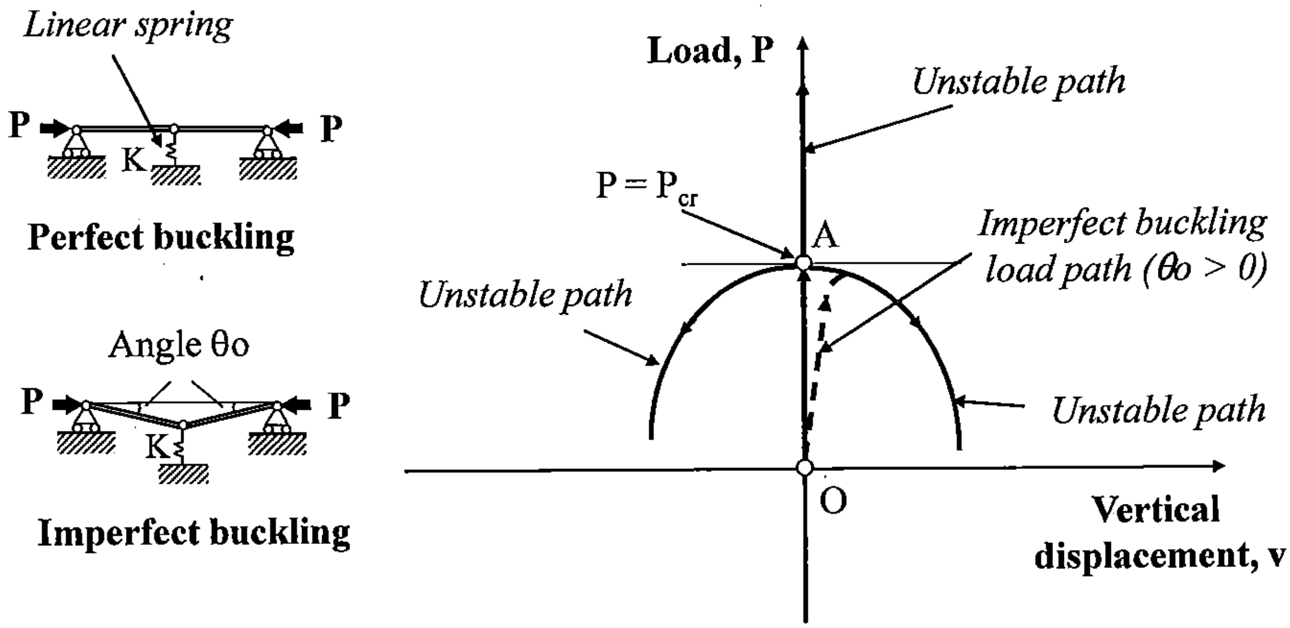


Imperfect buckling

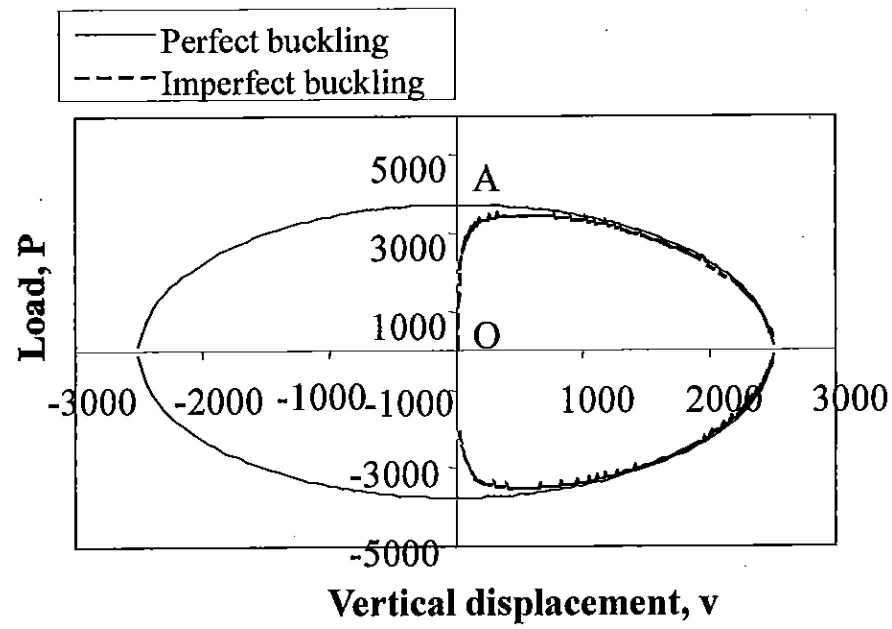


Load path in stable bifurcation (rotational spring)

(b) Unstable Bifurcation Example



Load path in unstable bifurcation (linear spring)



Reference solutions for the bifurcation problem

4.7 Summary of Key Points

- In GNL problems, the **updated geometry** (i.e. the deformed shape) is taken into account in the analysis.
- GNL problems are not easy to identify since they can involve **small or large** displacements, rotations and strains.
- The **stress and strain definitions** in GNL problems have to be carefully established since the conventional linear stress and strain definitions are inadequate when dealing with large displacements and large strains.
- The **direction of the applied loads** must be defined in GNL problems as either conservative or non-conservative (follower).
- GNL problems may involve **limit points** in the load-displacement curves, and may lead to snap-through (load peak) or snap-back (displacement peak).

(Summary/ Continued)

- **Bifurcation and buckling** problems are GNL problems in which the solution path may follow one or more solutions paths, some of which may be unstable load paths.
- **Displacement control** can be used for snap-through GNL problems, while load-control can be used for snap-back GNL problems.
- Special numerical procedures such as **arc-length** and **line search methods** can be used to overcome difficulties encountered around limit points in the load-displacement curve.
- **Buckling problems** usually require an eigen-value analysis in which the eigen-value is the multiple of the load, and the eigen-vector is the displacement vector describing the bucking mode.

Part 5

Contact Problems
(Boundary Non-Linearity)

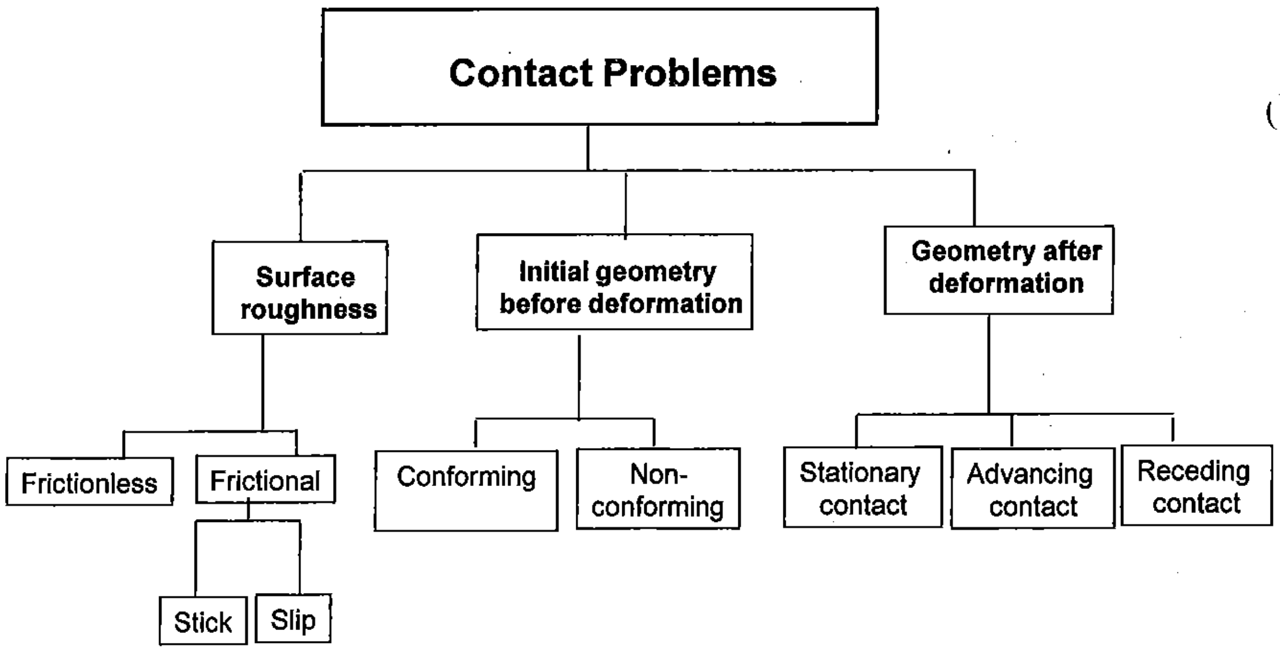
Lecture Outline

5.1 Classification of Contact Problems

- Contact problems are problems in which two or more bodies touch each other without being attached together and forces are transmitted between them.

- The analysis of contact problems is a major concern in many engineering applications such as:
  ball bearings, gears, rollers, mechanical seals and pressure vessel attachments.

- Contact problems have been traditionally classified according to:
  - Surface roughness
  - Initial geometry before contact
  - Behaviour after contact



Classification of contact problems

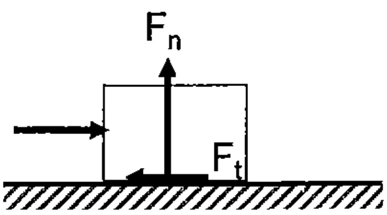
(a) Classification according to Surface Roughness

(i) Frictionless contact

- The contact surfaces are assumed to be perfectly lubricated, i.e. slipping in the tangential direction is not restricted.
- Only compressive normal forces can be transmitted across the contacting bodies.
- The surfaces are free to move in the tangential direction, but the continuity of displacement must be maintained in the normal direction.
- To satisfy equilibrium, the normal forces/tractions must be equal and opposite, but the tangential forces are zero.

(ii) Frictional contact

Coulomb Law $F_t \leq \mu F_n$



μ is the Coulomb coefficient of friction

Frictional stick $F_t < \mu F_n$

Frictional slip $F_t = \mu F_n$

(F_t cannot exceed the value of μF_n)
The direction of the relative slip must be such that it opposes motion.

Stick-Slip
In practical contact problems, contact surfaces may contain a mixture of frictionless and frictional interfaces, and may display a mixture of stick and slip zones in the frictional interfaces.

(b) Classification according to Initial Geometry (Before Deformation)

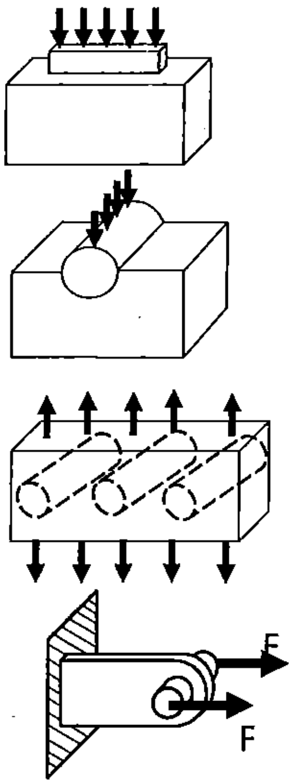
(i) Conforming contact

- The surfaces of the contacting bodies fit closely together prior to deformation.
- Examples: contact around a fibre in a matrix, or a loaded pin.

(ii) Non-conforming contact

- The contact surfaces have different surface profiles,
- Examples: Contact of spheres or cylindrical rollers.
- Contact may initially be at a point, e.g. in ball bearings, or along a line, e.g. in roller bearings.
- The contact area in non-conforming bodies is usually small compared to the overall dimensions of the bodies.

Examples of Conforming Contact Problems



Two flat plates

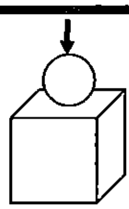
Cylinder in a cylindrical groove

Cylindrical fibers embedded in a homogeneous matrix

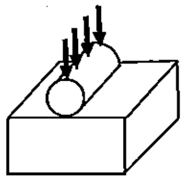
Loaded pin in a hole

Examples of conforming contact problems

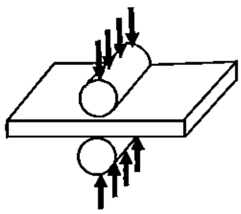
Examples of Non-Conforming Contact Problems



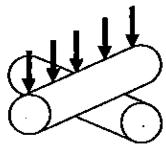
Sphere on a plate



Cylindrical roller on a plate



Two rollers and a flat plate



Perpendicular cylinders

Examples of non-conforming contact problems

(c) Classification according to Geometry After Contact

(i) Stationary Contact

- The size of the contact area does not change upon loading
- Note that if the direction of the applied load changes, then the contact area may change.
- Example: Flat punch on a foundation

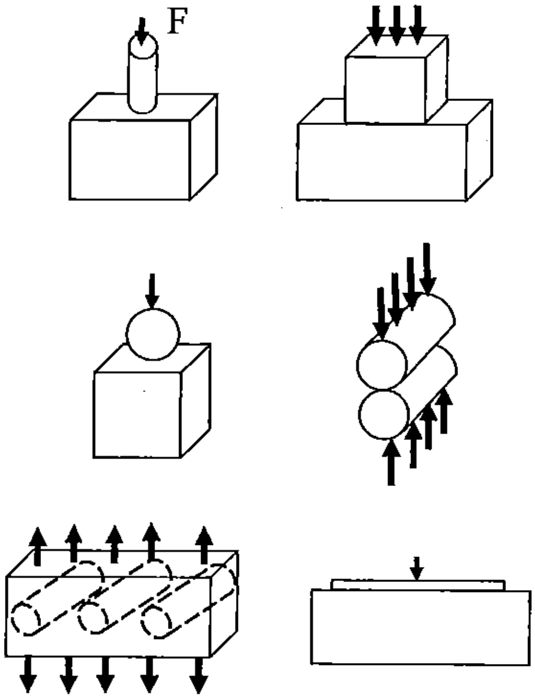
(ii) Advancing contact

- The contact area increases from its initial state as the external loads are applied.
- Examples: contact of spheres or cylindrical rollers

(iii) Receding contact

- The initial contact area becomes smaller as the external loads are applied.
- The size of the contact area does not change as the load is increased (provided that the direction of loading remains the same).
- Example: inclusion embedded in a large matrix

(Classification according to Geometry After Contact/ Continued)



(a) Stationary contact
(No change in contact area)

(b) Advancing contact
(Contact area increases)

(c) Receding contact
(Contact area decreases)

Examples of stationary, advancing and receding contact problems

5.2 Hertzian Contact Problems

- Based on the contact theories first established by Hertz [1896]
- Hertzian solutions provide simple expressions for contact pressures and contact areas, widely used in design.
- Hertzian solutions are very useful for predicting the patterns of contact stresses (and checking FE contact solutions)
- Although limited to frictionless contact problems, Hertzian contact expressions can be extended to approximate frictional contact problems

Heinrich Rudolf Hertz
(1857–1894)

Assumptions and Limitations of Hertz Theory

Hertzian contact problems must satisfy the following 5 conditions:

(i) The bodies are homogeneous, isotropic, satisfy Hooke's law and experience **'small' deformations, strains and rotations.**

(ii) Prior to deformation, the **contact surfaces are continuous and non-conforming**, and may be represented by second order polynomials (i.e. quadratic surfaces).

(iii) The contact surfaces are **frictionless.**

(iv) Each contacting body may be treated as **a half-space**, i.e. the contact areas are very small compared to the overall dimensions.

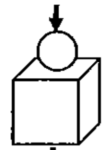
(v) The dimensions of the **deformed contact area remain small** compared to the principal radii of the undeformed surfaces.

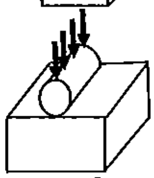
Comparison of Hertzian and Non-Hertzian Contact Problems

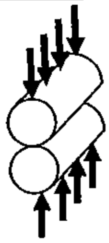
Contact Problems

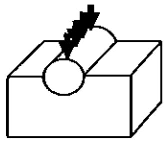
	Hertzian	**Non Hertzian**
Material behaviour	Linear elastic	Elastic or non-linear material behaviour
Displacement	Small displacement	Small or large displacement
Friction	Frictionless	Friction may be present
Surfaces in contact	Curved contact surfaces No sharp edges or corners	May be curved or flat May contain sharp edges or corners
Contact area as load is applied	Advancing contact	Stationary or receding contact
Geometry before contact	Non-conforming contact	Usually non-conforming contact

Examples of Hertzian (and Non-Hertzian) Contact Problems

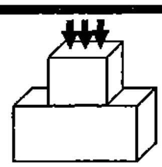

Sphere on a plate

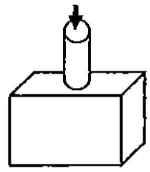

Cylinder on a plate

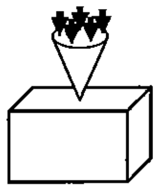

Two cylinders

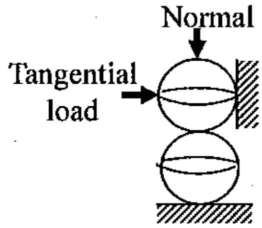

Cylinder in a socket

Examples of Hertzian contact problems


Flat punch on a foundation


Cylindrical punch on a foundation


Conical indenter



Tangential loading

Examples of Non-Hertzian contact problems

Derivation of the Hertz Contact Stresses for curved bodies

By examining the initial geometry (before deformation) in the vicinity of the contact area, it can be shown that the separation of the two bodies, measured along the z-axis, is given by:

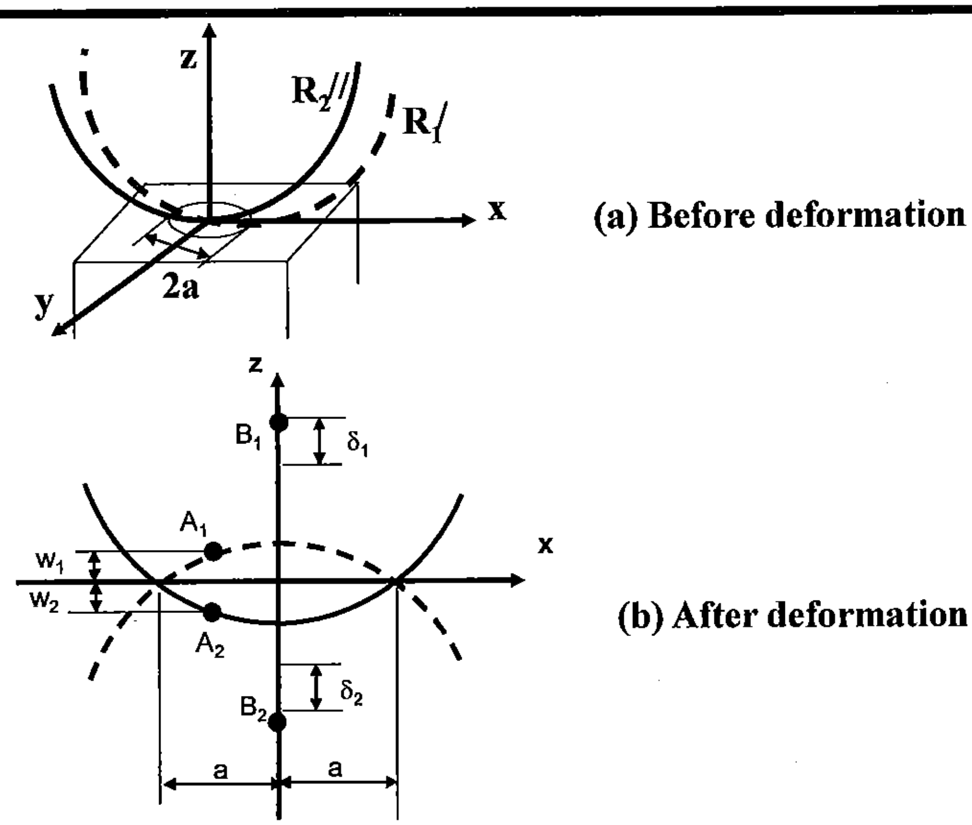
$$h(x,y) = \frac{x^2}{2R'} + \frac{y^2}{2R''}$$

where R' and R'' are the <u>relative</u> radii of curvature of the two bodies, defined as follows:

$$\frac{1}{R'} = \frac{1}{R_1'} + \frac{1}{R_2'}$$

$$\frac{1}{R''} = \frac{1}{R_1''} + \frac{1}{R_2''}$$

where R_1', R_1'' are the principal radii of curvature of body 1, while R_2', R_2'' are the principal radii of curvature of body 2.



Definition of the Hertz contact geometry before and after loading

After the application of the external loads, the deformation of the corresponding contact points A_1 and A_2 is given by:

$$w_1(x,y) + w_2(x,y) = \delta - h(x,y)$$

w is the displacement in the z-direction
δ is the approach of these two points in each body (i.e. the approach of two distant points B_1 and B_2), chosen as the datum for the elastic displacements, as follows:

$$\delta = \delta_1 + \delta_2 = w_1(0) + w_2(0)$$

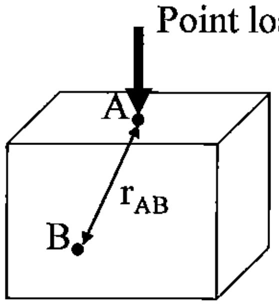
where δ_1 and δ_2 are the displacements in the z-direction of the two distant points, B_1 and B_2.

For a concentrated point load, F, applied on point A on the plane surface of a semi-infinite body, the displacement of any other point, B, can be derived as follows:

$$w(x, y) = \frac{1 - \nu^2}{\pi E} \frac{F}{r_{AB}}$$

where r_{AB} is the distance between points A and B.

This expression can be extended to cover loading situations in which the pressure is distributed over an area, S.



