

Computational geomechanics

Course notes

v2024.3

Brice Lecampion
Geo-Energy Lab - EPFL

November 7, 2024

This document is under construction. It is not complete neither full-proof yet (a coffee is offered for anyone spotting typos). Every week of the course correspond to one chapter of these notes. I will provide updates of these notes when required.

I ask you to read the chapter of the week **BEFORE** the class. During the class (Mondays pm), I will prepare a short number of questions to kick off the discussion and we will review together the content of the material. We will also review / correct problems you encountered during the exercices.

Chapters 7 and 12 - respectively on thermo-poroelasticity and multiphase flow are **NOT** directly part of the course (no quiz questions will be asked on this material) - but can be useful for advanced projects that you may do in the future.

Contents

1	Introduction / preliminaries	7
1.1	Notation	7
1.2	Mesh description	8
1.3	Iso-parametric element	10
1.4	Estimating the gradient from the knowledge of nodal values	11
1.4.1	Projection	12
2	Confined Steady-state groundwater flow	15
2.1	Mass conservation	15
2.2	Darcy's law	16
2.2.1	Piezometric head / Expression for 2D cases	17
2.2.2	Permeability tensor	17
2.3	Steady-state flow problem	17
2.3.1	Boundary conditions	18
2.4	Finite element discretization	18
2.4.1	Weak form	18
2.4.2	Finite element discretization	19
2.4.2.1	Final system	20
2.4.2.2	Application of Dirichlet boundary conditions	21
2.4.2.3	Evaluation of the flux	21
2.5	Extensions	22
3	Unconfined steady-state flow	23
3.1	Capturing unconfined flow with a relative permeability function	24
3.1.1	Simplifications	25
3.1.2	Pore pressure in the soil when $S_l < 1$	26
3.2	Boundary conditions on the potential seepage boundary	26
3.3	Solution of the corresponding non-linear problem	28
3.3.1	Solution of the non-linear system via fixed point iterations	29
3.3.1.1	Fixed point iterations scheme - known boundary conditions	29

3.3.1.2	Newton-Raphson scheme - known boundary conditions . . .	30
3.3.1.3	Complete fixed-point solver including the location of the seeping nodes	31
4	Transient flow	33
4.1	Governing equations	33
4.1.1	Storage coefficient / diffusion equation	34
4.2	Numerical Solution by Finite difference	35
4.2.1	A one-dimensional problem	35
4.2.2	Spatial discretization	36
4.2.3	Time integration	38
4.3	Finite Element	40
4.3.1	Weak form	40
4.3.2	Time integration	41
5	Quasi-static Poroelasticity I	43
5.1	Material Description	43
5.2	Constitutive law	44
5.2.0.1	Variation of porosity, factoring out the fluid type	45
5.2.1	Drained / undrained conditions	46
5.2.1.1	drained (or “slow” loading)	46
5.2.1.2	undrained (or “fast” loading)	46
5.2.2	Simple micromechanic relations	47
5.2.3	Relations between different poroelastic constants	49
5.2.4	Non-linearities	49
5.3	Fluid flow in porous media	50
5.4	Quasi-static poroelasticity	51
5.5	Solution of undrained problems	52
5.5.1	The undrained elasticity problem	53
5.5.1.1	Finite element solution	53
6	Quasi-static Poroelasticity II	57
6.1	The initial boundary value problem	57
6.1.1	Initial and boundary conditions	58
6.2	Weak form of the problem	59
6.2.1	Introducing the poroelastic constitutive equation	60
6.2.2	Final weak form	60
6.2.3	Discretized final system of ODEs	61
6.2.4	Solving for variation from the initial state	62
6.3	Spatial discretization	63
6.4	Time Integration	63