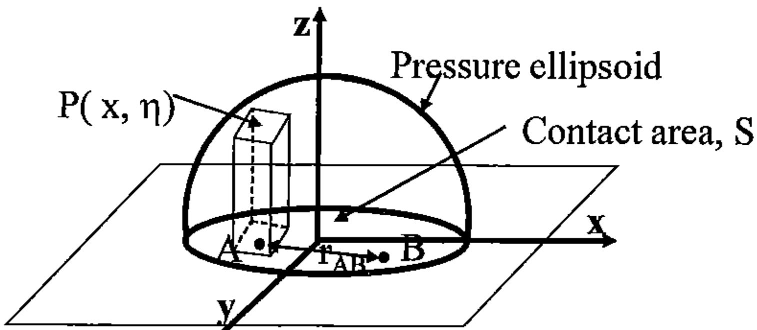
Point load on a surface

Considering a small area on which the distributed pressure *p(x,y)* applies, the force-displacement relationship can be written as follows:

$$\frac{1}{\pi E^*} \iint_S \frac{p(\xi, \eta)}{\left[(x-\xi)^2 + (y-\eta)^2\right]^{1/2}} \, d\xi \, d\eta = \delta - h(x \;\; y)$$

where E^* is the "effective" Young's modulus, defined as follows:

$$\frac{1}{E^*} = \frac{1-\nu_1{}^2}{E_1} + \frac{1-\nu_2{}^2}{E_2}$$



Distributed pressure on a half-space

Hertz Contact Solutions for Two Doubly-curved Bodies in Contact

Hertz worked out that an elliptical distribution of the pressure satisfies the above integral equation (based on his work on electrostatic potential problems). The Hertz solution is

$$p(x, y) = p_o \left(1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}\right)^{1/2}$$

where a and b are the semi-axes of the ellipse, and p_o is the maximum pressure, which occurs at the centre of the contact surface.

The total load, F, pressing the two bodies together, is equal to the volume of the semi-ellipsoid, i.e.

$$F = \frac{2}{3} \pi\, a\, b\, p_o$$

Hertz solutions for Axisymmetric Contact of Spheres

- The initial undeformed contact is at one point
- The contact area after the application of the external load is circular.
- Consider the contact between two spheres of different radii R_1 and R_2, pressed by a perpendicular point force F along the line joining their centres.
- Due to the spherical geometry, $R' = R'' = R$, and the contact area is a **circle**. Therefore, the contact pressure can be written as

$$p(r) = p_o \left(1 - \frac{r^2}{a^2}\right)^{1/2}$$

where r is the radius of any point within the contact area, i.e.

$$r = \sqrt{\left(x^2 + y^2\right)}$$

The Hertz solutions for spherical contact are:

$$a = \left(\frac{3}{4} \frac{F\,R}{E^*} \right)^{1/3}$$

$$p_o = \left(\frac{6}{\pi^3} \frac{F\,E^{*2}}{R^2} \right)^{1/3}$$

$$\delta = \left(\frac{9}{16} \frac{F^2}{R\,E^{*2}} \right)^{1/3}$$

where R is the relative radius of curvature, defined as follows:

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}$$

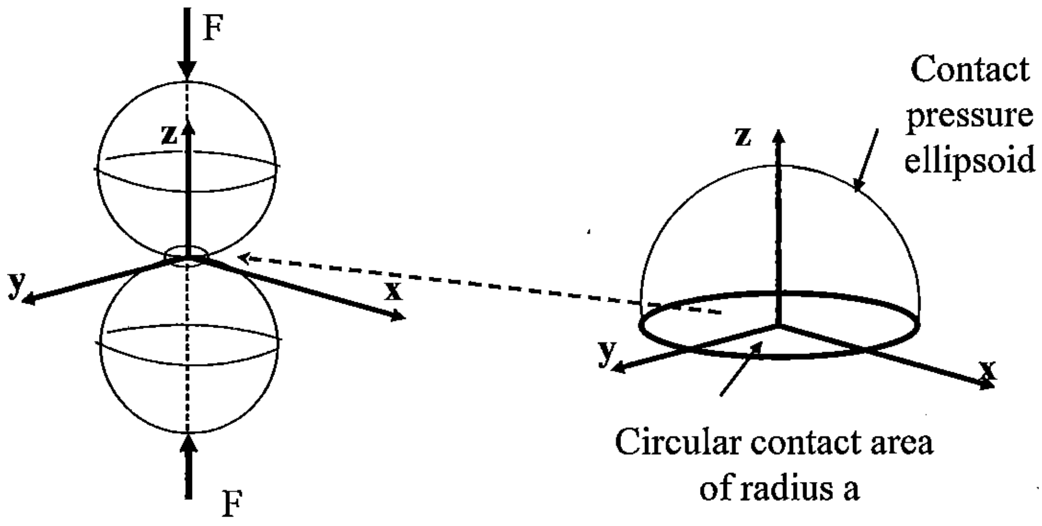
- The above expressions can be easily modified for **other spherical contact problems.**

- **Sphere on a flat plane**
  Here the radius of curvature of a flat surface is infinite, i.e. $1/R = 0$

$$\frac{1}{R} = \frac{1}{R_1} + 0 = \frac{1}{R_1}$$

- **Sphere in a spherical socket**
  Here the radius of curvature is negative i.e.

$$\frac{1}{R} = \frac{1}{R_1} - \frac{1}{R_2}$$



Axisymmetric contact of two spheres

Hertz solutions for 2D Contact of Cylinders

- The initial contact is a **straight line** along the length of the cylinder
- After the application of the external loads, the contact area becomes **rectangular.**
- Consider the contact of two parallel cylinders of different radii R_1 and R_2 of the same length L, pressed by a perpendicular (total) force F along the line joining their centres.

  Here $R' = R$ and $R'' =$ infinite, and the contact area is a rectangular strip of width $2a$ and length L.

  Therefore, the contact pressure can be written as follows:

$$p(x) = p_o \left(1 - \frac{x^2}{a^2}\right)^{1/2}$$

where x is the distance from the centre of the contact area.

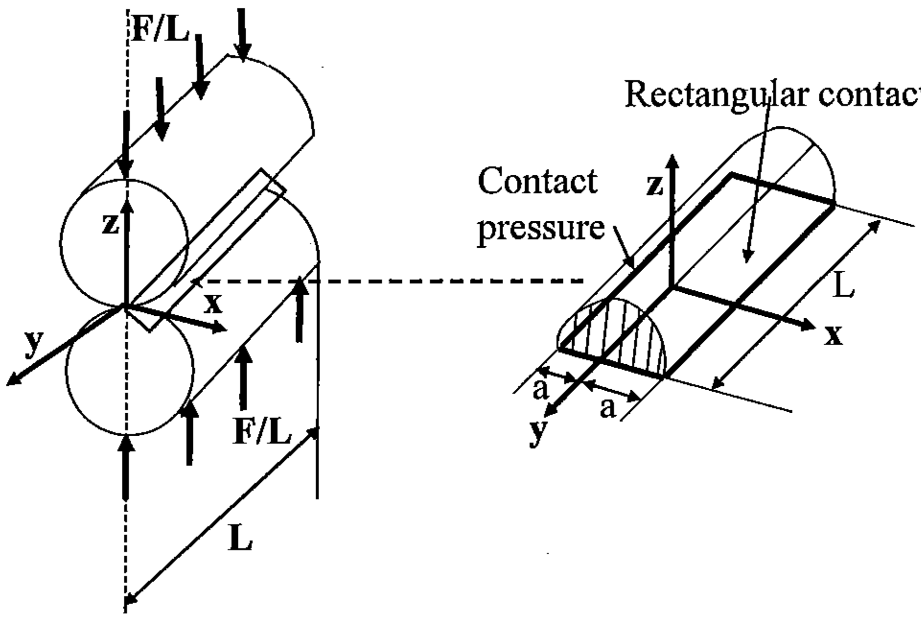
- Hertz solutions for cylindrical contact are:

$$a = \left(\frac{4}{\pi} \frac{F R}{L E^*} \right)^{1/2}$$

$$p_o = \left(\frac{1}{\pi} \frac{F E^*}{L R} \right)^{1/2}$$

$$\delta = \frac{F}{\pi L} \left[\frac{1-\nu_1{}^2}{E_1} \left(2 \ ln \frac{4 R_1}{a} - 1 \right) + \frac{1-\nu_2{}^2}{E_2} \left(2 \, ln \frac{2 R_2}{a} - 1 \right) \right]$$

- For a **cylinder on a flat plane,** R_2 = infinite
- For a **cylinder on a cylindrical groove,** R_2 is negative



2DI contact of two cylinders

First Yield Point In Hertzian Contact Problems

- Hertz solutions can be used to derive expressions for the for the stresses beneath the contact surface
- An interesting feature of Hertzian contact problems is that **the maximum principal shear stress** (which is responsible for the initiation of plastic yielding) **occurs beneath the surface** (rather than on the surface).
- Therefore, **first yield occurs beneath the surface.** ← *Important Observation*
- This is very fortunate since the interior of a homogeneous body is unlikely to contain cracks or imperfections.

First Yield in Contact of Two Spheres

- An analytical solution for the stresses along the *z*-axis of two spheres in contact, beneath the contact surface, can be derived as follows:

$$\sigma_{rr} = \sigma_{\theta\theta} = -p_o\left[(1+\nu)\left(1 - s\ \tan^{-1}\frac{1}{s}\right) + \frac{1}{2\left(1+s^2\right)}\right]$$

$$\sigma_{zz} = \frac{-p_o}{\left(1+s^2\right)}$$

where s is the normalised distance beneath the *z*-axis ($s = z/a$).

- The stresses decrease as the distance from the z-axis increases (away from the contact origin, underneath the surface).

The principal stresses along the z-axis are equal to σ_{rr} , $\sigma_{\theta\theta}$ and σ_{zz}. In order to predict the onset of plastic yield, it is necessary to calculate the **maximum principal shear stress,** τ_1, which in this case is given by:

$$\tau_1 = \frac{1}{2}\left(\sigma_{zz} - \sigma_{rr}\right)$$

Substituting the expressions for the stresses in the above expression for τ_1, and differentiating with respect to the distance s, the maximum value of τ_1 , for $v = 0.3$, can be obtained as follows:

$$(\tau_1)max = 0.31\ p_o$$

This occurs beneath the surface at a distance of:

$$z = 0.48\,a$$

This means that plastic behaviour is expected to occur underneath the surface, rather than on it.

Tresca yield criterion can be used as follows:

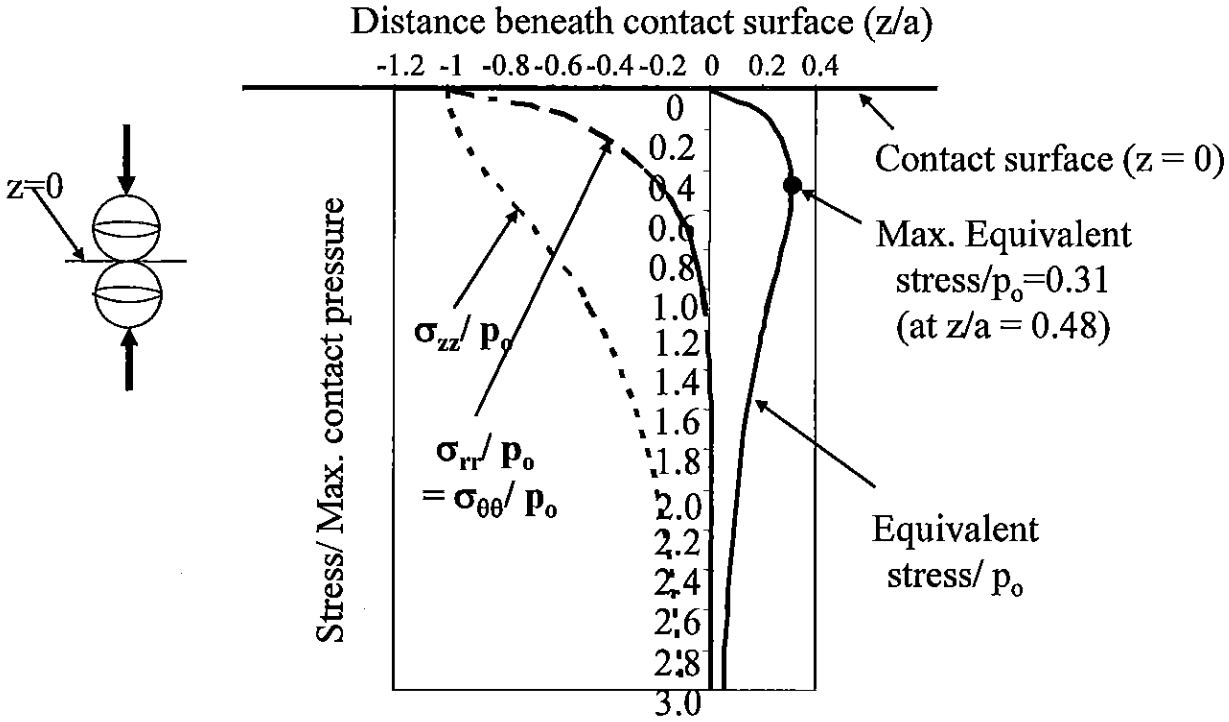
$$Maximum\ shear\ stress = \frac{\sigma_y}{2} = \frac{(\sigma_1 - \sigma_2)}{2}\ or\ \frac{(\sigma_1 - \sigma_3)}{2}\ or\ \frac{(\sigma_2 - \sigma_3)}{2}$$

In the contact of spheres, the maximum value of the shear stress is 0.31 p_o which gives:

$$p_o = 1.61\ \sigma_y$$

The value of the contact force to initiate plastic yield, F_{yield}, is:

$$F_{yield} = 0.6955\ \pi\ \frac{\sigma_y{}^3\ R^2}{E^{*2}}$$



Axisymmetric stresses beneath the surface of two spheres in contact

First Yield in Contact of Two Cylinders

- An analytical solution for the stresses along the z-axis of two parallel cylinders in contact, beneath the contact surface, can be derived as follows:

$$\sigma_{xx} = -p_o\left[\frac{1+2s^2}{\left(1+s^2\right)^{1/2}} - 2s\right]$$

$$\sigma_{zz} = \frac{-p_o}{\left(1+s^2\right)}$$

where $s = z/a$.

- The stresses decrease as the distance from the z-axis increases (away from the contact origin, underneath the surface).

- Note that the third stress, σ_{yy}, is either zero (plane stress problems) or equal to $\nu\,(\sigma_{xx}+\sigma_{zz})$ in plane strain problems (since $\varepsilon_{yy} = 0.0$).

The principal stresses along the z-axis are equal to σ_{xx} , σ_{yy} and σ_{zz}.

In order to predict the onset of plastic yield, it is necessary to calculate the maximum principal shear stress, τ_1, which in this case is given by:

$$\tau_1 = \frac{1}{2}\left(\sigma_{zz} - \sigma_{xx}\right)$$

Substituting the expressions for the stresses in the above expression for τ_1, and differentiating with respect to the distance s, the maximum value of τ_1 , for $\underline{\nu} = 0.3$, can be obtained as follows:

$$\left(\tau_1\right) max = 0.3 \; p_o$$

This occurs beneath the surface at a distance of:

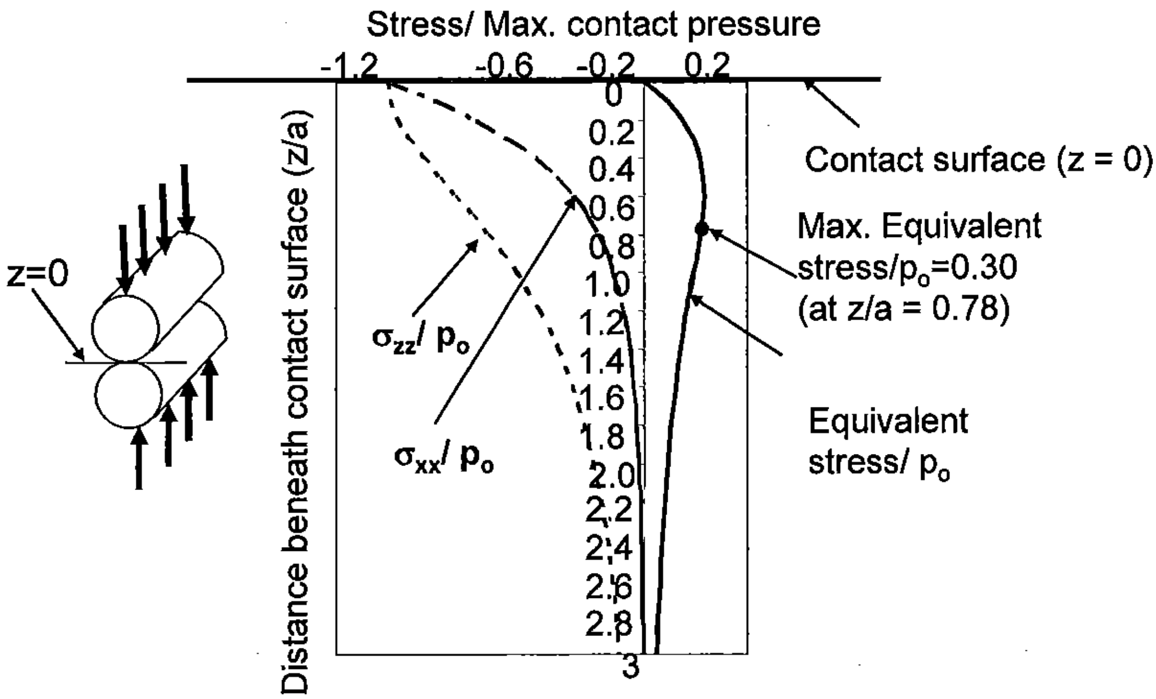
$$z = 0.78\, a$$

In order to predict the value of the contact force F that will cause initial plastic yield underneath the contact surface, the Tresca yield criterion can be used to give:

$$p_o = 1.67 \; \sigma_y$$

The value of the contact force to initiate plastic yield, F_{yield}, is:

$$F_{yield} = 2.78 \; \pi \; \frac{\sigma_y{}^2 \; L \; R}{E^*}$$



2D stresses beneath the surface of two cylinders in contact

5.3 Non-Hertzian Contact Problems

Many practical engineering problems involve the contact of bodies with sharp corners such as punches, large contact areas such as in conforming contact situations, friction in the contact surfaces or are loaded by tangential as well as normal forces.

Such problems are often called "Non-Hertzian" to distinguish them from the Hertz contact problems.

This section covers a range of non-Hertzian contact problems including:

- **(a) Frictional Stick-Slip Behaviour**
- **(b) Tangential loading**
- **(c) Receding Contact Problems**
- **(d) Contact of punches**
- **(e) Contact of wedges**

(a) Frictional Stick-Slip Behaviour-Normal Loading only

- Using the classical definition of static tangential slip, the relationship between the tangential traction and the normal pressure is given by:

$$q \leq \mu p$$

where μ is the coulomb coefficient of friction. **Note that q cannot exceed μp.**

- The direction of q must be **opposite to the direction of motion**, i.e. friction must oppose the sliding motion.

- In most practical problems involving only normal loading, **stick usually occurs in a central region** of the contact area whereas slip occurs at the edges of the contact area.

- If the coefficient of friction is very high, i.e. very rough surfaces in contact, slip may be prevented at all points in the contact area resulting in a **complete stick situation.**

(b) Tangential Loading

If tangential as well as normal loads are applied to the contact bodies, an **irreversible stick-slip behaviour** occurs.

Assumptions:

(i) The development of the tangential traction does not affect the shape and size of the contact area.

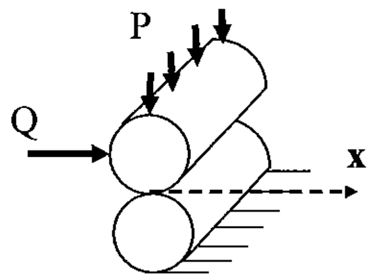
(ii) The effects of normal loading and tangential loading do not interact with each other.

Hence, it will be assumed that the size of the contact area and the normal contact pressure are the same as those of the Hertz theory.

Example 1: Tangential Loading on Two Parallel Cylinders

Consider the stick-slip situation caused by a normal load P followed by a tangential load Q which is less than μP.

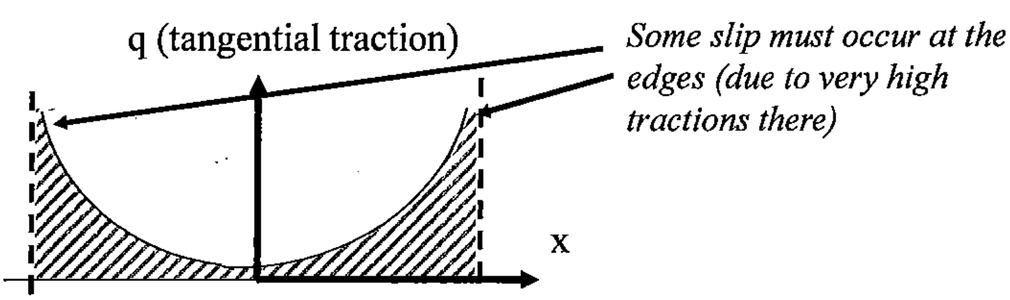
In order to arrive at the final distribution of the tangential traction, the principle of superposition can be applied as discussed below.



Tangential Loading on Two Parallel Cylinders

(i) Assumption of Complete Stick

- If all points stick in the contact area, an infinite traction occurs at the edges of the contact area.
- This in obviously unrealistic since such high tangential tractions are bound to cause slipping near the edges.
- Therefore, slip must occur at the contact edges with a central stick region.



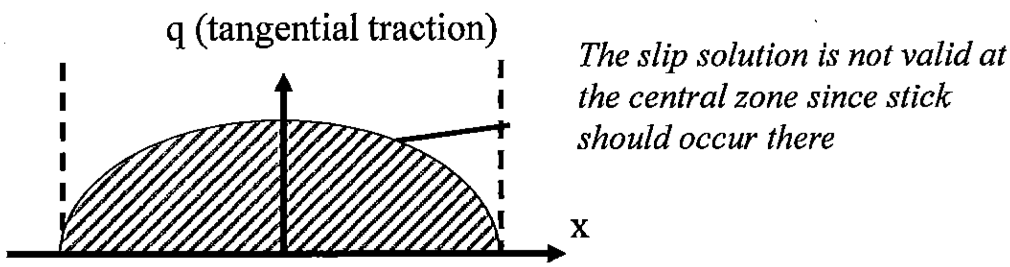
Tangential traction in a complete stick situation

(ii) Assumption of Complete Slip

If the force Q is increased to reach its limiting value of μP, such that the two bodies are on the point of complete slip (sliding), the tangential traction is given by Coulomb Law ($q = \mu\ p$), i.e.

$$q_{slip} = \mu\ po\left(1 - \frac{x^2}{a^2}\right)^{1/2}$$

where $p_o = 2P/(\pi a)$. The tangential traction in a complete slip situation is proportional to the Hertz contact pressure.



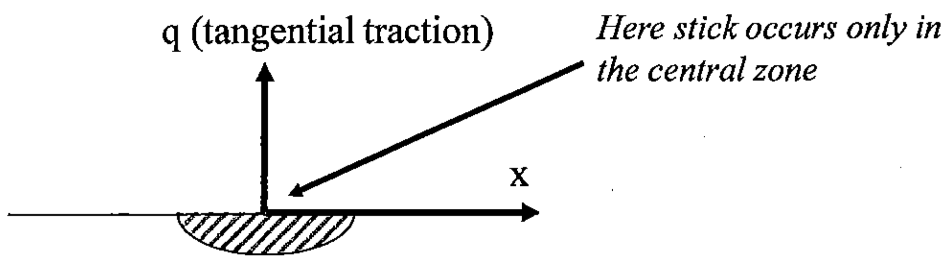
Tangential traction in a complete slip situation

(Tangential Loading on Two Cylinders/ Continued)

(iii) Central stick only

If stick is only in the central zone of width $2c$ (i.e. slip is now restricted to the outer regions only), the tangential traction is given by:

$$q_{stick} = -\frac{c}{a}\mu\ p_o\left(1 - \frac{x^2}{c^2}\right)^{1/2}$$



Stick only in the central zone

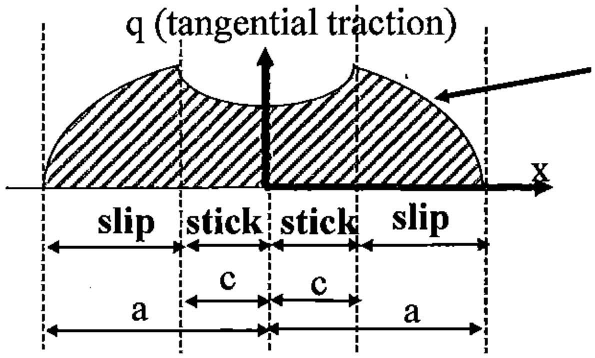
(iv) Combined stick-slip

By superposition, the final solutions for the tangential tractions is:

$$q_{stick-slip} = \mu\, po \left[\left(1 - \frac{x^2}{a^2}\right)^{1/2} - \left(\frac{c^2}{a^2} - \frac{x^2}{a^2}\right)^{1/2} \right]$$

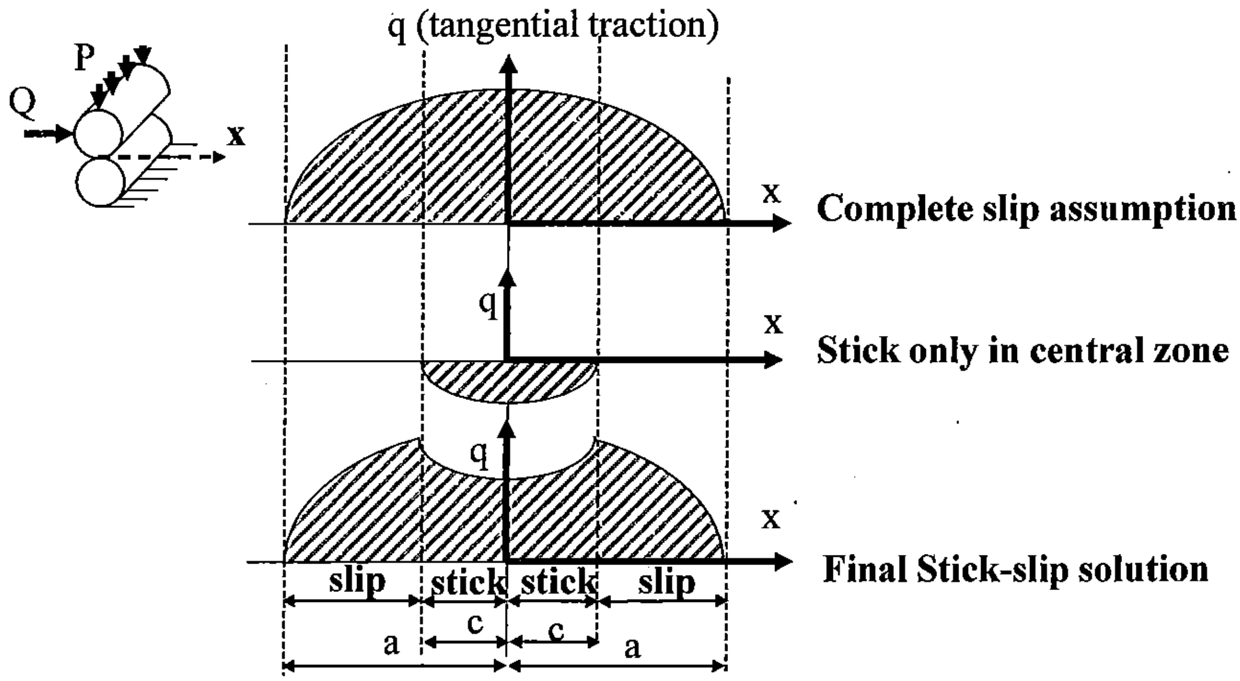
The half-width of the stick zone, *c*, is given as follows:

$$\frac{c}{a} = \left(1 - \frac{Q}{\mu P}\right)^{1/2}$$



The resultant tangential traction is obtained by superposition

Tangential traction in combined stick-slip

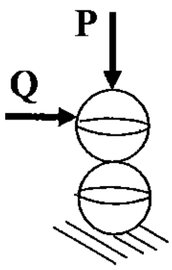


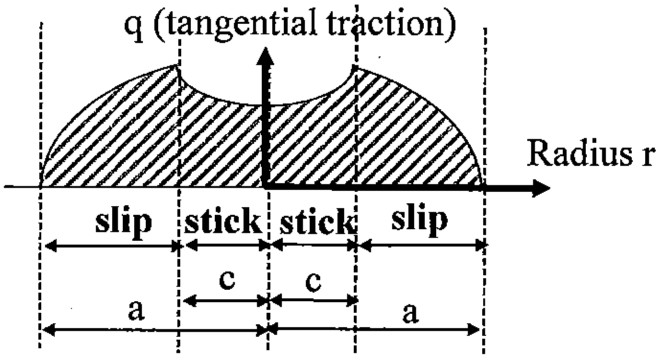
Summary of tangential tractions in two cylinders under normal and tangential loading

Example 2: Tangential Loading on Two Spheres

- The tangential traction distribution is similar to the two-cylinders solution.
- Stick occurs in the central zone, with slip occurring in the outer contact zone.
- The radius c of the central stick circle is given by

$$\frac{c}{a} = \left(1 - \frac{Q}{\mu P}\right)^{1/3}$$

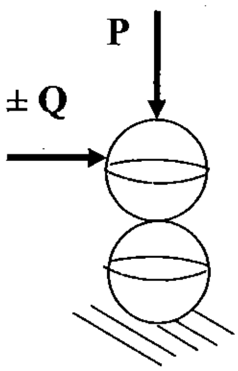


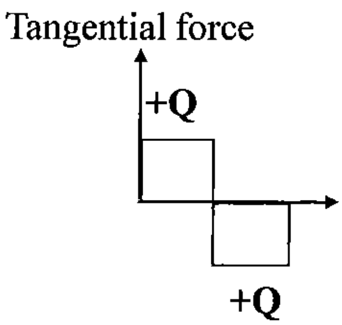


Stick-slip behaviour in two spheres under tangential loading

Example 3: Cyclic Tangential Load on Two Spheres

- In many engineering contact problems, components may be subjected to a cyclic tangential load which will affect the stick-slip behaviour per cycle in the contact area.
- Consider **two spheres** in contact under a normal load P followed by a cyclic tangential load Q ranging from $+Q$ to $-Q$.

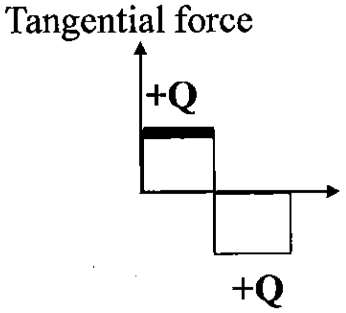


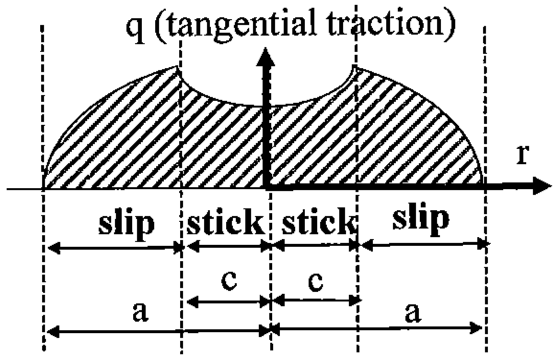


Cyclic Tangential Load on Two Spheres

(i) Apply +Q

After the application of $+Q$, stick occurs in the central region of radius c and slip occurs in the annulus area at the edges of the contact area.

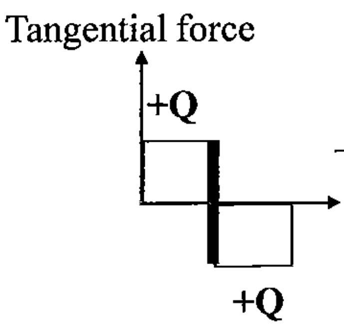


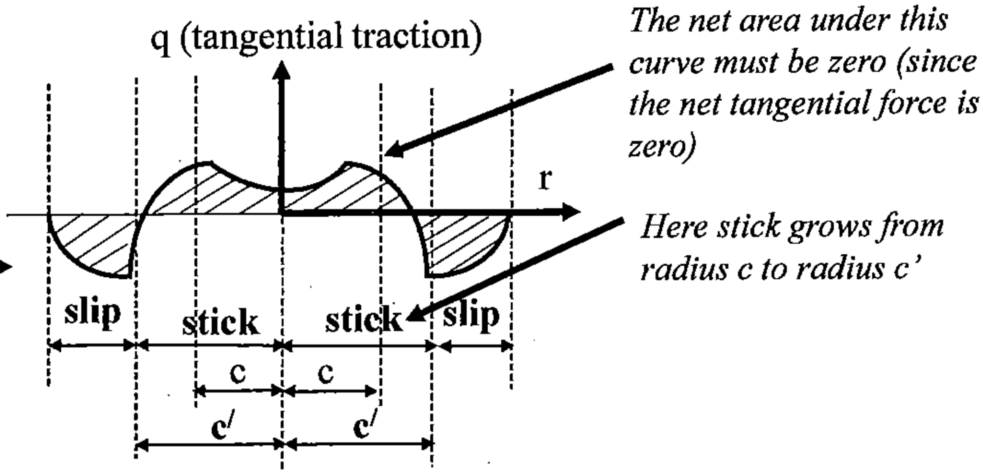


Tangential traction due to +Q

(ii) Unload (i.e. Q=0)

- On unloading, i.e. returning to a zero tangential load, a negative slip occurs at the contact edges and the **stick area grows** from a radius of c to a radius of c'.
- Note that since the resultant tangential load must be zero, the net area under the tangential traction curve must be zero.

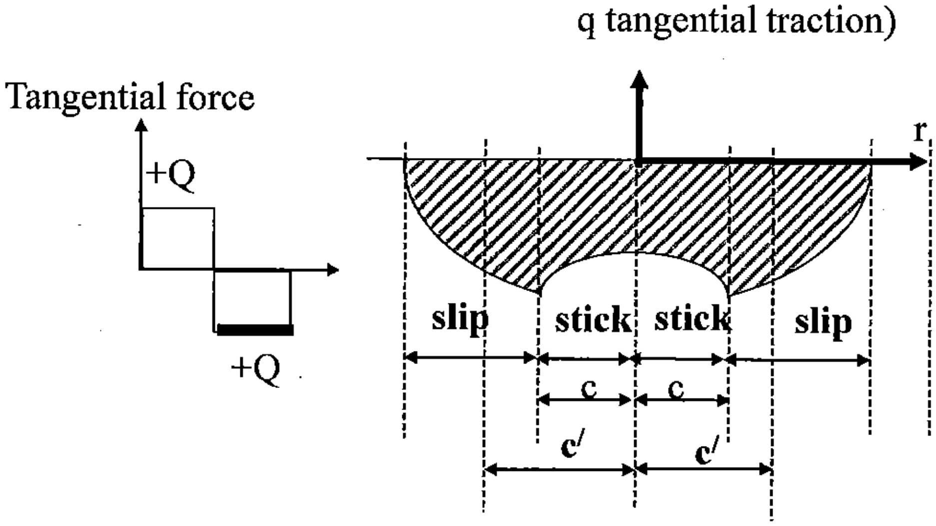




Tangential traction due to unloading Q to 0

(iii) Reverse load to –Q

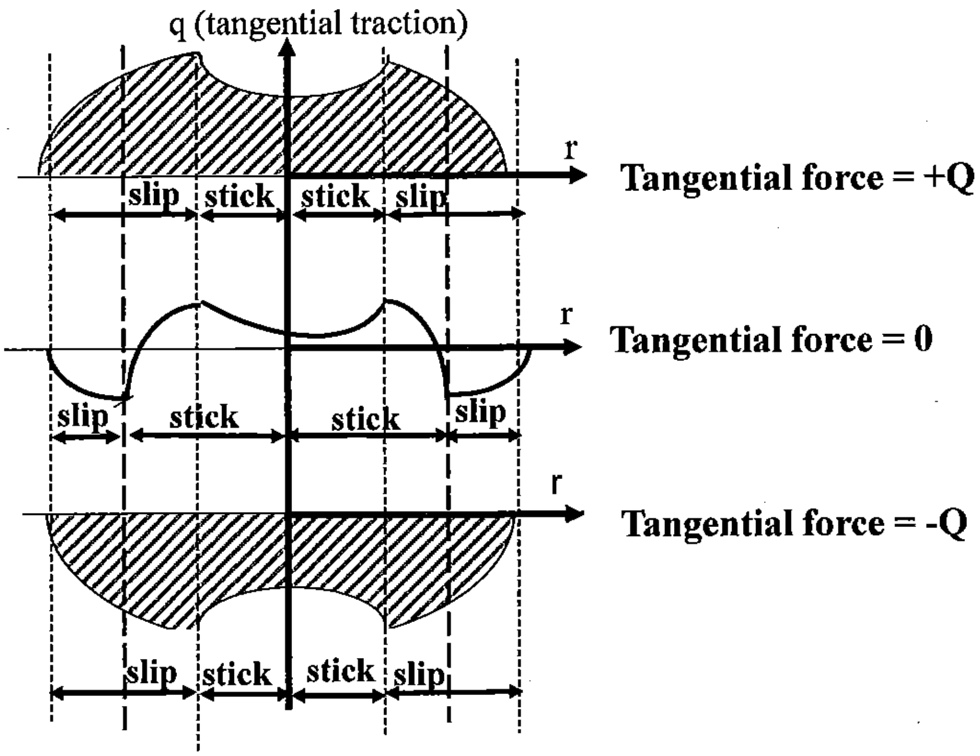
When a reversal of the tangential load occurs, i.e. a force of -Q is applied, the tangential traction is the reverse of that caused by +Q.



Tangential traction due to -Q

(iv) Repeat cyclic behaviour of Q

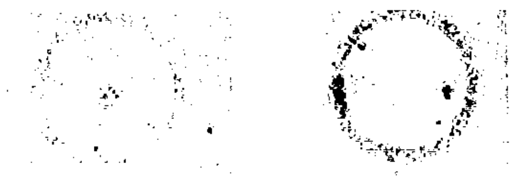
- Repeating the cyclic tangential load from +Q to –Q causes the micro-slip region to oscillate between c and c'.
- This behaviour has been observed in experiments and can produce **fretting surface damage** in the annulus zone between c and c', which can become serious if the normal contact load is high.



Summary of frictional stick-slip under cyclic tangential loading

Fretting Fatigue Phenomenon

Experimental tests have shown that slip and fretting regions occur between two spheres in contact, loaded by an oscillating oblique force Q at an angle of α to the vertical axis, (Reference: K.L. Johnson [1985]).



$\alpha = 20°$ $\alpha = 30°$

$\alpha = 60°$ $\alpha = 90°$

Fretting regions in two spheres under tangential loading

(c) Receding Contact Problems

- Receding contact problems usually occur in **closely conforming contacts** in which the contact area is large in relation to the overall dimensions of the bodies.

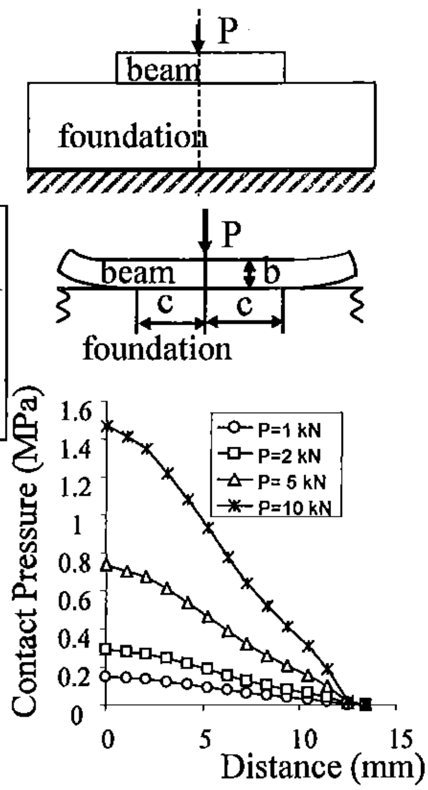
- Upon the application of external loads **the contact area begins to decrease**, i.e. the contact surfaces begin to separate.

- An interesting feature of receding contact problems is that **the size and shape of the contact area is independent of the magnitude of the applied load**, provided its direction remains the same. Therefore, the displacements and stresses increase in direct proportion to the increase in contact load.

- Example: Separation of spherical inclusions from an infinite matrix, which occurs materials such as ceramics which contain inclusions or second phase materials.

Example: Receding Contact of a Beam on a Foundation

Beam width= 100 mm
Beam thickness =10 mm
E1=E2= 200 x10^3 MPa
ν = 0.3
μ =0



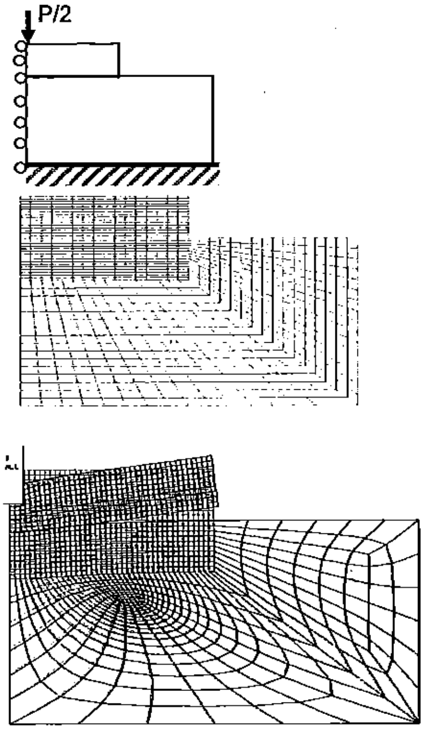
(a) Before loading

(b) After loading

(c) Contact pressure

Receding contact of a beam on a foundation

(Receding Contact of a Beam on a Foundation/ Continued)



Deformed shape of a beam on a foundation

(d) Contact of Punches

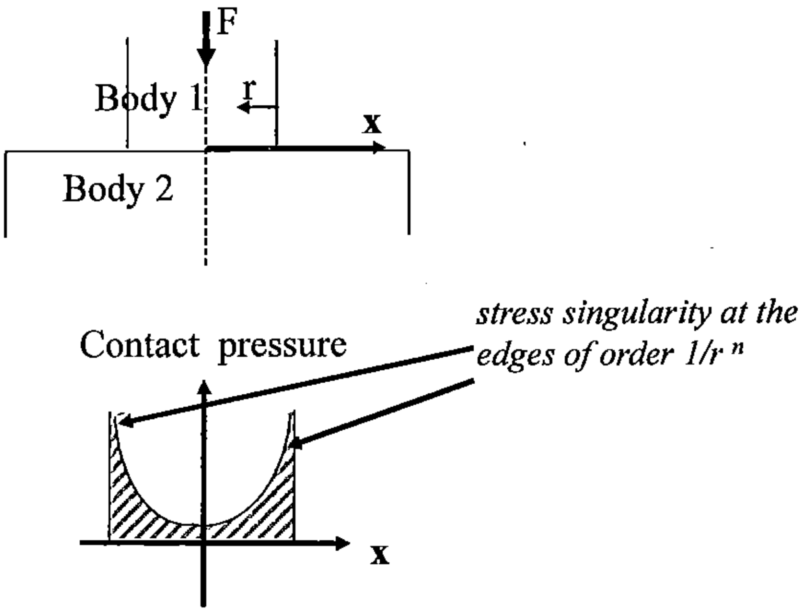
- In problems involving **sharp corners**, such as the contact of a punch on a foundation, the analytical solutions for the contact pressure show that **stresses are infinite at the edge** (corner) of the contact area.
- This causes a "**stress singularity**" at the edge of the contact area of the order $1/\rho^n$, where ρ is the distance from the contact edge. For a rigid square punch, n= 0.5.
- This hypothetical stress-singularity at the contact edge is very similar to that used in defining stresses at the tip of a crack.
- In real life, such stresses cannot exist because the material will undergo **plastic yielding** before reaching very high stresses

Dunders Constants

$$\alpha = \frac{\left(\frac{1-\nu_1}{G_1}\right) - \left(\frac{1-\nu_2}{G_2}\right)}{\left(\frac{1-\nu_1}{G_1}\right) + \left(\frac{1-\nu_2}{G_2}\right)} \quad ; \quad \beta = 0.5\left[\frac{\left(\frac{1-2\nu_1}{G_1}\right) - \left(\frac{1-2\nu_2}{G_2}\right)}{\left(\frac{1-\nu_1}{G_1}\right) + \left(\frac{1-\nu_2}{G_2}\right)}\right]$$

where G is the shear modulus defined as follows: $G = \frac{E}{2(1+\nu)}$

- Note that the subscript 1 refers to the foundation while subscript 2 refers to the wedge or punch.
- The parameter α ranges from
  α= +1 (rigid punch) to α= -1 (rigid foundation)
- β ranges from
  β= +0.5 (rigid punch) to β= -0.5 (rigid foundation



Contact of a punch and a foundation

(e) Contact of Wedges

In the contact of wedges, a discontinuity occurs in the profile of the contact surfaces within the contact area.