6.4.1	Undrained problem	64
7	Extension to Thermoporoelasticity	65
7.1	Constitutive equations	65
7.1.1	The solid porous skeleton	65
7.1.2	The saturating fluid	66
7.1.3	The porous material	66
7.1.4	Recapitulation of the linear thermoporoelastic constitutive laws . . .	67
7.2	Conservation laws	68
7.3	Summary	69
7.4	The geothermal gradient	70
8	Failure of geomaterials - recaps	71
8.1	Geomaterials fail in term of Terzahi's effective stress	71
8.2	Elasto-plastic constitutive law	72
8.3	Typical failure envelopes for geomaterials	74
8.3.1	Mohr-Coulomb	75
8.3.1.1	Drucker-Prager equivalent	76
8.3.2	Tresca	78
8.3.3	Modified Cam-Clay	78
8.3.4	Failure of faults / pre-existing fractures	78
9	Elasto-plasticity I	80
9.1	Elasto-plastic mathematical description - recaps	80
9.2	Tri-axial test of a Mohr-Coulomb material	81
9.2.1	Hydrostatic confinement stage	82
9.2.2	Deviatoric loading stage	83
9.2.3	Full Stress-strain curve	85
9.2.4	Summary	85
9.3	Toward the solution of elasto-plastic problems with FEM: non-linear small strain elasticity	87
9.4	Solution of elasto-plastic problem via FEM	89
9.4.1	Incremental stress update	90
9.4.1.1	Plastic corrector step (Radial return mapping)	91
9.4.1.2	Consistent tangent operator	92
9.5	Unstable equilibrium - The arc length method	93
10	Poro-plasticity	94
10.1	Undrained poroplastic response	95
10.2	Drained poroplastic response	96

11 Limit analysis / rigid plasticity	97
11.1 The rigid plastic boundary value problem	97
11.1.1 Principle of virtual power	98
11.1.2 Load multiplier	99
11.2 Lower bound and upper bound theorems	99
11.2.1 Lower bound	100
11.2.2 Upper bound	101
11.3 Numerical limit analysis	102
11.3.1 Lower bound problem	102
11.3.2 Upper bound problem	103
11.3.3 Mixed formulation	105
12 Two-phases flow in porous media	107
12.1 Formulation	107
12.1.1 Saturation and relative permeability	107
12.1.1.1 Mass balances	108
12.1.1.2 Phases fluxes and total fluxes	108
12.1.1.3 System in term of S_1 and p_2	109
12.1.1.4 Initial and boundary conditions	110
12.2 FEM	110
A Displacement, Strain, Stress - some recaps	114
A.1 Transformation, displacement and strain	114
A.1.1 Small Strain	116
A.1.2 Material derivatives - Lagrangian vs Eulerian	117
A.1.3 Small strain & compatibility conditions	118
A.2 Stress and equations of motion	118
A.2.1 Principal stresses - Mohr Circles etc.	120
A.2.1.1 Coordinates transform	120
A.2.1.2 Principal stresses & invariants	120
A.2.1.3 Normal and shear stress to a plane	121
A.2.1.4 Mohr Circles	121
A.2.1.5 Mean and deviatoric stress	122
A.3 Exercises	122
B Coordinates systems	124
C Gauss integration rules	134

Organization

In this course, we will use the finite element method to solve geomechanical problem - using iso-parametric continuous Galerkin techniques. You should all have followed an introductory course in finite element to properly follow this course (as well as continuum mechanics, soil, groundwater seepage and rock mechanics course). For background, I refer you to the textbooks of Zienkiewicz & Taylor (2005) and Hughes, T.J.R. (1987) ¹.

During the first half of the course, we will use and further develop a set of finite element routines in Python for i) steady state groundwater flow problem (confined and unconfined cases), ii) transient groundwater flow problem, iii) elasticity and finally iv) poro-elasticity (i.e. when flow and mechanical deformation becomes coupled and leads to undrained / drained responses). During that first half of the semester, we will go in-depth in the required finite element routines and algorithms. We will distribute a set of already coded routine via GitHub - you will have to code up scripts and some specific functions / code parts. Every week, we will issue a new release (containing the corrections), therefore slowly building everything we need to tackle the next week exercise. You will have an homework assignment for which you will be using this set of routines (with a short report and the code as deliverables). This assignment will be given in week 7 of the course and will have to be handed out at the end of the semester (December 23).

In the second part of the course (from week 8 on), we will account for the non-linear mechanical behavior of geomaterials. We will discuss and derive numerical schemes for plasticity and poro-plasticity, but instead of coding up everything ourselves, we will use a commercial Finite Element software: OptumG2, tailored for the solution of rigid-plastic and elastoplastic problems encountered in geotechnical engineering.

Although this course focus solely on the finite element method, other numerical methods exist to solve geomechanical problems (all methods have their advantages and drawbacks depending on the problem to solve). Therefore, in team of two, you will have to research and prepare a 15 minutes presentation to all the class on a given numerical method not covered during this course (we will assign randomly different numerical methods to the different teams by week 3). These presentations will take place during weeks 12/13.

Two multiple choice tests will occur (week 6 and week 13).

¹A large number of other good textbooks also exist.

Chapter 1

Introduction / preliminaries

We will use the finite element method (FEM) to solve porous media flow and poromechanical problems encountered in geomechanics. I will assume a knowledge of continuum mechanics, fluid mechanics, but also description and mechanical behavior of soils and rocks, theory of groundwater flow. I will provide quick recaps when needed. Appendix A contains basic recap of continuum mechanics, the other appendices recalls the different PDE¹ operators in cartesian, cylindrical and spherical coordinates systems.

1.1 Notation

The laws of conservation of mass and balance of momentum combined with constitutive relations for the behavior of (here porous) material yield sets of partial differential equations (PDEs). We will use indicial notation when deriving / writing partial differential equation prior to their discretization.

We will use an orthornormal cartesian coordinates system: $\mathbf{e}_1, \mathbf{e}_2, (\mathbf{e}_3)^2$ - otherwise stated. We will denote $n_d = 2$ for 2D and $n_d = 3$ for 3D problems. For sake of clarity, I may often 'confuse' the notation writting $x_1 = x, x_2 = y$ (, $x_3 = z$) in order to lighten the presentation - it should be readily understandable from the context.

Vector $\mathbf{v}$ with components (v_1, v_2, v_3) will be noted as $v_i, i = 1, 3$ (or $i = 1, 2$ in 2D), second order tensor with components $\sigma_{ij}, i, j = 1, 3$ ($i, j = 1, 2$ in 2D), fourth order tensor c_{ijkl} . We will use the convention of summation on repeated (so called dummy) indices otherwise stated, i.e. $\sum_{j=1,3} \sigma_{ij} n_j \equiv \sigma_{ij} n_j$ etc.

Spatial gradient of scalar field a and divergence of vector $\mathbf{v}$ will be written as :

$$\nabla a = \frac{\partial a}{\partial x_i} \equiv a_{,i} \quad \nabla \cdot \mathbf{v} = \frac{\partial v_i}{\partial x_i} \equiv v_{i,i}$$

¹Partial Differential Equations

²We will mostly do 2D problems.

and gradient of vector $\mathbf{v}$ and divergence of tensor σ as:

$$\nabla \mathbf{v} = \frac{\partial v_i}{\partial x_j} \equiv a_{i,j} \quad \nabla \cdot \sigma = \frac{\partial \sigma_{ij}}{\partial x_j} \equiv \sigma_{ij,j}$$

Times derivatives on the other hand will be written as usual e.g.

$$\frac{\partial f}{\partial t} \equiv \partial_t f$$

When discussing finite element matrices, the use of indicial notation quickly becomes too heavy. It is useful to then adopt an *index-free* notation. Aside from avoiding the proliferation of indices, we will see that this notation will help better understanding the computer implementation. In particular, we will write vector in bold and lower letter, e.g.

$$\mathbf{v} \equiv \begin{bmatrix} v_1 \\ v_2 \\ v_2 \end{bmatrix}$$

and matrices in bold and capitals

$$\mathbf{D} = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix}$$

There is a direct correspondance with the indicial notation, and matrix product vector are written as

$$\mathbf{D} \cdot \mathbf{v} = d_{ij}v_j \quad \mathbf{v}^T \cdot \mathbf{v} = v_i v_i \quad \mathbf{v}^T \cdot \mathbf{D} \cdot \mathbf{v} = v_i d_{ij}v_j = d_{ij}v_i v_j$$