Example 1 : Blunt conical indentor on a foundation

For a conical indenter with angle 2θ, the elastic contact stresses, provided the contact area remains small compared to the dimensions of the contacting bodies, are:

$$p_o = \frac{E^*}{2\tan\theta}\cosh^{-1}\left(\frac{a}{r}\right)$$

where r is the radial distance from the centre of circular contact area and a is the outer radius of the contact area.

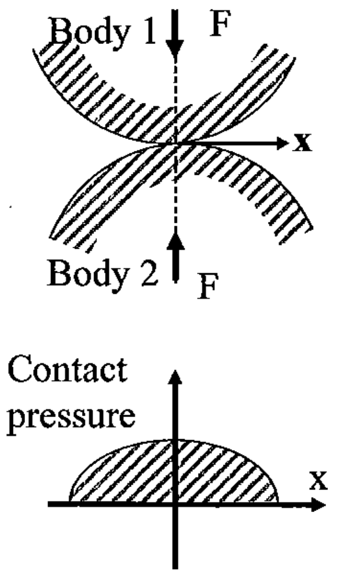
Note that this solution contains a stress singularity, i.e. a theoretically infinite stress, at the apex of the wedge.

Example 2 : A 2D blunt wedge pressed into a foundation

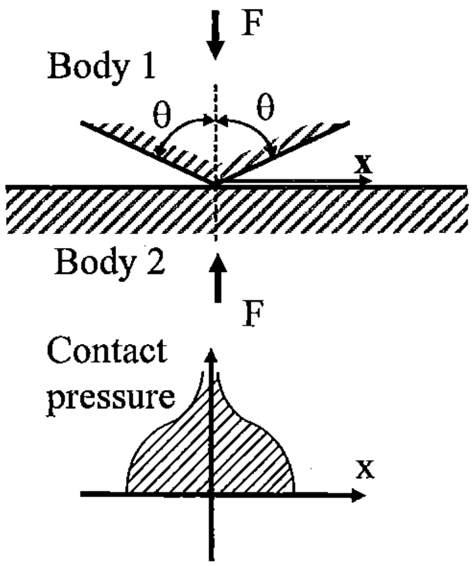
For a 2D wedge with angle 2θ, a similar analytical solution for the contact stresses can be obtained, provided the contact area remains small compared to the dimensions of the contacting bodies, as follows:

$$p_o = \frac{E^*}{\pi\tan\theta}\cosh^{-1}\left(\frac{a}{x}\right)$$

where x is the distance from the centre of contact and a is the semi-width of the contact area.



(a) Contact of continuous curved surfaces (Hertzian contact)



(b) Contact of a wedge on a foundation (stress singularity at the apex)

Contact of a 2D sharp wedge on a foundation

5.4 Relationships Between Contact Variables

- In contact problems, points on the contact interface (unlike the other boundary surfaces) **do not have prescribed displacements or forces**.
- **Tractions** may be considered as "stress vectors".
- **Three relationships** must be satisfied:

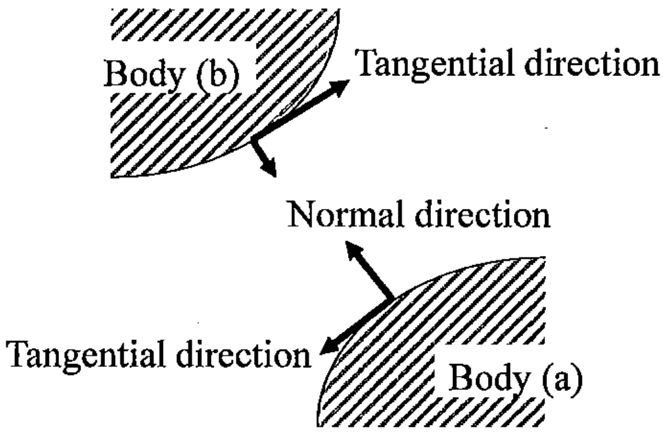
  (i) **Continuity of displacements** in the normal contact direction (i.e. no overlap).

  (ii) **Equilibrium conditions** (i.e. equal and opposite tractions or forces).

  (iii) **Stick or slip conditions** in the presence of friction.

(Relationships Between Contact Variables/ Continued)

- In **heat conduction** problems, only two variables must be linked, temperature (T) and temperature gradient or flux.
- It is convenient to refer to the local **normal and tangential directions**, n and t, respectively, rather than the Cartesian x and y-axes.



Local normal and tangential directions in contact problems

Contact Interface Conditions

(a) Thermal contact conditions

In perfect thermal conduction between two contacting bodies, denoted by superscripts (a) and (b), **continuity of temperatures and heat fluxes** must be maintained at the contacting nodes.

$$T^{(a)} = T^{(b)}$$

$$k^{(a)} \left(\frac{dT}{dn} \right)^{(a)} = k^{(b)} \left(\frac{dT}{dn} \right)^{(b)}$$

where T and dT/dn are temperature and normal temperature gradient, respectively, while k is the thermal conductivity.

(b) Frictional Stick Conditions

- When the contacting surfaces are stuck together (not allowed to slip in any direction), **continuity of displacements** in the contact area must be maintained resulting in the following relationships:

$$u_n^{(a)} = u_n^{(b)} + \delta_n$$

$$u_t^{(a)} = u_t^{(b)}$$

where u_n and u_t are the normal and tangential displacement components, respectively, and δ_n is the normal gap or clearance between the contact points which must be accounted for in the contact of curved non-conforming bodies such as spheres.

- The **equilibrium conditions** must be satisfied by the following relationships:

$$t_n^{(a)} = -\, t_n^{(b)}$$

$$t_t^{(a)} = -\, t_t^{(b)}$$

where t_n and t_t are the normal and tangential traction components, respectively.

(c) Frictional slip conditions

- Slip occurs when the ratio of tangential to normal traction equals or exceeds the value of the Coulomb coefficient of friction, μ.

- The **continuity of displacements** must be maintained in the normal direction, but the **tangential displacements are not restricted**, resulting in the following equations:

$$u_n^{(a)} = u_n^{(b)} + \delta_n$$

$$u_t^{(a)} = u_t^{(b)} + \delta_t$$

where δ_t is the amount of slip in the tangential direction, which is an unknown variable that requires a further equation. This extra equation is provided by the **slip conditions**, resulting in the following traction conditions:

$$t_t^{(a)} = \pm\, \mu\, t_n^{(a)}$$

$$t_n^{(a)} = -\, t_n^{(b)}$$

$$t_t^{(a)} = -\, t_t^{(b)}$$

- The sign of the tangential traction must be opposite to the direction of the sliding motion.

(d) <u>Interference or clearance conditions</u>

- This condition occurs when there is an **interference (initial overlap) or clearance (initial gap)** between the contact surfaces.

- Interference Example: A cylinder or shaft sleeve can be heated and then fitted on another shaft (with a slightly larger diameter). The assembly is allowed to cool creating favourable compressive hoop stresses on the inner shaft.

- The coupling of the contact variables is the same as above.
  Contact gap : δ_n is positive
  Contact interference: δ_n is negative

5.5 FE Contact Formulations

Practical contact problems require some form of **iterations** and, if friction is present, **load incrementation**.

The analysis of practical contact problems is **more complicated** than other problems for the following reasons:

- Contact problems are **non-linear** because the contact area does not change linearly with the applied load (even if the material behaviour of each contacting body may be assumed linear elastic)

- In most practical problems, the actual **contact area is unknown in advance** and can only be estimated at first.

- In frictional contact problems, the behaviour may be **history-dependent.** which requires the application of small load increments.

- In frictional contact problems where a stick-slip behaviour occurs, the **direction of slip** is difficult to arrive at.

Difficulties in FE modelling contact problems

(i) Arriving at the **correct contact status** may require many load increments and iterations.

(ii) If the contact bodies are **insufficiently restrained**, i.e. under-restrained, a singular solution matrix may be obtained.

(iii) **Complete sliding** or large sliding may occur, i.e. the initial contact nodes move away from each other, which may affect the accuracy of the initial contact constraints.

(iv) In **multiple contact problems** in which more than two bodies are in contact, it is difficult for the FE code to monitor the contact status in more than one contact region.

(v) **In 3D problems**, slipping occurs in a plane (rather than a line), which makes it difficult to determine the direction of the tangential displacement vector.

(vi) Contact analysis requires a **fine mesh discretisation** of the contact surfaces, which may be impractical in large 3D problems. **Sub-structuring** may be used to reduce the degrees of freedom of the system.

Load Incrementation in Contact Analysis

- In complex frictional contact problems, the behaviour of the contacting bodies is usually **dependent on the load history**.
- This means that the load must be applied in **small increments**, and for each increment the **iterations are performed** to satisfy the contact conditions.
- The computed displacements and tractions are **updated at the end of each increment**, until the final load is reached.
- The **size of the increments can be varied**, and can be adjusted depending on the latest state of the contact area.

Contact Iterations and Checks

An iterative scheme is necessary to check that the final solution (contact pressure and the extent of the contact area) is physically acceptable.

After each iteration, the following three checks should be performed.

(i) Overlap Check

The displacements of the nodes just outside the assumed contact area should be checked to determine whether overlap has occurred.

The detection of overlap is an indication that the contact area is too small for the given load, and the overlapped nodes must be included in the contact area in the next iteration.

(ii) Tensile Contact Stress Check

All contacting nodes should be checked to determine if any tensile stresses have been generated.

The detection of tensile stresses in contact area is an indication that the contact area is too large for the given load, and these nodes must be released from contact in the next iteration.

(iii) Friction Slip Check

If friction is present, the ratio of the tangential to normal tractions must not exceed the value of μ.

If it does, the relevant points should be allowed to slip in the next iteration and the Coulomb friction slip condition is imposed on the tangential traction.

FE contact algorithms

(a) **Transformation Matrix Method**

- This method works by enforcing the displacement continuity condition between the contacting nodes
- For each node in contact, the displacement is considered identical to its contacting node, allowing one of the displacements to be directly eliminated.
- The size of the structural stiffness matrix is therfore reduced.
- This method is considered suitable for problems where the contact region is well defined and known not to change during the loading.

Advantages and Disadvanyages of the Transformation Matrix Method

Advantages

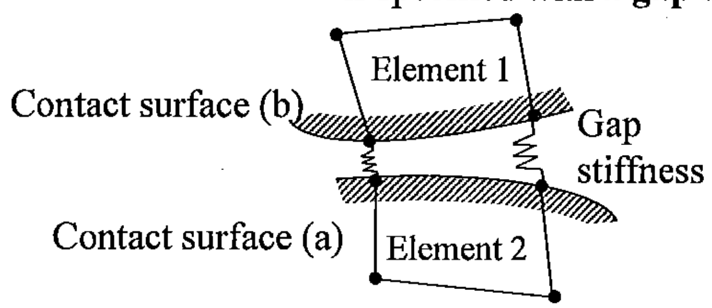
- Exact application of the displacement constraint
- Reduction in the size of solution matrix.
- Does not require major changes in the FE Coding.

Disadvantages

- Many solution stages are required before convergence is obtained.
- Not suitable for problems where the size of the contact area changes during the analysis, this approach requires many solution stages and is not recommended.

(b) Penalty Function Method (Gap Elements)

- A **penalty function** is used to approximately impose the contact constraints through the use of additional **'gap' elements** inserted between the contacting surfaces.
- Two types of gap elements can be used:
  (a) **Line elements** linking node-pairs
  (b) **Area or interface elements** between element faces
- **Node-on-node contact** is specified with **a gap stiffness**



Gap Stiffness in the Penalty Function Method

- The **gap stiffness** can be either defined by the user or automatically chosen by the FE code.

- The **magnitude of the gap stiffness influences the enforcement of the contact constraints**, as follows:

  (a) When the nodes are not in contact, the gap stiffness is negligible and has no effect on the deformation of the bodies.

  (b) When the nodes are in contact, the gap stiffness becomes very high in order to prevent any overlap or penetration of the two bodies, and to allow the compressive contact pressures to be transmitted across the interface.

- **Frictional contact analysis** can be incorporated in this method since the tangential gap stiffness can be related to the coefficient of friction.

Advantages and Disadvanyages of the the Penalty Function Method

Advantages

- Relative ease of implementation and monitoring of the contact status
- The size and bandwidth of the solution matrix is maintained.

Disadvantages

- Contact node pairs must be relatively aligned with each other.
- Only limited amounts of sliding are allowed.
- The penalty method is the penalty number (gap stiffness) must be chosen so that it is not very low (which can cause poor convergence and poor enforcement of the contact constraints) or very high (which can cause a singular solution matrix or ill-conditioning).

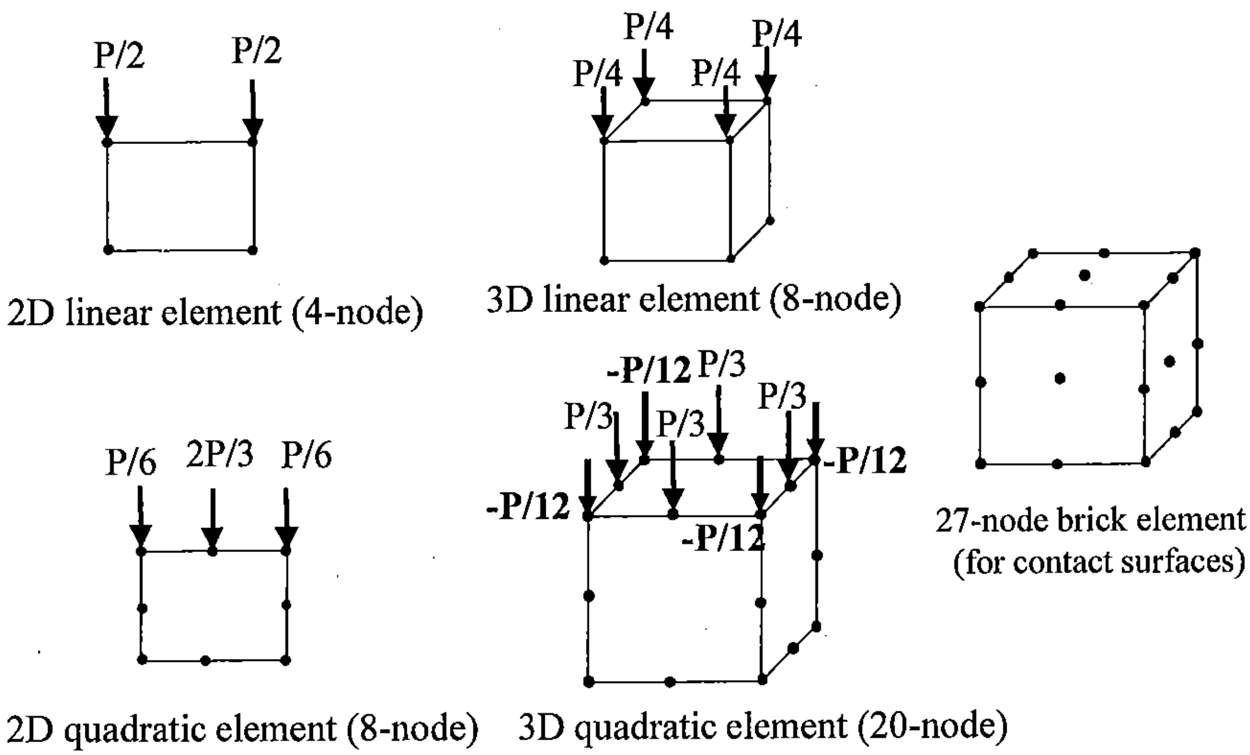
A warning about 3D Quadratic (20-node) brick elements in contact

- In 3D problems in which 20-node brick elements are used in the contact interfaces, an **incorrect contact separation** may be reported.

- This is due to the **3D quadratic shape function** used in these elements, which converts pressure into **"equivalent nodal forces"** which are not equally distributed on the contacting nodes.

- For 20-node brick elements, the equivalent forces on the corner nodes are in an **opposite direction** to the forces on the mid-point nodes.

- To overcome this problem, **27-node brick elements** can be used at the contact interface. with 20-node brick elements used elsewhere. These elements have an additional node at each face of the brick element, and an extra point at the centroid of the element.



Kinematically equivalent nodal forces on element faces

(c) Lagrange Multiplier Constraint Method

- Lagrange multipliers are used to **exactly enforce the contact constraints** between the contacting surfaces.

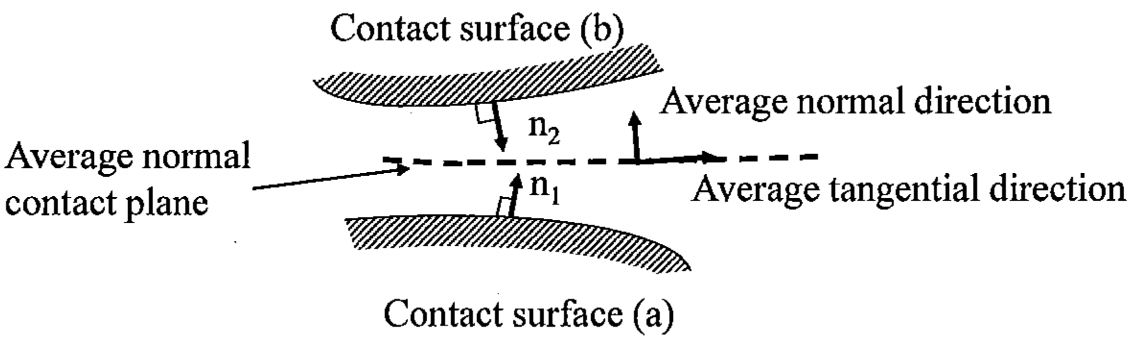
- **Additional variables are introduced** in the solution matrix in the form of symmetric terms added as extra rows and columns to the stiffness matrix.

- The **size of the solution matrix is increased** and zero diagonal terms are introduced into the stiffness matrix.

- **Independent or free meshing** of the contacting surfaces is allowed.

- **Frictional conditions** can be imposed at the interface, usually by specifying the contact region as initially sticking, and subsequently allowing the nodes to slip if the ratio of tangential to normal force exceeds the value of the coefficient of friction. The direction of sliding is chosen to oppose the tangential forces.

- Sliding around curved surfaces can be allowed by defining average normal and tangential directions.



Average normal contact plane

Advantages and Disadvantages of Lagrange Multiplier Method

Advantages

- Nodal alignment of meshes is not necessary
- Large scale sliding problems can be modelled.
- Accurate build-up of the load history can be obtained.

Disadvantages

- Convergence difficulties when coarse meshes are used.
- Complexity of programming of the Lagrange multipliers in the FE code.
- Many solution stages may be required if the number of contacting nodes is large.

Some Guidelines for using FE codes in Contact Problems

(a) Plotting Deformed shapes in Contact Problems

- FE codes usually automatically multiply the actual displacements by a large **magnification factor** in order to plot the deformed shapes.

- In contact problems, this may result in a **misleading plot showing an overlap** of the contact surfaces.

- The FE user can set the magnification factor used in plotting deformed shapes to unity, i.e. show the actual deformations. With a magnification factor of 1, the user can zoom on the contact area and check that overlap of the contacting elements has not occurred.

(b) Specifying the initial contact surfaces

- In modelling advancing contact problems, such as a sphere on a plate, the user should identify an **approximate contact area** (i.e. an initial guess of the contact surfaces).

- The FE code will use the initial guess as a starting point to perform contact iterations to arrive at the correct contact area.

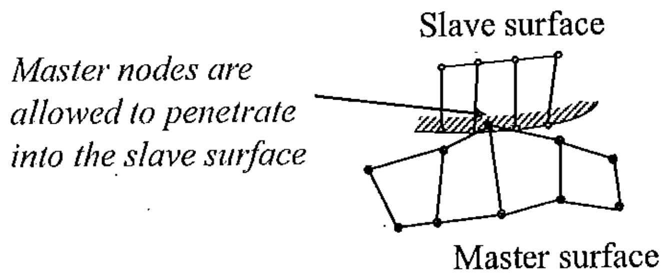
(c) Mesh refinement of the contact surfaces

- Meshes should be refined at the contact surfaces, Ideally, a **similar mesh spacing** should be used on both sides of the contact surfaces (even when the Lagrange Multiplier approach is used)

- However, this may not be possible in large 3D problems. If different mesh refinements are used, then the surface with the less refined mesh should be designated as the Master surface.

(d) Master-Slave Contact Algorithms

- In order to improve the solution convergence in practical contact problems, a master-slave algorithm can be used which allows a limited amount of violation of the contact conditions.
- Slave nodes are not allowed to penetrate into the Master surface
- The master surface is used to define the normal and tangential contact directions.
- Rigid surfaces should be chosen as the master surfaces.
- If both bodies are deformable, the stiffer body should be chosen as the master surface.
- If both bodies are of a similar stiffness, the one with the less refined mesh in the contact region should be chosen as the master surface.



Master-slave contact surfaces

5.6 Contact Examples

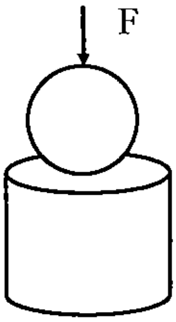
5.6.1 Contact Example 1: Sphere on a Plate (Advancing Contact)

Physical Attributes

This example features an advancing contact problem in which the applied load increases (advances) the contact area.

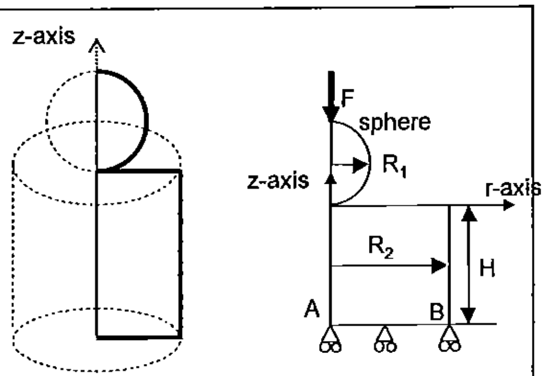
The main attributes of the problem are:

- Advancing contact with a closing gap
- Hertzian contact
- Frictionless contact
- Deformable bodies
- Curved contact surfaces

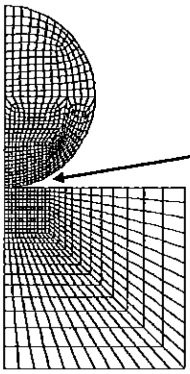


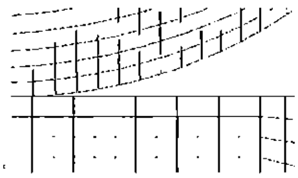
(Sphere on a Plate Example/ Continued)

Problem Definition

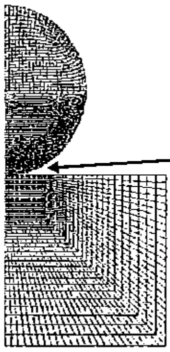
Geometry	Axisymmetric problem $R_1 = 1$ m $R_2 = 2$ m $H = 2$ m $F = -1.172 \times 10^9$ N 
Material Properties	Young's modulus, $E = 200 \times 10^9$ N/m^2 Poisson's ratio, $\nu = 0.3$
Boundary Conditions	Prescribed zero z-displacement on line AB
Loading	A prescribed point load (F) acting on the top node

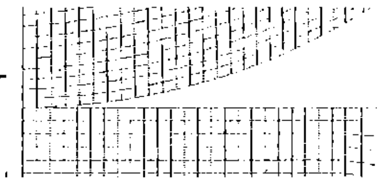
(Sphere on a Plate Example/ Continued)





Magnified view of the contact area

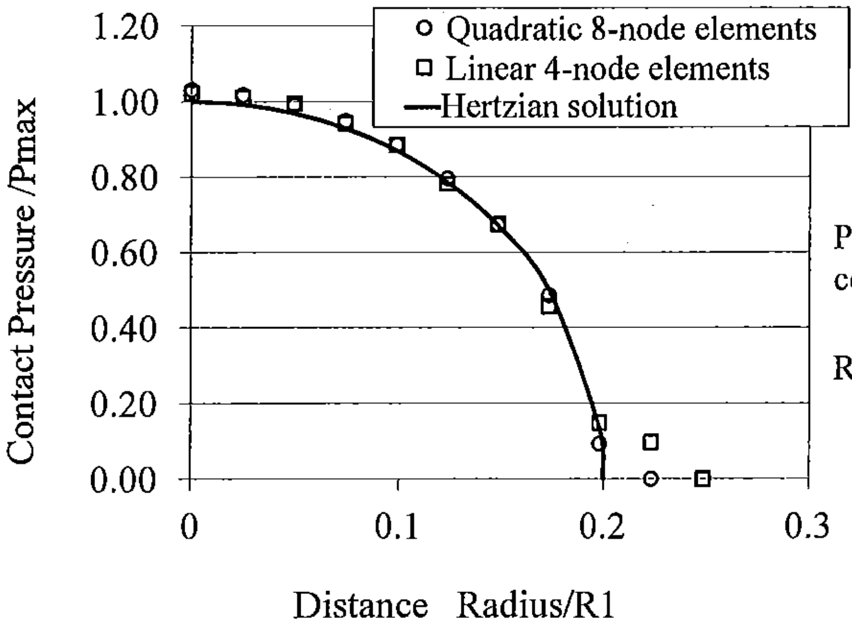




Magnified view of the contact area

FE meshes for the sphere on a plate problem

(Sphere on a Plate Example/ Continued)



Pmax=Maximum contact pressure

R1= Sphere radius

Reference solutions for the sphere on a plate problem

5.6.2 Contact Example 2: Punch And Foundation (Stationary contact with friction slipping)

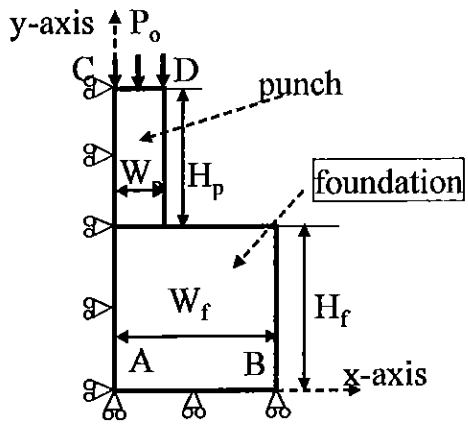
Physical Attributes

This example features a two-dimensional contact problem with friction in which the contact area remains constant during contact (stationary contact).

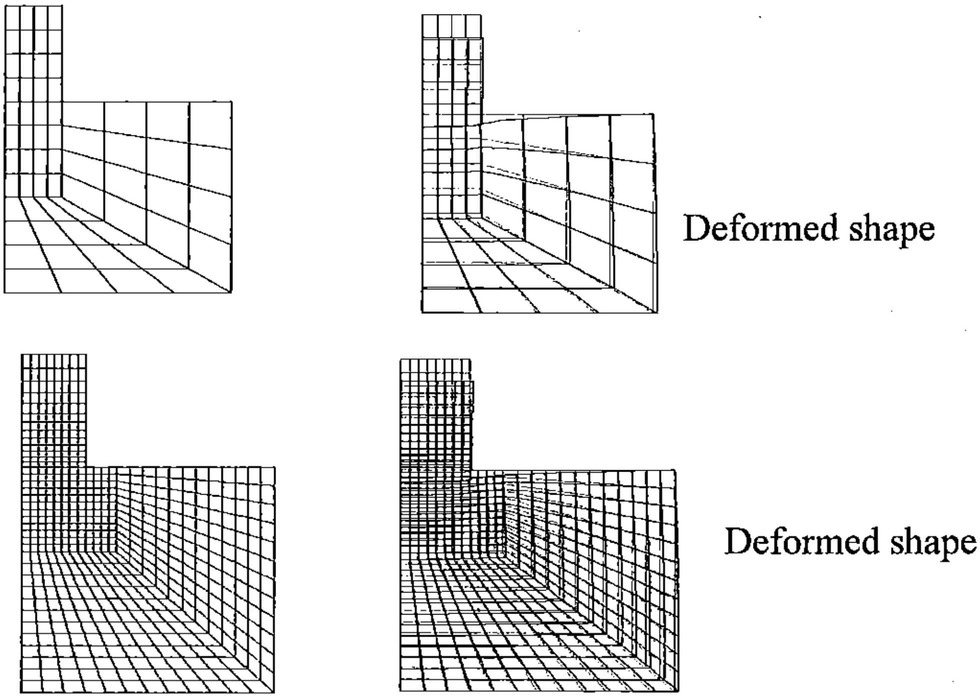
The main attributes of the problem are:
- Non-Hertzian contact
- Stationary contact area
- Non-zero coefficient of friction
- Stick-slip contact region
- Infinite stresses at the edge of contact

(Punch on Foundation Example/ Continued)

Problem definition

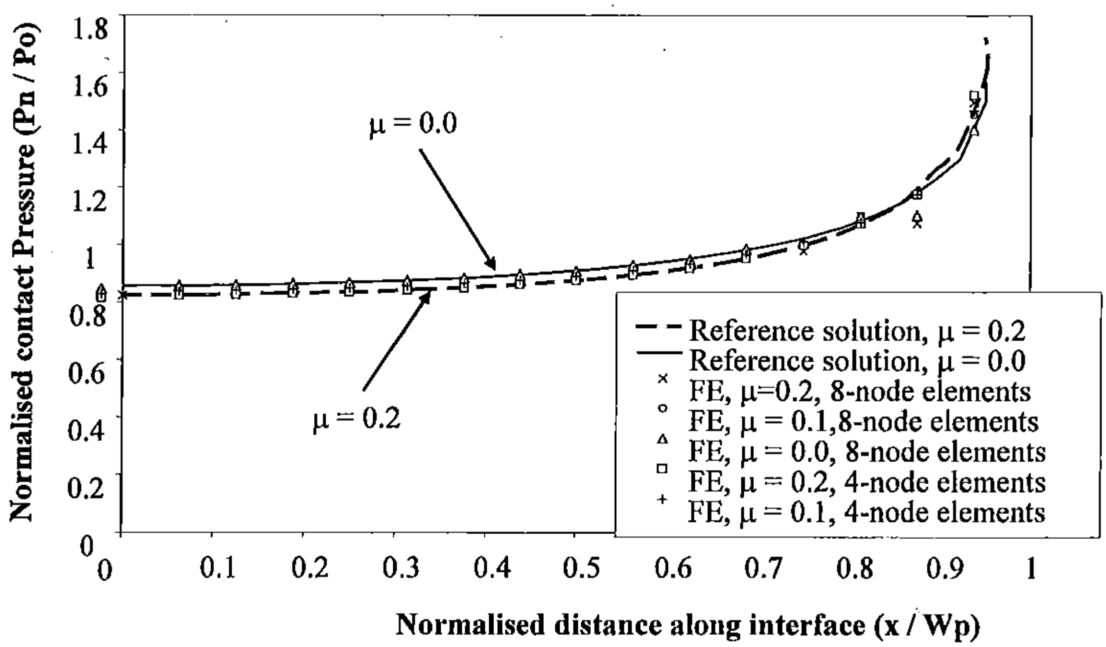
Geometry	2D Plane Strain $W_p = 1.0$ m $H_p = 2.0$ m $W_f = 4.0$ m $H_f = 4.0$ m $P_o = 1.0$ N/m²	
Material Properties	Young's modulus, $E_p = E_f = 1.0$ Poisson's ratio, $\nu_p = \nu_f = 0.3$	
Boundary Conditions	Prescribed zero y-displacement on line AB Prescribed zero x-displacement on line AC	
Loading	A prescribed pressure (P_o) on line CD	

(Punch on Foundation Example/ Continued)

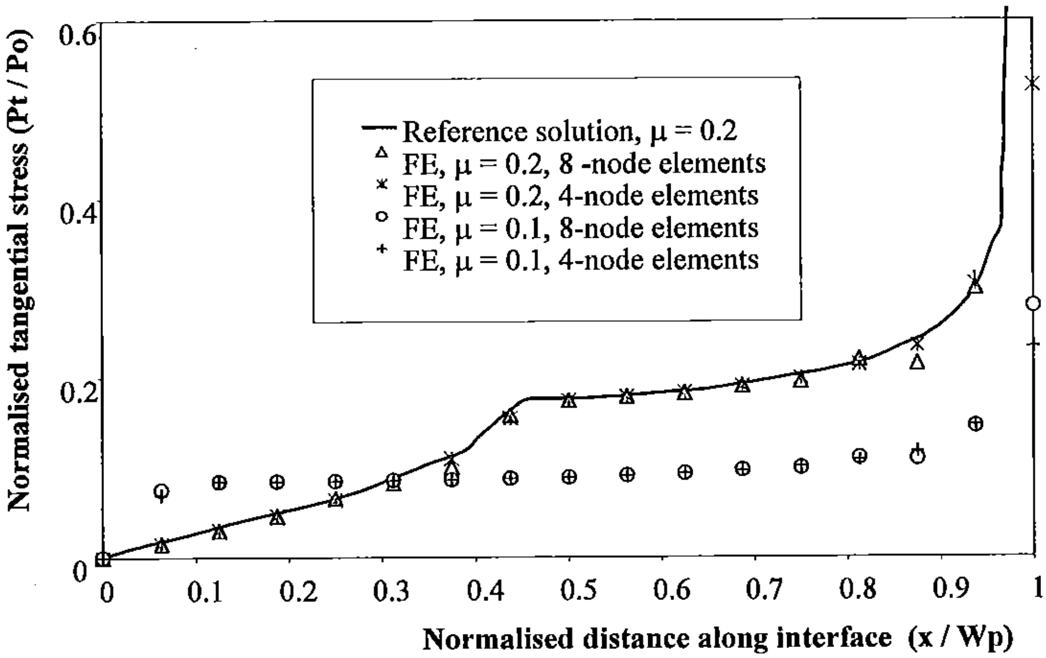


FE meshes for the punch on foundation problem

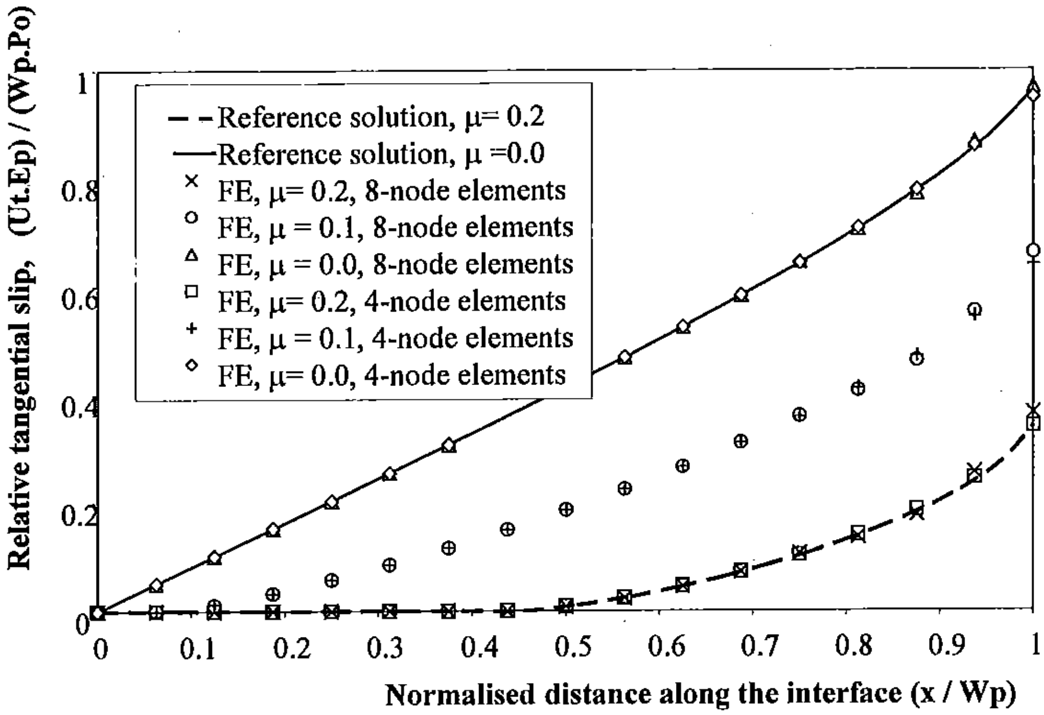
(Punch on Foundation Example/ Continued)



Contact Pressure for the punch on foundation problem

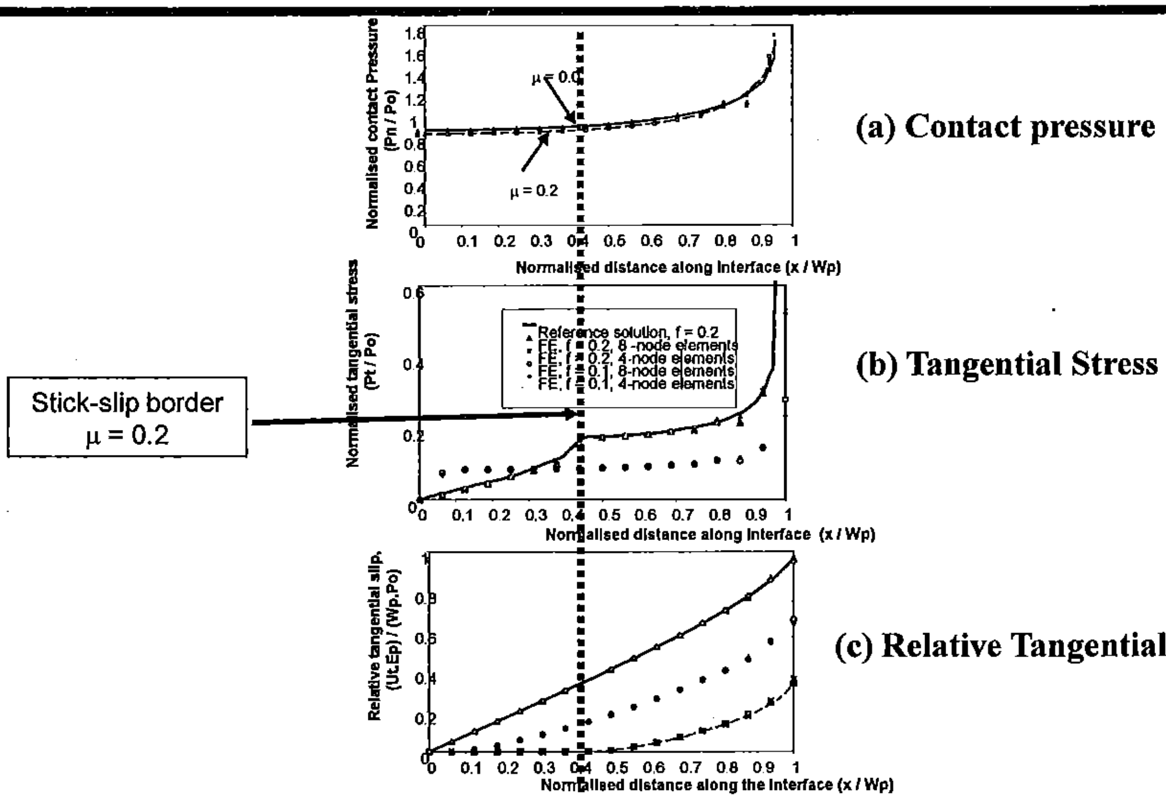


Tangential stress for the punch on foundation problem



Relative slip for the punch on foundation problem

(Punch on Foundation Example/ Continued)



Reference solutions for the punch on foundation problem

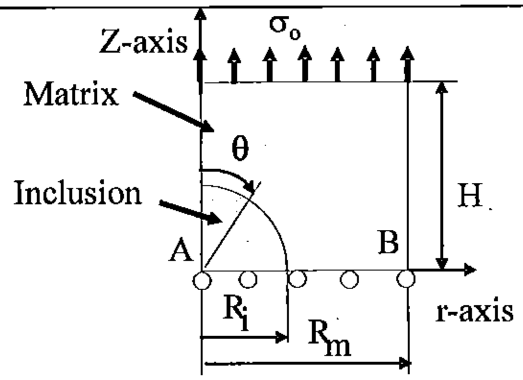
5.6.3 Contact Example 3: Inclusion In a Matrix (Receding Contact)

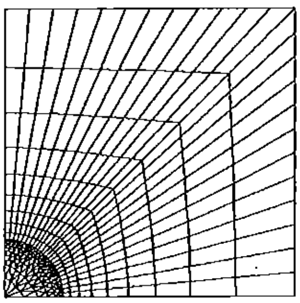
<u>Physical Attributes</u>

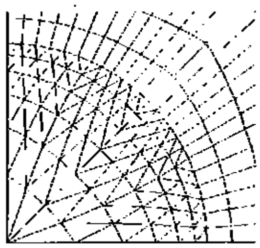
This example features a receding contact problem between a spherical inclusion and a homogenous infinite matrix.