1.2 Mesh description

We will use mostly 2D unstructured meshes made of triangles as per figure 1.1. In the following notes, if the problem dimension is not stated explicitly, it is 2D. A domain Ω is meshed by a tessellation $\mathcal{T}$ defined as the combination of n_e triangular elements and there is n_n nodes in such a mesh. It is described by:

1. a table containing the spatial coordinates of all the nodes in the mesh, e.g. in 2D an array of size $n_n \times 2$:

$$\begin{array}{cc} x_1^1 & x_2^1 \\ x_1^2 & x_2^2 \\ \dots & \dots \\ x_1^i & x_2^i \end{array}$$

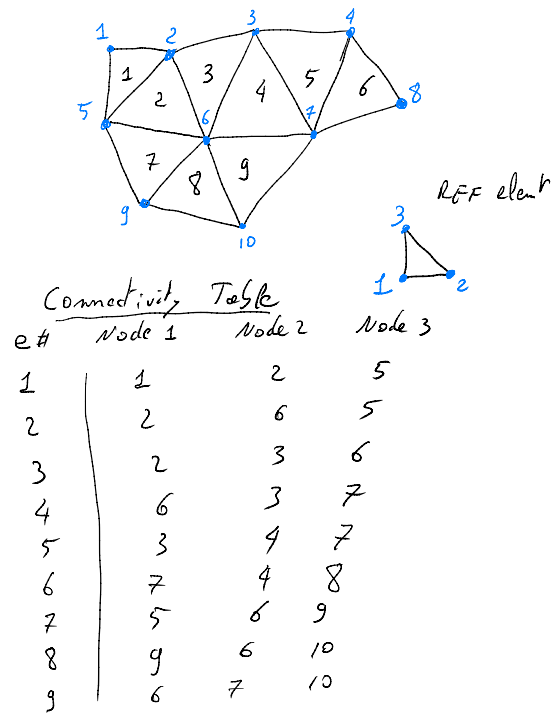


Figure 1.1: Exemple of a triangular mesh - connectivity table.

2. a connectivity table, describing the vertex (nodes) of each element. For a triangular *linear* mesh, it is an array of size $n_e \times 3$: row i corresponding to element i , and the columns correspond to the index of the first, second and third nodes of that triangle. (see fig. 1.1 for an example). Note that for a mesh made of quadratic triangle the connectivity array will have a size $n_e \times 6$ (and so on for other elements).

We will use Mesh2D, a matlab library to generate unstructured triangular meshes. We ship it with our finite element routines, a number of examples are included. In this first week, you will build a mesh for the configuration of a sheet pile wall. We will use that mesh to solve the steady-state groundwater flow next week.

1.3 Iso-parametric element

We denote by isoparametric element, an element for which the same shape functions are used to define the element's geometrical shape and the spatial variation of the unknown within the element.

The domain Ω^e of a given triangle can be mapped to a reference/unit triangular element by a map³ - see Figure 1.2. That means that a point $\xi = \begin{pmatrix} \xi \\ \eta \end{pmatrix}$ in the unit reference triangle can be related to the coordinate of a point $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \equiv \begin{pmatrix} x \\ y \end{pmatrix}$ by a linear mapping

$$\mathbf{x} = \sum_{a=1}^3 N_a(\xi) \mathbf{x}_a^e$$

where $\mathbf{x}_a^e$ is the coordinates of node a on the linear triangle. N_a are the corresponding shape functions. Figure 1.2 provides the definition of the linear 3 nodes triangle (also named CST for constant strain triangle as the derivatives are uniform inside the element). We recall that the jacobian matrix of the transformation $\mathbf{x} \rightarrow \xi$ is