The main attributes of this problem are:
- Receding contact situation
- Curved contact surfaces
- Tensile and compressive applied loads

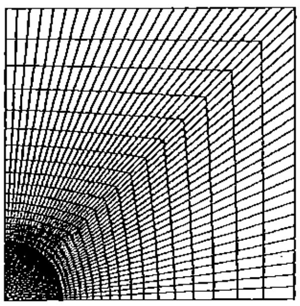
Problem definition

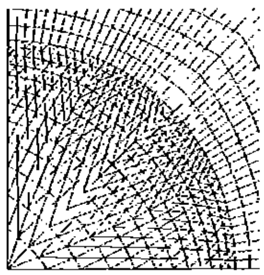
<table>
<tr><td>Geometry</td><td>Axisymmetric
$R_m / R_i = 5.0$
$H / R_i = 5.0$
$\sigma_o = 1.0$ N/m^2

</td></tr>
<tr><td>Material Properties</td><td>Young's modulus, $E_i = E_m = 1.0$
Poisson's ratio, $\nu_i = \nu_m = 0.3$</td></tr>
<tr><td>Boundary Conditions</td><td>Prescribed zero z-displacement on the bottom surface</td></tr>
<tr><td>Loading</td><td>A prescribed stress (s_o) on the top surface</td></tr>
</table>





Magnified view of the contact area

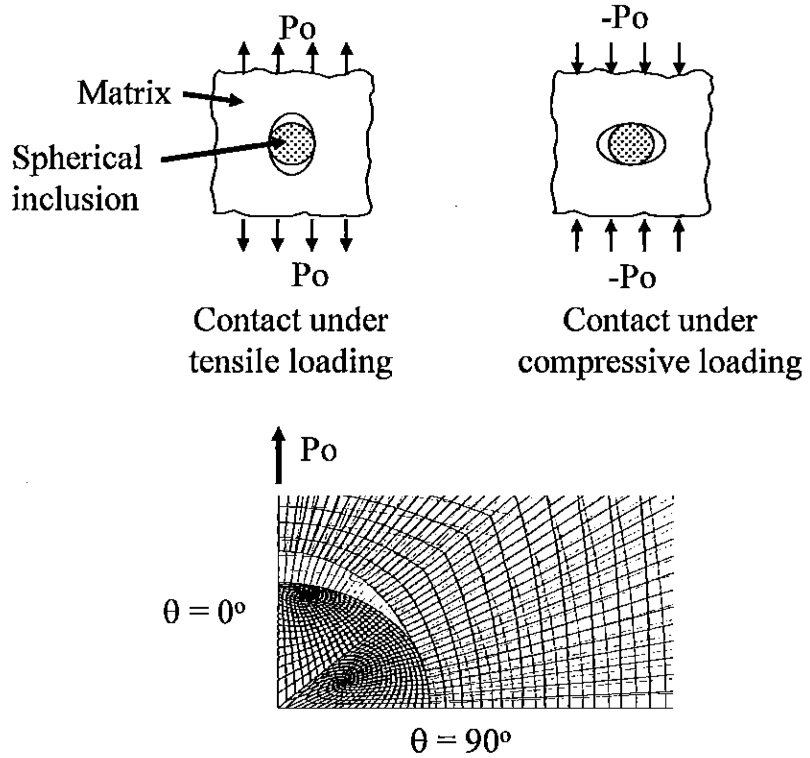




Magnified view of the contact area

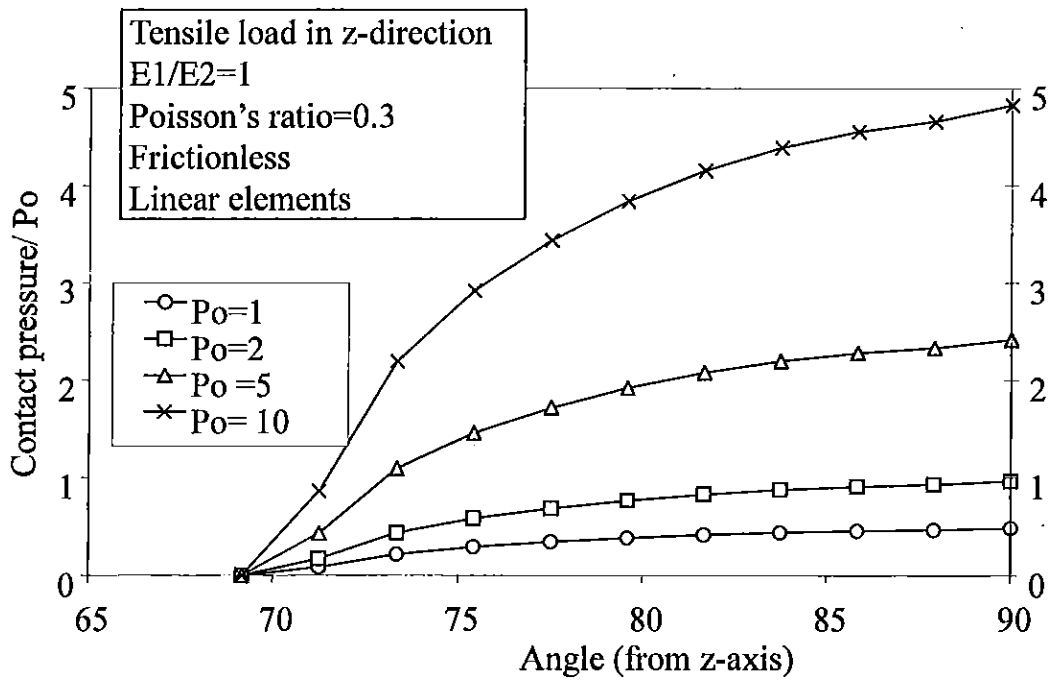
FE mesh for the inclusion in a matrix problem

(Inclusion in a Matrix Example/ Continued)

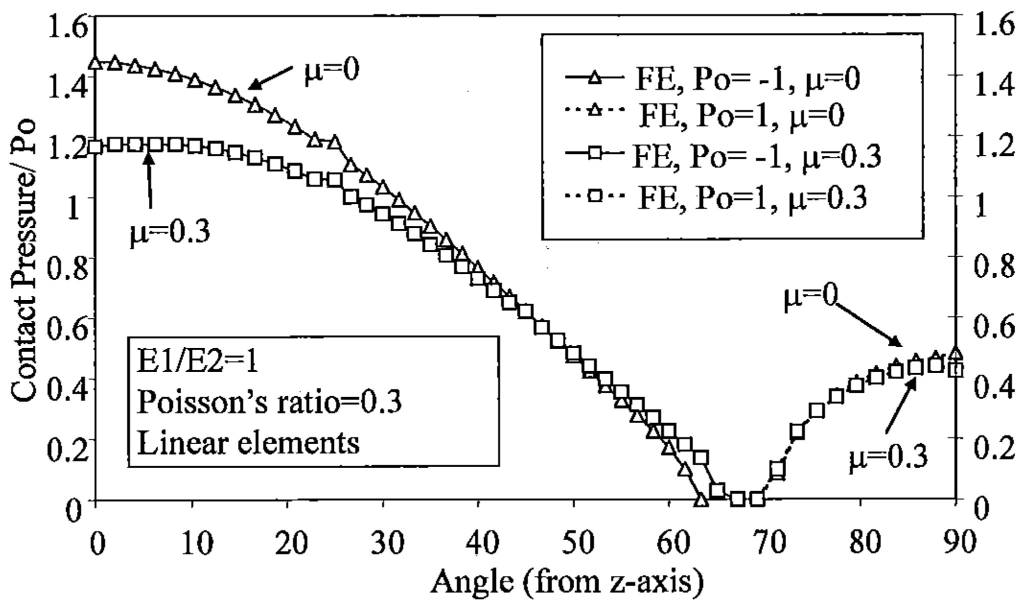


Deformed shape for the inclusion in a matrix problem

(Inclusion in a Matrix Example/ Continued)



Contact pressure (frictionless) for the inclusion in a matrix problem



Contact pressure (with friction) for the inclusion in a matrix problem

5.7 Summary of Key Points

- Contact problems can be classified according to the **surface roughness** (frictionless or frictional), **initial geometry** before contact (conforming or non-conforming) or **behaviour after contact** (stationary, advancing or receding).
- **Hertzian contact solutions** are limited to frictionless contact of curved bodies, but are very useful in arriving at approximate contact solutions.
- In Hertzian contact problems, the maximum contact pressure occurs at the centre of the contact area, and the **maximum shear stresses that cause yielding are beneath the contact surface**.
- In **frictional problems**, the application of normal and tangential loads may cause a stick-slip behaviour in the contact area.
- **Cyclic tangential loading** in frictional contact problems may induce a micro-slip situation which can result in fretting surface damage.

- In **receding contact** problems, the size and shape of the contact area is independent of the magnitude of the applied load, provided its direction remains the same.
- **Sharp edges and punches** usually cause a stress singularity (i.e. infinite or very high stress) at the corners.
- **Contact interface variables** have to satisfy equilibrium (equal and opposite forces), compatibility (no overlap) and slip condition (e.g. Coulomb friction law).
- Two main approaches are used in FE contact analysis; **Gap elements** inserted between the contacting nodes and **Lagrange Multipliers** which directly couple the contact equations.
- In **3D contact problems**, 20-node quadratic brick elements may result in negative nodal contact forces. Linear 8-node brick elements or 27-node elements are considered more suitable.

Part 6

Other Advanced FE Applications

Lecture Outline

6.1 Fracture Mechanics Problems

6.2 Fatigue Problems

6.3 Thermo-mechanical Problems

6.4 Viscoelastic Materials (Polymers, plastics, rubber)

6.5 Explicit FE Analysis

6.1 Fracture Mechanics Problems

Some Definitions

- **Fracture mechanics**
  The study of the strength of materials containing pre-existing flaws or cracks under the action of externally applied loads. Fracture mechanics is not concerned with crack nucleation.

- **Stress Intensity Factor (K)**
  A factor used to characterise the stresses at the tip of the crack which are assumed to be infinite.

- **Fracture Toughness (K_c)**
  This is a material property independent of the geometry or applied loads. As K reaches a critical value (known as K_c), catastrophic failure (i.e. fast fracture) occurs. This K_c value is referred to as the "Fracture Toughness" of the material.

Linear Elastic and Post-Yield Fracture Mechanics

- **LEFM (Linear Elastic Fracture Mechanics)**

  LEFM is usually used for brittle fracture. Assumptions of LEFM:

  (i) Stresses are infinite at the crack tip of the order $1/r^{0.5}$ where r is the distance from the crack tip.

  (ii) The size of the plastic zone is small compared to other crack dimensions

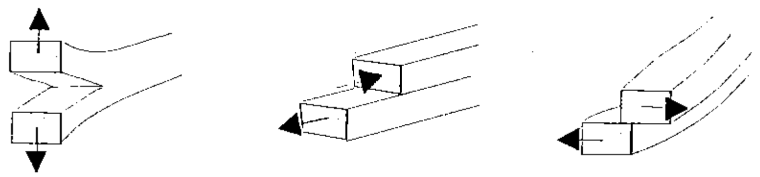
  (iii) The crack is situated in a homogenous material.

- **PYFM (Post-Yield Fracture Mechanics)**

  PYFM covers fracture theories that take into consideration the plastic zone around the crack tip. PYFM is used for ductile fracture.

Three Modes of Fracture

- Three modes of fracture:
  Mode I : Opening mode
  Mode II : Sliding mode
  Mode III : Tearing mode
- The three modes of fracture are usually treated separately.
- Mixed mode fracture involves the interaction of more than one fracture mode.
- Analysis is usually concerned with Mode I which is considered the dominant mode. Fracture toughness is usually concerned with Mode I fracture.



Mode I (Opening) **Mode II (Sliding)** **Mode III (Tearing)**

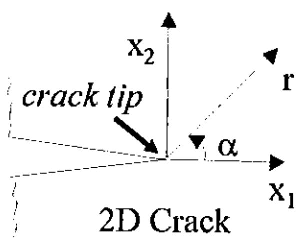
The three modes of Fracture

Displacements and Stresses at the Crack tip

$$\textit{Stress at the crack tip} = \frac{1}{\sqrt{r}}\, f(\alpha, K)$$

$$\textit{Displacement at the crack tip} = \sqrt{r}\ f(\alpha, K)$$

Local direction 1 = Tangential direction
Local direction 2 = Normal direction
Local direction 3 = Normal to the x-y plane



Crack tip geometry

(a) 2D Problems

The following expressions exclude terms with higher order of r (where r is the distance from the crack tip), since r is considered very small compared to other dimensions of the crack.

$$\sigma_{11}=\frac{1}{\sqrt{2\pi r}}\left[K_I\cos\frac{\alpha}{2}\left(1-\sin\frac{\alpha}{2}\sin\frac{3\alpha}{2}\right)+K_{II}\sin\frac{\alpha}{2}\left(-2-\cos\frac{\alpha}{2}\cos\frac{3\alpha}{2}\right)\right]$$

$$\sigma_{22}=\frac{1}{\sqrt{2\pi r}}\left[K_I\cos\frac{\alpha}{2}\left(1+\sin\frac{\alpha}{2}\sin\frac{3\alpha}{2}\right)+K_{II}\sin\frac{\alpha}{2}\cos\frac{\alpha}{2}\cos\frac{3\alpha}{2}\right]$$

$$\sigma_{33}=\frac{3-\nu-\kappa(1+\nu)}{4\nu}(\sigma_{11}+\sigma_{22})$$

$$\sigma_{12}=\frac{1}{\sqrt{2\pi r}}\left[K_I\sin\frac{\alpha}{2}\cos\frac{\alpha}{2}\cos\frac{3\alpha}{2}+K_{II}\cos\frac{3\alpha}{2}\left(1-\sin\frac{\alpha}{2}\sin\frac{3\alpha}{2}\right)\right]$$

$$\sigma_{23}=\frac{1}{\sqrt{2\pi r}}K_{III}\cos\frac{\alpha}{2}$$

$$\sigma_{13}=\frac{-1}{\sqrt{2\pi r}}K_{III}\sin\frac{\alpha}{2}$$

$$u_1=\frac{\sqrt{2\pi r}}{8\pi\mu}\left[K_I\left[(2\kappa-1)\cos\frac{\alpha}{2}-\cos\frac{3\alpha}{2}\right]+K_{II}\left[(2\kappa+3)\sin\frac{\alpha}{2}+\sin\frac{3\alpha}{2}\right]\right]$$

$$u_2=\frac{\sqrt{2\pi r}}{8\pi\mu}\left[K_I\left[(2\kappa+1)\sin\frac{\alpha}{2}-\sin\frac{3\alpha}{2}\right]+K_{II}\left[(3-2\kappa)\cos\frac{\alpha}{2}-\cos\frac{3\alpha}{2}\right]\right]$$

$$u_3=\frac{2\sqrt{2\pi r}}{\pi\mu}K_{III}\sin\frac{\alpha}{2}$$

where α is the angle measured anticlockwise from the plane of expected crack growth.

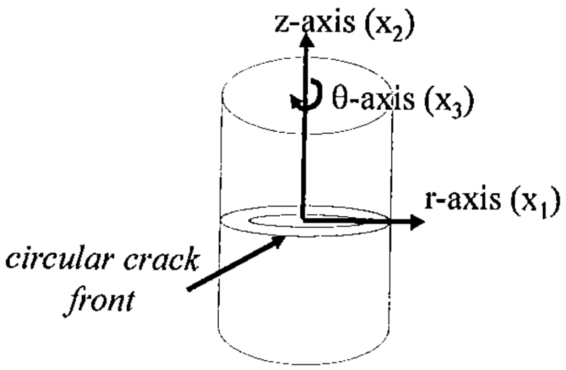
The parameter κ is defined as follows.

$$\kappa=3-4\nu \quad \textit{(plane strain)}$$

$$\kappa=\frac{3-\nu}{1+\nu} \quad \textit{(plane stress)}$$

(b) Axisymmetric problems

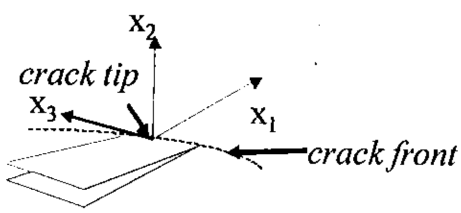
- The behaviour in the vicinity of axisymmetric (circular) cracks is usually assumed to be the same as that of 2D plane strain.
- Typical axisymmetric cracks include circumferential cracks on the outer surface of a cylinder or an embedded "penny-shaped" crack.



Axisymmetric Crack Geometry

(c) 3D problems

- The normal 2D definition of K does not hold because the crack tip is no longer a point but a line or a curve.
- In practice, it is assumed that most 3D cracks embedded far away from the external surfaces may be treated as plane strain cracks. The region where a 3D crack cuts a free surface is usually assumed to be a plane stress crack region.
- For 3D cracks, K will generally vary from point to point on the crack front. Hence, K must be calculated for all points on the crack front to determine if any K value exceeds the critical value , K_c.



3D Crack Tip Geometry

Strain Energy Release Rate

- The strain energy release rate , G, is another parameter that can be used to characterise crack behaviour. It is defined as the rate of change of strain energy for an incremental change in crack length, ∂a, as follows.

$$G = \frac{\partial U}{\partial a}$$

  where U is the strain energy stored due to the application of the external forces.

- Relationships between G and K:

$$G_I = \frac{\kappa + 1}{8\mu}(K_I)^2 \qquad G_{II} = \frac{\kappa + 1}{8\mu}(K_{II})^2 \qquad G_{III} = \frac{1}{2\mu}(K_{III})^2$$

  where G_I, G_{II} and G_{III} are the strain energy release rates for modes I, II and III, respectively.

- As K approaches K_c, rapid (or catastrophic) fracture occurs. Similarly, a critical value of the strain energy rate, G_c, can be used.

Methods for Calculating SIF

Analytical solutions

These are usually limited to relatively simple geometries and loading conditions, and require a good knowledge of sophisticated mathematics. They include:

- Complex stress functions
- Boundary collocation
- Conformal mapping
- Green's function
- Integral transforms

Experimental Methods

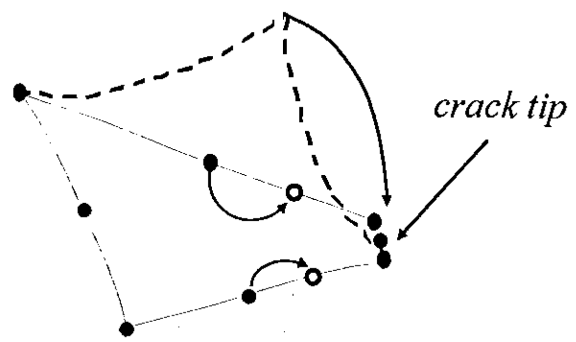
Experimental procedures are essential in determining the critical value of K_I (by testing to destruction) , but are very costly and time consuming in determining K for a range of crack shapes and depths. They include :

- Measuring compliance for several crack lengths
- Photoelasticity
- Interferometry

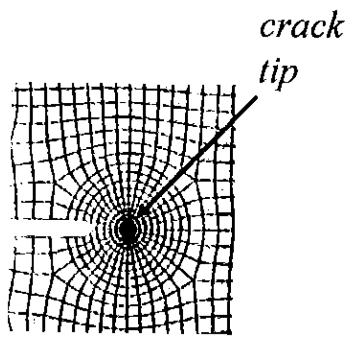
Crack-tip Singularity Elements (Collapsed elements)

- Since the **stresses are infinite** at the crack tip, even using a very fine mesh of higher-order elements will not produce sufficient accuracy.
- Because the nature of the singularity of the crack tip is known (usually of the order of $1/r^{0.5}$), it is possible to develop special crack-tip elements that have a **built-in singularity**.
- Best elements to use are **isoparametric quadratic elements**, e.g.
  For 2D problems : 6-node triangles or 8-node quadrilateral elements
  For 3D : 15-node wedge elements or 20-node brick elements
- Usually only a **ring of special crack tip elements** (or a tube of elements in 3D cracks) are used around the crack tip.
- To create the singularity at the crack tip, **one side of the quadratic element is "collapsed"** such all three nodes on that side coincide. In LEFM the collapsed nodes are usually "tied" together for better accuracy.

(Crack-tip Singularity Elements / Continued)



2D Collapsed crack-tip element

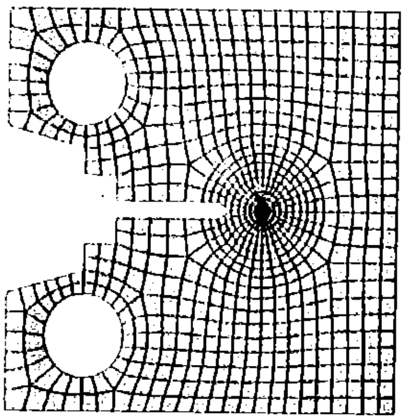


Typical FE mesh showing a ring of collaped elements around the crack tip

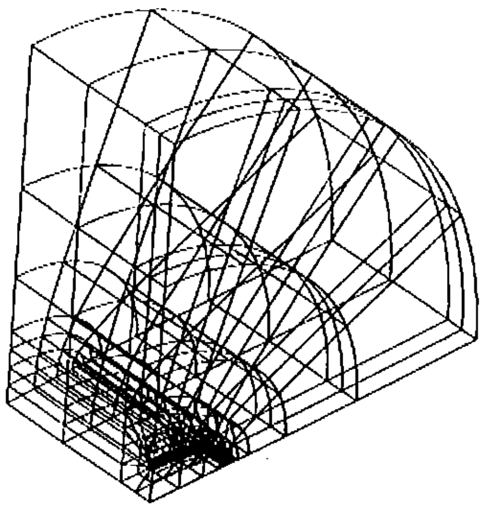
Collapsed crack-tip elements

Notes on Collapsed Elements

- The collapsed elements **can be used to model elastic-plastic analysis** because they create a singularity similar to that in an elastic-perfectly-plastic material.
- In PYFM, the collapsed nodes can be left "untied", which simulates the **"blunting" of the crack tip in PYFM**.
- The crack tip singularity is **not always of the order of $1/r^{0.5}$**. PYFM and interface cracks between similar materials have different singularities.
- For **3D cracks**, singularity is different at or near the free surface.
- In some FE codes, **"variable singularity" crack tip elements** may be offered, where the position of the shifted nodes can be adjusted to suit the problem. However, these elements are not commonly available in FE codes.



2D FE mesh used to model fracture in a compact tension specimen



3D FE mesh used to model a semi-elliptical crack in a pipe

Examples of FE Crack Tip Meshes

Numerical Methods for Calculating SIF

(a) Displacement and Stress Substitution Methods

- The classical crack tip displacement and streess equations are used to obtain expressions for K for a number of points near the crack tip.
- The points must be located within the region of the crack dominance, i.e. not very far away from the tip.
- The values of K can be extrapolated to the crack tip.
- **Disadvantage:** The displacement and stress values are not very accurate very close to the crack tip.

(b) Energy Difference Method

- The differential of the strain energy with respect to the crack length can be expressed as a simple difference formula as follows:

$$G_I = \frac{\Delta U}{\Delta a} = \frac{U_2 - U_2}{a_2 - a_1}$$

where U_1 and U_2 are the strain energies associated with the crack lengths a_1 and a_2. This approximation is only valid if the **difference between the two crack lengths is very small.**

- The simplest approach is to **perform two FE runs** with about 1-2% change in the crack length. This requires no modification of the software, but care must be taken to ensure that the increment in the crack length is small enough for accuracy but not too small to cause numerical inaccuracies.
- **Disadvantage:** Uncertainty in choosing a suitable size of crack increment to be used. Furthermore, it is inefficient because two or more FE runs are necessary.

(c) J-Contour Integral

The J-contour integral [Rice, 1968] is a two-dimensional path independent integral defined as follows:

$$J = \int_{\Gamma^*} \left(W \, dx_2 - t_i \frac{\partial u_i}{\partial x_1} dS^* \right)$$

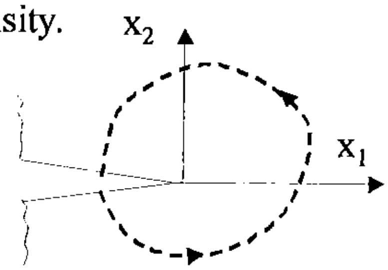
where

Γ^* is a path from one surface of the crack to the other inside the solution domain in the x1-x2 plane,

dS^* is the differential distance on this path,

t_i is the traction vector

W is the strain energy density.



J-Contour integral

Applications of the J-Contour Integral

- The J-integral is widely used in commercial FE codes.
- For LEFM, the J-integral is equal to the strain energy release rate G.
- A modified J-contour integral can used for PYFM.
- It is advisable to use paths away from the crack tip since values of displacements, strains and stresses are usually are more accurate.
- In practice, the average value of J over a number of contours is used.

Limitations of the J-integral method

(i) Axisymmetric problems: The J-integral is not path independent except for regions very close to the crack tip. This is because the behaviour of stresses near axisymmetric crack tips is similar to that in 2D plane strain conditions.

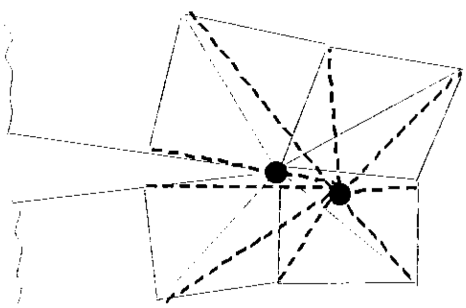
(ii) 3D problems: The J-integral is determined on a 2D plane perpendicular to the crack front, i.e. over a tubular surface around the crack tip.

(iii) Thermal Problems: The J-integral is not path independent. A modified form of the J contour is required.

(iv) PYFM: The J-integral is not path independent. An extension to the J contour is needed to keep it path independent by excluding a small circular area of radius r around the crack tip.

(d) VCE (Virtual Crack Extension)

- The change in strain energy, ΔU, for a virtual crack extension of the crack tip, Δa. can be calculated without the need to perform two FE runs.
- A selected ring or a layer of rings of elements surrounding the crack tip can be used to calculate the change in energy.
- These methods require extra coding in the FE software.



Virtual crack extension method

(Numerical Methods for Calculating SIF/ Continued)

Different Approaches for VCE (Virtual Crack Extension)

(i) **Stiffness derivative method** [Parks, 1974)]
This method uses a derivative of the stiffness matrix with respect to the crack length, a, over several layers of elements surrounding the crack tip

(ii) **Strain energy accumulation** [Hellen, 1975]
This method calculates the strain energy over only the crack tip elements.

(iii) **Modified Crack Closure Method** [Rybicki and Kanninen, 1977]
This method calculates the work required to close a small increment in of the crack length

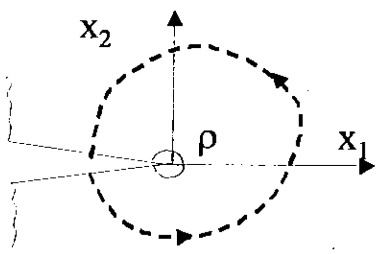
Methods for Calculating SIF in PYFM

(i) Modified J-Integral for PYFM

An extra term is needed to make the J-integral path independent in PYFM, as follows [Hellen, 1987].

$$J = \int_{\Gamma^*} \left(W\, dx_2 - t_i \frac{\partial u_i}{\partial x_1} dS^* \right) +$$

$$\lim_{\rho \to 0} \iint_A \left[\frac{\partial W}{\partial x_1} - \frac{\partial}{\partial x_i}\left(\sigma_{ij} \frac{\partial u_j}{\partial x_1} \right) + \frac{\partial}{\partial x_3}\left(\sigma_{i3} \frac{\partial u_i}{\partial x_1} \right) \right] dA$$



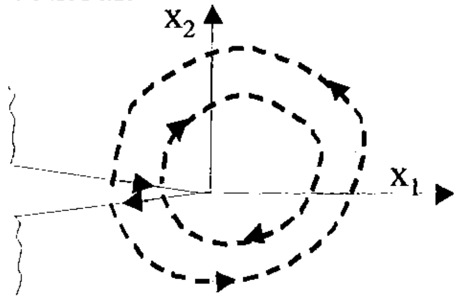
Modified J-Contour integral for PYFM

(ii) Domain Integrals for PYFM

The integration is performed over a thick band of the domain (Li, Shih and Needleman [1985])

$$J_{domain} = \int_{A^*} \left(-\sigma_{ij} \frac{\partial u_i}{\partial x_1} - W \delta_{i_j} \right) \frac{\partial q_1}{\partial x_j} \, dA^*$$

δ_{ij} is the Kronecker Delta (equal to 1 if $i = j$, and zero otherwise)
A is the area of the band of domain
q is a function that equals 1 on the inside of the contour and zero on the outside of the contour.



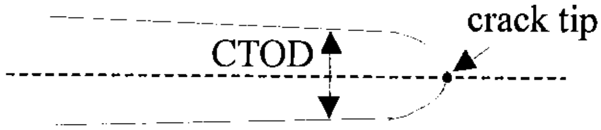
Domain integral for PYFM

(iii) CTOD (Crack Tip Opening Displacement) for PYFM

- The CTOD can be determined by extrapolating the crack opening displacement away from the crack tip to obtain the CTOD at or near the crack tip.

- In some FE analyses, CTOD is measured by taking the value of the opening displacement at the intersection of the crack front and a plane at 45° to the crack plane.

- CTOD can be measured at the crack opening using experimental techniques.



Crack Tip Opening Displacement for PYFM

Mixed Mode Fracture

The separation of mode I and mode II stress intensity factors can be obtained from the normal and tangential displacement components, respectively.

$$K_{II} = \frac{E\sqrt{2\pi}}{4\left(1-\nu^2\right)}\left[\frac{u_t}{\sqrt{r}}\right]_{r\to 0}$$

$$K_{II} = \sqrt{2\pi}\left[\tau_{xy}\sqrt{r}\right]_{r\to 0}$$

$$G_{total} = \frac{(K_I)^2 + (K_{II})^2}{E} \qquad (plane\ stress)$$

$$G_{total} = \frac{(K_I)^2 + (K_{II})^2}{\left(\dfrac{E}{1-\nu^2}\right)} \qquad (plane\ strain)$$

Fatigue Crack Growth

- In LEFM is it assumed that cyclic loading causes **an incremental increase in the crack size.**
- FE analyses can be used to calculate the **maximum and minimum values of K** associated with the maximum and minimum values of the applied load in a constant amplitude fatigue cycle.
- The difference between the values of K is the **range of the stress intensity factor, ΔK.**
- The **Paris Law** is often used to calculate the rate of change of crack length per cycle, *da/dN*, as follows:

$$\frac{da}{dN} = C(\Delta K)^n$$

where *C* and *n* are material constants, and D*K* is stress intensity factor range over each cycle.

Creep Crack Growth

- The strain energy release rate can be used to calculate K in creep problems, provided that the **creep strains are included** in the strain calculations.

- A **creep fracture parameter, C^*,** can be calculated as the rate of change of the J-integral with time, as follows:

$$C^* = \int_{\Gamma^*} \left(\frac{dW}{dt} dx_2 - t_i \frac{d}{dt} \left[\frac{\partial u_i}{\partial x_1} \right] dS^* \right)$$

- Alternatively, the C^* parameter can be calculated from the change in strain energy as follows:

$$C^* = \frac{d}{dt} \left(\frac{dU}{da} \right)$$

Summary of Fracture Mechanics

- Several techniques for the **calculation of the stress intensity factor (SIF)** exist, such as stress and displacement substitution and energy techniques.

- Special **crack tip singularity elements** can be used to produce very accurate results without a very fine degree of mesh refinement, and are widely used in FE software.

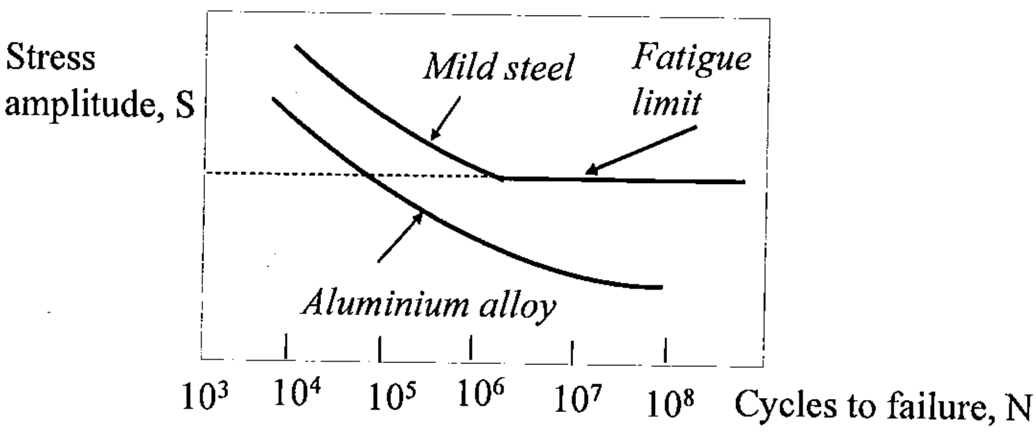
- Care must be taken in analysing non-linear fracture problems, such as **PYFM and creep fracture**, as the parameters used in LEFM may not be appropriate.

- Some FE codes contain a number of **automated procedures** for calculating the Stress Intensity Factors. However, the user must check the FE solutions carefully.

- The user must gain **a good understanding of fracture mechanics** before attempting to analyse a cracked geometry using FE.
- It is recommended that the user should run a number of **fracture benchmarks** in order to verify the accuracy of the FE code, and to ensure that the FE code has been used correctly.

6.2 Fatigue Problems

- Fatigue occurs when a structure is subjected to **a repeated or varying load** which is below the level that would cause failure in a single load application.
- Fatigue strength is measured by testing a number of identical uniaxial specimens under different loading cycles until they fail after N Cycles under a given stress amplitude S. The resulting curves are called **the S-N curves**, as shown below.



Typical S-N Curves

Fatigue Limit

- In some metals, there is flattening of the curve at lower stress levels, and a **"fatigue limit"** (a safe working stress) is reached below which fatigue failure would not occur.
- In real life applications, however, the fatigue limit would depend on the **mean stress** and other factors such as surface finish and corrosion).
- Fatigue failure originates from **small cracks** that initiate in weaker grains or grain boundaries in metals under repeated stress cycles.

Three Types of Fatigue Analysis

(a) **<u>High-cycle fatigue (Stress-Life Approach)</u>**
This is usually associated with relatively low loads and long life (approximately over 10,000 cycles). Such problems are commonly analyzed with "Stress-Life" or "Total-Life" methods, which predict the number of cycles sustained before failure.

(b) **<u>Low-cycle fatigue (Strain-Life Approach)</u>**
Here significant plastic deformation occurs, which produces a relatively short life (less than 10,000 cycles). Such problems are usually analysed with "Strain-Life" or "Crack-Initiation" methods.

(c) **<u>Crack growth</u>**
Here the propagation of a crack is predicted using linear elastic fracture mechanics (LEFM) or elasto-plastic fracture mechanics (EPFM). The rate of crack growth per cycle can be predicted

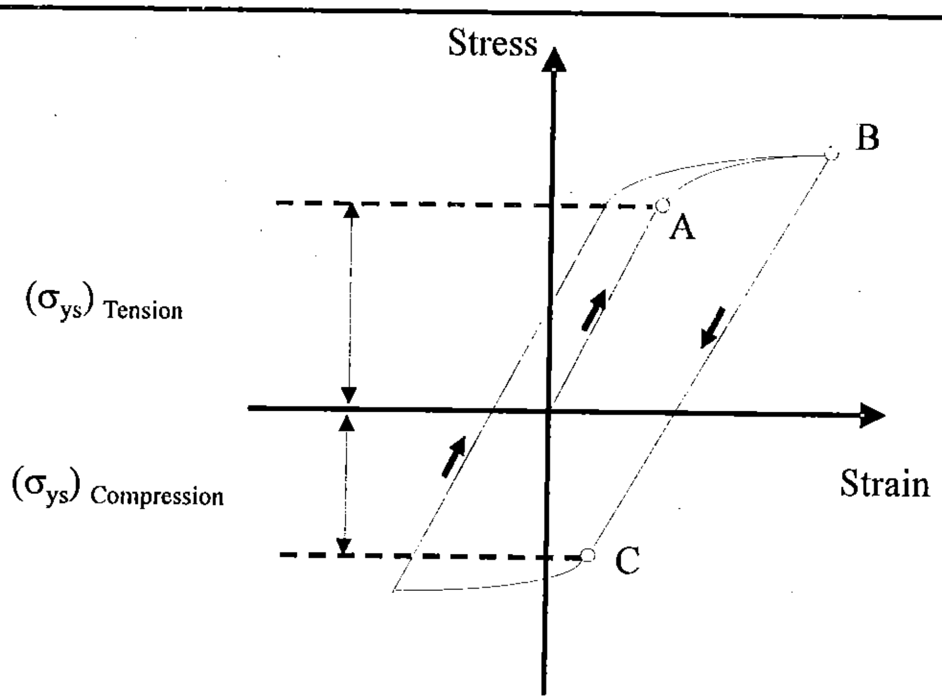
The Bauschinger Effect

- When some materials are subjected to uniaxial tension beyond yield, then unloaded and reloaded in compression, it is found that the yield stress in compression is less than the equivalent value in tension. This effect is called the ***Bauschinger effect*** and occurs because of the permanent strains and residual stresses remaining after the first yield point is reached. These residual stresses add to the reversed stresses in compression loading thus lowering the second yield point.

- A typical **'hysteresis loop'** forms when reversed loading is applied to a metal, where the Bauschinger effect is exhibited by the fact that $(\sigma_{ys})_{\text{compression}}$ is less than $(\sigma_{ys})_{\text{tension}}$.

- In **isotropic hardening** it is assumed that the yield point and the effects of work hardening are the same in tension and compression (i.e. no Bauschinger effect), whereas **in kinematic hardening** the yield point in compression is lower.

(The Bauschinger Effect/ Continued)



Hysteresis loop

Fatigue Analysis Using FEA

- FE analysis of fatigue problems is usually performed by incorporating the computed FE stresses into a further **post-processing operation** to estimate the fatigue life. The effect of each applied load can be **summed at every node** to derive a total stress or strain response history.

- Modern FE fatigue software may incorporate a **database of fatigue properties,** which can be used to predict the location of the **fatigue "hotspots",** i.e. the locations where fatigue cracks may initiate. FE solutions may be presented as colour contour plots of fatigue lives and/or probabilities of failure.

- Since peak stresses are important in fatigue life estimation, it is good practice to look at **un-averaged stress values at the Gauss integration points**, rather than nodal averaged stresses (which are usually averaged over a group of adjacent elements).

- In **fatigue fracture simulations**, special crack tip elements are used at the crack tip to simulate the stress singularity, i.e. the theoretically infinite stress, at the crack tip. A fine well-graded mesh should be used around the crack.

- The **plane of maximum principal stress** may be used to estimate the **direction of crack growth**. Some FE codes automatically re-mesh the new position of the propagating crack tip and continue the analysis.

- Nonlinear **plasticity FE analysis with kinematic hardening** is required to simulate cyclic loading.

6.3 Thermo-mechanical Problems

- FE analysis can be used to calculate the temperatures in the structure, using a thermal (heat conduction) analysis, and then feed the computed temperatures into the structural analysis to calculate thermal stresses and strains. Thermal strain is defined as follows:

$$\varepsilon_{thermal} = \alpha\left(\Delta T\right)$$

$\varepsilon_{thermal}$ is the thermal strain
α is the coefficient of thermal expansion (Units: $^{o}C^{-1}$)
ΔT is the change in temperature from a datum reference temperature (Units: ^{o}C)

- In most applications, the thermal strain can be usually superimposed on the mechanical strain.

- Thermal and mechanical analyses are either **uncoupled or fully coupled.**

(Thermo-mechanical Problems/ Continued)

(a) Uncoupled (or sequentially coupled) thermo-mechanical problems

- In uncoupled analysis, it is assumed that the **deformations of the body do not affect the temperature gradients** inside the body. Therefore, the thermal stresses can be simply obtained from the temperatures.

- The FE analysis of uncoupled themo-mechanical problems is performed in **two sequential stages.**

- In the **first stage**, the heat conduction analysis is performed to compute all the temperatures inside the body.

- In the **second stage**, the temperatures are treated as input to the structural analysis as part of the loading.

(b) Fully Coupled thermo-mechanical problems

- Fully coupled analysis is needed when the thermal and mechanical solutions are strongly dependent on each other, i.e. the **deformations of the body directly affect the temperature distribution** inside the body.

- For example, the **frictional contact of a brake disc pad** generates temperatures that are dependent on the amount of sliding.

- Another example is the **shrinking of a casting as it cools and separates from the mould**, thus affecting the heat transfer between the casting and the mould. In addition, significant heating of the casting may affect the material properties.

- Fully coupled thermo-mechanical requires the use of **both temperatures and displacements as degrees of freedom**, and is usually non-linear.

6.4 Viscoelastic Materials

(Polymers, Plastics, Rubber)

- Materials such as polymers exhibit a rate and time dependency. Such behaviour is usually referred to as **'viscoelastic' or 'creep' behaviour.**

- The term **'creep'** is usually used to describe the behaviour of metals at very high temperatures, whereas **'viscoelasticity'** is often used to describe the behaviour of glass and plastics. However, both terms are frequently used to describe any material with time and rate dependent behaviour.

- The mechanical properties of viscoelastic materials are **dependent on temperature, time and frequency,** with some material properties presented as 'complex numbers' (with a real and an imaginary number) in the frequency domain.

Effect of Temperature on Viscoelatic Materials

- The behaviour of a viscoelastic material as a 'rubber' or a 'plastic' depends on the temperature range of its application.
- A viscoelastic material which can be classified as rubber at room temperature can be made to behave as a glassy or plastic material at a very low temperature of -40 °C, whereas a viscoelastic material which can be classified as plastic at room temperature, can be made to behave as rubber at temperatures above 150 °C.
- The effect of temperature on the behaviour of viscoelastic materials can be divided into three distinct regions: Glassy, Transition and Rubber regions.

Glassy, Transition and Rubber regions

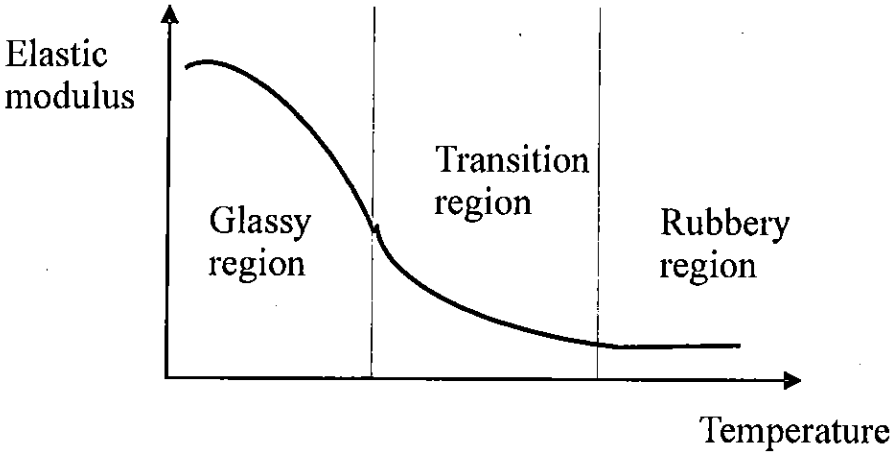
(i) **Glassy region** in which the motion of the long polymer chains of the material is highly restricted, and the elastic modulus is highest.

(ii) **Transition region** in which the highest rate of molecular motion occurs

(iii) **Rubber region** in which there is only a slight increase in molecular mobility at high temperatures.

(Glassy, Transition and Rubber regions/ Continued)

- The variation of the **elastic modulus** can be represented schematically as shown below.
- A similar variation with temperature occurs for the **shear and bulk moduli.**



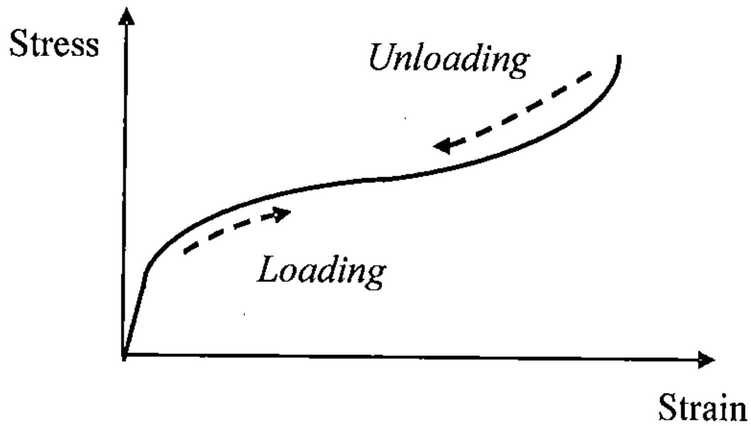
Schematic representation of the variation of elastic modulus with temperature

Ruhbers and Elastomers

- Rubber and elastomer materials are referred to as **'hyperelastic'** because the strains can reach very large values, e.g. over 400%, but still remain elastic.
- The deformation is called **'incompressible'**, i.e. there is no change of volume during deformation, because the value of Poisson's ratio is equal or nearly equal to the limiting value of 0.5.
- Although hyperelasticity is a non-linear behaviour, the body often returns to its original state after unloading, i.e. a reversible behaviour, and the **deformations are independent of the loading path.**

Stress-Strain Curves for Rubbers and Elastomers

- The stress-strain curve for rubbers and elastomers is often a **continuous curve** which can be represented by a single mathematical expression.
- The material model is often expressed as a **function of the strain energy**, i.e. the area under the stress-strain curve.



Hyperelastic material behaviour

Typical of Viscoelastic Materials

Typical applications in which viscoelastic materials are used, include the following:

- Door seals in cars and domestic appliances.
- Body panels for cars
- Engine mounts
- Diaphragms
- Rubber isolators and shock absorbers
- Mechanics shaft seals

FE Analysis of Viscoelastic materials

(a) Viscoelastic material properties

- Since viscoelastic materials require a complex set of material constants, it is important to use accurate and reliable sets of material properties in the FE analysis.

- Material handbooks do not usually contain the full range of viscoelastic material properties needed in FE analysis (frequency, temperature and time variation of constants).

- **Experimental techniques** are usually needed to obtain the viscoelastic material constants. These usually include uniaxial tension, uniaxial compression, equiaxial loading, shear loading and frequency tests.

(b) Hybrid Elements for incompressible materials

- In problems involving incompressible or nearly-incompressible materials, such as rubber, the conventional (displacement-based) FE formulation **fails or becomes inaccurate when Poisson's ratio is nearly equal to 0.5.**

- In such applications, **hybrid elements** are used where the hydrostatic stress is defined as an additional degree of freedom in the element formulation.

(c) Frequency domain analysis for dynamic viscoelastic problems

- In free vibration analysis of viscoelastic materials, a frequency domain FE analysis can be used to compute the mode shapes and natural frequencies.

- This is suitable mainly for linear problems and for viscoelastic materials whose material constants are independent of frequency. Only the real part of the complex modulus definition is used.

Modal superposition and direct integration methods

- For forced vibration analysis, either a modal superposition method or a direct integration method can be used.
- In the modal superposition method, the frequency response functions are obtained from the modal deformations, which results in a fast run time.
- The direct integration method uses a time-stepping scheme to approximate the solution over each time step, and is suitable for both linear and non-linear dynamic problems, and for viscoelastic materials with frequency-dependent properties and which have a complex function to describe the constants. However, this results in run times which are considerably longer than the modal superposition method.

6.5 Explicit FE Analysis

- Explicit FE codes are popular in modelling **dynamics or transient problems** in which material, geometric or contact non-linearities are included in the analysis.

- **Typical problems** include impact, crash worthiness of vehicles, earthquake analysis and metal forming processes.

- In **time integration** or **time-marching** schemes, the total time is divided into small time steps, and integration is performed over each time step to predict the solution for the next time step.

- The term **'explicit'** is used to describe the time integration scheme where the new values of displacements (and other variables) are evaluated only from the values at the start of the new time step.

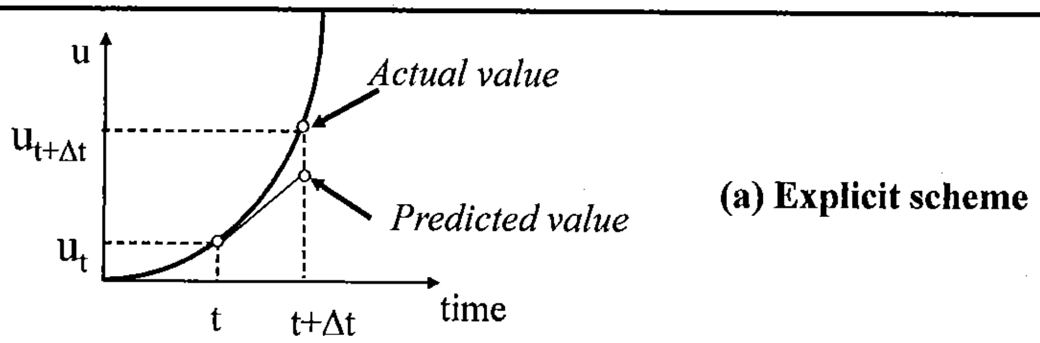
Explicit vs. Implicit Methods for Time Integration

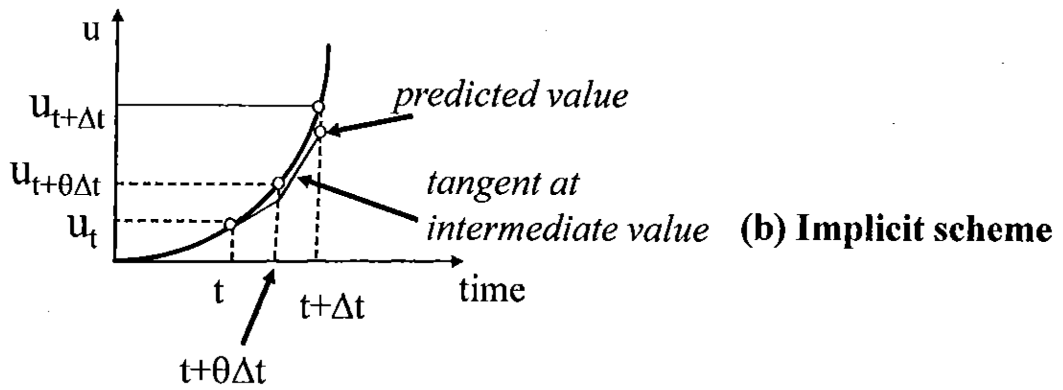
<u>Implicit time integration</u>

- Traditionally used in most FE codes.
- More complex than the explicit scheme
- Size of the time step is unrestricted
- Unconditionally stable
- **Disadvantage :** Global equilibrium must be checked, and a new stiffness matrix must be computed before checking that equilibrium is satisfied. Furthermore, if equilibrium is not satisfied within a given increment or time step, the implicit scheme requires that a new time step is chosen, and a new stiffness matrix computed.

<u>Explicit time integration</u>

- Simpler to implement
- Must use small time steps
- Does not require frequent equilibrium convergence checks
- **Disadvantage:** If the time step is too large, instability can occur leading to large errors





Schematic representation of explicit and implicit time integration schemes

Problems suitable for Explicit FE analysis

- **Dynamic high velocity contact/impact** problems where the load is applied over a very short time, e.g. a fraction of a second. In such problems, the solution can be propagated through one element at a time as a stress wave.
- **Vehicle crashworthiness** analysis, which involves contact of many components and joints and very large meshes
- **Metal forming** processes in which severe plastic strains occur with changing contact and friction conditions, often requiring adaptive meshing algorithms
- **Missile impact** problems where the loading duration is extremely small (e.g. less than 1 ms)
- **Rigid body motion** where the bodies are either unconstrained or experience little or no strains during motion
- **Buckling, post-buckling and snap-through** problems, where the material experiences a sudden loss of stiffness

Note: **Implicit analysis is more efficient for linear problems** where the loads are slowly varying and dynamic effects are small. This is because large time steps can be tolerated, and equilibrium checks are not frequently needed.

Computational issues in Explicit FE analysis

(i) Restriction on the size of the time step in Explicit FE analysis

- In explicit FE analysis, the size of the time step must be kept small. If the time step is too large, significant errors may occur.
- In the central-difference explicit scheme, an automatic time scheme is usually used in FE codes. A restriction is imposed on the size of the time step, as follows:

$$(\Delta t)_{critical} \leq \frac{2\,\rho\, c_p}{D\,\pi^2\, k} L^2$$

k_z is the thermal conductivity
ρ is the density of the solid (units kg m^{-3})
c_p is the specific heat of the solid (units: J kg^{-1} K^{-1})
L is "characteristic length" usually determined as the minimum distance between any two nodes in the FE mesh
D is either 1,2 or 3 corresponding to one-dimensional, two-dimensional or three-dimensional problems, respectively

(ii) Round-off errors

- Computers store real number sin a fixed number of 'bits', i.e. numbers are truncated after a fixed number of digits after the decimal point, which results in "round-off" errors. In most FE analyses, round-off error is too small to be significant.
- However, in some applications, the units used for the material constants may give rise to variables that may suffer from large round-off errors. This can happen in an explicit analysis, where, for example, a variable may become so small that its truncated value may appear as unchanged during an iteration or a time step.

(iii) Single- and double-precision

- Single-precision is commonly used to describe a word length of 32 bits, whereas double- precision is usually used to describe a word length of 64 bits. In some FE codes, the user may be allowed to select either single- or double-precision.

- In a highly non-linear analysis which involves a very large mesh, the choice of double-precision may significantly increase the memory requirements and slow down the run times.

(iv) Adaptive meshing

- **Adaptive meshing** should be used in explicit FE analysis of problems where severe changes in strain or element deformation occur that would render the original mesh unsuitable.

- **Sub-modelling** can also be used to map the results from an initial coarse mesh onto a selected region of very fine elements.

(v) Storage Requirement

The memory storage requirements for explicit schemes are proportional to N where N is the number of unknowns in the structure. For implicit schemes, the storage requirements are proportional to N^2 or N^3, which, for large FE meshes, can substantially increase the computational time.

(vi) Units of Material constants

A` consistent set of units must be used in explicit analysis, particularly if the numerical values of material constants are extremely small or extremely large. For example, it may be appropriate to use micro-seconds, rather than hours, as units of time in high velocity impact problems.

REFERENCES / FURTHER READING

Allen, H.G. and Buslon, P.S. [1980] *Background to Buckling*, McGraw-Hill, London.

Argyris, J.H. and Symeonidis, S. [1981], Non-linear finite element analysis of elastic systems under non-conservative loading – Natural formulation, Part 1 Quasi-static problems, *Computer Methods in Applied Mechanics and Engineering*, Vol. 26, pp. 75-123.

Bathe, K.J. [1982] *Finite element procedures in engineering analysis*, Prentice-Hall, New Jersey.

Becker, A.A. [1992], *The Boundary Element Method in Engineering*, McGraw-Hill, London.

Becker, A.A. [2001] *Understanding Non-linear Finite Element Analysis*, NAFEMS, Glasgow.

Becker, A.A., *An Introductory Guide to Finite Element Analysis*, Wiley, London (2004).

Becker, A.A. and Hyde, T.H. [1993] *Fundamental tests of creep behaviour*, NAFEMS Report, R0027.

Belytschko, T., Liu, W.K. and Moran, B. [2000], *Nonlinear Finite Elements for Continua and Structures*, Wiley, New York.

Bittnar, Z. and Sejnoha, J. [1996] *Numerical methods in structural mechanics*, ASCE Press, New York.

Blazynski, T.Z. [1983] *Applied elasto-plasticity of solids*, Macmillan Press, London.

Bonet, J. and Wood, R.D. [1997] *Non-linear continuum mechanics for finite element analysis*, Cambridge University Press.

Calladine, C.R. [1983], *Theory of shell structures*, Cambridge University Press, Cambridge.

Crisfield, M.A. [1980] Incremental/iterative solution procedures for non-linear structural analysis, in *Numerical Methods for Non-linear Problems*, edited by C. Taylor, E. Hinton, D.R.J. Owen and E. Onate, pp. 261-290, Pineridge Press, Swansea.

Crisfield, M.A. [1991] *Non-linear finite element analysis of solids and structures - Volume 1*, John Wiley, Chichester.

Crisfield, M.A. [1997] *Non-linear finite element analysis of solids and structures - Volume 2*, John Wiley, Chichester.

Crisfield, M.A., Hunt, G.W. and Duxbury, P.G. [1987] *Benchmarks for geometric non-linearity*, *NAFEMS Report*, P03.

Hayhurst, D.R., Dimmer, P.R. and Morrison, C.J. [1984] Development of continuum damage in the creep rupture of notched bars, *Phil. Trans. R. Soc. London*, Vol. A311, pp. 103-129.

Hill, R. [1950] *The mathematical theory of plasticity*, Oxford University Press, Oxford.

Hills, D.A. , Nowell, D. and Sackfield, A. [1993] "Mechanics of elastic contacts", Butterworth-Heinemann, Oxford.

Hinton, E. [1992] *Introduction to non-linear finite element analysis*, NAFEMS.

Hinton, E. and Ezatt, M.H. [1987] Fundamental tests for two and three dimensional small strain elastoplastic finite element analysis, *NAFEMS Report*, P06.

Johnson, K.L. [1982], "One hundred years of Hertz contact", *Proc. I. Mech. E.*, Vol. 196, pp. 363-378.

Johnson, K.L. [1985], "Contact mechanics", Cambridge University Press, Cambridge.

Knowles, N.C. [1990] A review of benchmark problems for geometric *non-lin*ear behaviour of 3D beams and shells, *NAFEMS Report*, R0009.

Kraus, H. [1980] *Creep analysis*, John Wiley, New York.

Kreyszig, E. [1979] Advanced engineering mathematics, John Wiley, New York.

Krieg, R. D. and Kreig, D. B. [1977] Accuracies of numerical solution methods for the elastic perfectly plastic model, *J. Pressure Vessel Technology*, Vol. 99, pp. 510-513.

Linkens, D. [1993] Selected benchmarks for material non-linearity, *NAFEMS Report*, R0026.

Lyons, P. and Holsgrove, S. [1989] Finite element benchmarks for 2D beams and axisymmetric shells involving geometric non-linearity, *NAFEMS Report*, P10.

Mindlin, R.D. & Deresiewicz, H. [1953], "Elastic spheres in contact under varying oblique forces", *J. Applied Mechanics*, Trans. ASME, Vol. 20, pp. 327-344.

Olukoko, O.A., Becker, A.A. and Fenner, R.T. [1994] "A review of three alternative approaches to modelling frictional contact problems using the boundary element method", *Proc. Royal Society of London*, Series A, Vol. 444, 37-51.

Prinja, N.K. and Clegg, R.A. [1993] Assembly benchmark tests for 3D beams and shells exhibiting geometric non-linear behaviour, *NAFEMS Report*, R0025.

Riks, E. [1979] An incremental approach to the solution of snapping and buckling problems, *Int. J. Solids & Structures*, Vol. 15, pp. 529-551.

Timoshenko, S.P. and Gere, J.M. [1961] *Theory of elastic instability*, McGraw-Hill, London.

Zienkiewicz, O.C. and Taylor, R.L. [1991] *The Finite element method-Volume 2*, McGraw-Hill.

GLOSSARY

<u>**Alternating plasticity**</u> occurs in cyclic loading when there is a progressive increase in total strain with each cycle.

<u>**Arc-length method**</u> is a non-linear iterative technique used to arrive at a reliable and accurate solution at or near limit points where the slope of the load-displacement curve changes sign. The solution is advanced by a suitable 'arc length' measured along the load path itself, rather than a load or displacement increment.

<u>**Axisymmetric problems**</u> are formed by rotating a two-dimensional flat plane through 360° about the z-axis. For an axisymmetric assumption to be valid, the geometry and all the variables must be axisymmetric.

<u>**Bauschinger effect**</u> is observed when the yield stress in compression is less than the equivalent value in tension

<u>**Benchmarks**</u> are problems with known and reliable reference solutions, preferably non-FE solutions, that can be used to test the accuracy of FE programs.

<u>**Bifurcation**</u> occurs on the load-displacement curve as the load path forks into two or more solution paths that satisfy equilibrium. Only one path is stable, while all others are unstable.

<u>**Buckling**</u> is a geometric instability usually caused by compressive forces. Buckling can be analysed as a special case of geometric non-linearity involving equations of the type classified in mathematics as 'eigenvalue' problems.

<u>**Cauchy stress**</u> (also called true stress) is defined as the force over the current (instantaneous) area.

<u>**Conservative load**</u> is a load which always applies in a fixed direction regardless of the deformation of the body. A typical example is a gravitational load, which always applies vertically.

<u>**Constitutive Equations**</u> are the relationships between stress and strain incorporating the material properties, such as Young's modulus and Poisson's ratio.

<u>**Contact Problems**</u> are non-linear problems in which two or more surfaces come into or out of contact. Elements in the contact areas may stick or slip according to the prescribed coefficient of friction.

<u>**Convergence**</u> refers to the accuracy and reliability of the non-linear solution procedures. Convergence is achieved when equilibrium is satisfied

Creep is a time-dependent non-linear material behaviour which usually occurs in metals at high temperatures in which the effect of the variation of stress/strain with time is of interest.

Deviatoric stress represents the shear component of the stress, i.e. the remainder of the stress after deducting the hydrostatic stress component. The deviatoric stress is responsible for the plastic flow and creep behaviour through the shear component

Displacement control refers to problems where the displacement of a specified node or a set of nodes is increased by a small increment. It is recommended for snap-through problems, because the load-displacement tangent becomes horizontal and the load must subsequently remain constant or decrease in order to follow the load-displacement curve

Engineering strain (also called nominal strain) is defined as the ratio of the change in length over a given gauge length to the original length.

Explicit Euler Method (also known as forward-difference) is a time-marching scheme in which the new value of the variable is arrived at by using the tangent at the previous time step. This method needs very small time steps to remain stable.

Flow rule is used to define the constitutive relationship between the plastic strain increment and the stress increment.

Follower load is a load which changes its direction during the deformation of the structure, e.g. an internal pressure in a vessel which changes its position and direction to remain perpendicular to the surface as the vessel deforms.

Gaussian Points are strategically placed integration points inside the element at which the stress values can be evaluated.

Geometric Non-linearity occurs when the changes in the geometry of a structure due to its displacement under load are taken into account in analysing its behaviour. Geometric non-linearity problems can involve large or small deformations and strains.

Hydrostatic stress is the average of the direct stresses with no shear component. The hydrostatic stress is composed of principal stresses with no shear component and thus is responsible only for the change in volume (but not shape) of an element of the material.

Implicit Euler Method (also known as backward-difference) is a time-marching scheme in which the new value of the variable is arrived at using the tangent at the next time step. This method is unconditionally stable, i.e. it remains stable even if large time steps are used, although small time steps are more accurate.

Isotropic hardening occurs when the original yield surface increases in size with increasing plastic strain but maintains its original shape

Iterations are indirect ways of solving non-linear problems, based on successive corrections of an initial guess (trial solution).

Kinematic hardening occurs when the original yield surface is translated to a new position in the stress space as the plastic strain is increased, with no change in size or shape

Large Deformations usually occur in problems involving non-linear geometric behaviour, such as buckling problems, in which the deformations are so large that a new stiffness matrix has to be calculated.

Limit point is a point on the load-displacement curve where the tangent is horizontal. and the structure cannot sustain a further increase in the load.

Line search is a method used in geometric non-linearity problems to accelerate the incremental-iterative procedure, but at the same time maintaining the reliability of the solutions by minimising the total potential energy.

Load control refers to problems where the load at a specified node or a set of nodes is increased by a small increment. Load control can be used in most application, except where snap-through behaviour is encountered, because the load must subsequently remain constant or decrease in order to follow the load-displacement curve

Material Non-linearity occurs when the stress-strain constitutive relationships are non-linear as in plastic or creep behaviour.

Mixed hardening is a combination of isotropic and kinematic hardening where the original yield surface both expands and translates to a new position with increasing plastic strain.

Newton-Raphson Method is a special method of solving non-linear equations using an incremental-iterative approach based on either a constant or an updated slope in each iteration.

Normality rule is used in plasticity to ensure that the plastic strain components are in a ratio such that their resultant is in a direction normal to the plastic potential surface

Norton-Bailey equation is a creep law in which the creep strain rate is proportional to the power of the stress and the time.

Perfectly plastic material is a material where there is no further increase in the yield stress after initial yielding.

Plane strain is used to define very 'thick' geometries where the strain across the thickness is neglected, but the stress there is non-zero.

Plane stress is used to define 'thin' geometries in the z-direction where the stress across the thickness is neglected.

Plasticity is a non-linear material behaviour which is usually assumed to be independent of time such as the elastic-plastic behaviour of metals in which the structure is loaded past the yield point.

Primary creep is the initial stage of creep where the strain rate decreases.

Principal planes are planes on which only the normal component of the stress remains, i.e. the shear component of the stress is zero. The stresses and strains acting on these planes are called principal stresses and principal strains.

Principal stresses/strains are normal stresses/strains with no shear component that act on the principal planes. The magnitudes of the principal stresses/strains are independent of the coordinate system used.

Proof stress is used in some materials, such as alloy steels and non-ferrous metals, where there is no clearly identifiable yield stress. Proof stress is defined at a given plastic strains, typically at 0.1% and 0.2%.

Ratchetting occurs in cyclic loading where plastic strains keep on accumulating incrementally with each cycle, leading to eventual failure.

Reduced Integration is an integration scheme in which the number of Gaussian points is reduced. Despite being a less accurate technique than full integration, reduced integration can improve the overall accuracy of several types of elements and consumes less computation time.

Secondary creep is the creep stage where the creep strain rate is constant (or approximately constant).

Shakedown occurs in cyclic loading where the plastic strain in each cycle is relatively small, i.e. the total strain is less than twice the yield strain (the strain when the stress reaches the yield stress).. The structure is assumed to have settled down to an elastic state.

Snap-back usually occurs where there is a vertical tangent in the load-displacement curve where the load on the body suddenly drops, but the displacement remains constant

Snap-through usually occurs where the deformed shape suddenly jumps from one position to another, but the load remains constant. Dynamic effects may be caused by the sudden movement. In the load-displacement path, snap-through is exhibited by a horizontal tangent in the load-displacement curve.

Strain hardening law is used in analysing creep behaviour under a variable load where the creep strain rate is assumed to depend on the current stress and the accumulated creep strain.

Strain hardening material is a material where the stress after initial yielding increases with continuing plastic strains.

Stress Relaxation usually occurs in creep problems when the structure is loaded up to a stress level and then held at constant strain

Tertiary creep is the creep stage where the strain rate increases very rapidly, followed by eventual failure.

Time Hardening law is used in analysing creep behaviour under a variable load where the creep strain rate is assumed to depend on the current stress and the time from the start of the test.

Tresca criterion is a yield criterion for metals which assumes that yielding starts when the maximum value of the shear stress is reached.

True strain (also called natural or logarithmic strain) is defined as the integral of the incremental change of length over the current length.

True stress is defined as the force over the current (instantaneous) area.

Viscoelasticity is a non-linear material behaviour in which both the effects of plasticity and creep are exhibited, i.e. the stress is dependent on the strain rate.

Von Mises yield criterion is a criterion of yielding for metals which relates the multi-axial stresses in plasticity to the uniaxial behaviour by using the concept of the critical value of the shear strain energy stored in the material.

Yield criterion is used to define how the multi-axial behaviour of the material is related to the uniaxial behaviour.

Yield function (or yield surface) defines when initial yielding occurs.

Yield stress is the stress at which yielding occurs, i.e. the stress value which separates the stress-strain curve into an elastic portion and a plastic portion.