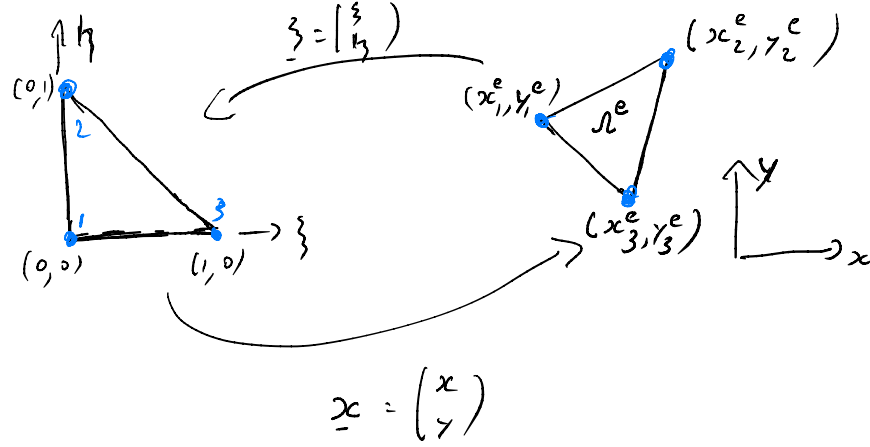
$$J_{i,j} = \frac{\partial x_i}{\partial \xi_j}$$

and the jacobian determinant of the transformation is denoted $j = \det(J)$.

We will use iso-parametric element, which means that we will discretize partial differential equations using the same interpolation for the unknowns -say a scalar field $u(x, y)$ - as for the spatial coordinates. As a result, the unknowns will be located at the node of the mesh, and we will write for $\mathbf{x}$ in the domain Ω^e of triangle e

$$u(x, y) = u(\mathbf{x}) = \sum_{a=1}^3 N_a(\xi(\mathbf{x})) u_a^e$$

³Such a map is linear for a linear 3 nodes triangle.



Node	coord. (ξ, η)	$N_a(\xi, \eta)$	$(\partial_\xi N_a, \partial_\eta N_a)$
1	$(0, 0)$	$1 - \xi - \eta$	$(-1, -1)$
2	$(0, 1)$	η	$(0, 1)$
3	$(1, 0)$	ξ	$(1, 0)$

Figure 1.2: Linear Triangle - Tri3

where u_a^e denotes the value of the scalar field at the node a of element e . It can be directly extended to a vector of unknowns $\mathbf{u}$ (containing for example the displacement in the x and y directions), e.g.:

$$\mathbf{u}(\mathbf{x}) = \sum_{a=1}^3 N_a(\xi(\mathbf{x})) \mathbf{u}_a^e$$

Note that a recall of Gaussian integration rules in 1 and 2D dimensions can be found in appendix C.

1.4 Estimating the gradient from the knowledge of nodal values

In a number of cases, we will solve a pde (partial differential equation) in terms of a primary variable, say for example a scalar field h . After solution of the resulting system, we will have values of this field at the nodes of the mesh. It is often important to then estimate quantities related to the spatial gradient (the spatial derivatives) of the primary unknown. This is notably the case in elasticity where one solve for the displacement vector and then want to estimate the stress tensor. Similarly in the case of groundwater flow, we will solve for the piezometric head h (or the pore pressure p) and then often wants to estimate the flux

(the relative discharge) $\mathbf{q}$ which is directly related to the gradient of the piezometric head via Darcy's law:

$$q_i = -K_{ij} \times h_{,j} \quad (1.1)$$

where K_{ij} is the hydraulic conductivity tensor (dimensions of m/s) - sometimes in geotechnics referred to as the permeability coefficient tensor.

We see that if we use linear triangle (TRI3) finite element as described in figure 1.2, the spatial derivatives of the shape function are uniform over the element. In other words, the flux q_i is uniform over the element. We clearly see that generally if shape functions of order p are used, the derivatives of the field will vary as $p - 1$ over the element. Gauss integration allow to integrate exactly a polynomial of order $p = 2n_p - 1$ with n_p integration point. A constant field over the triangle can thus be integrated by simply multiplying its constant value by the element area - and the location of the corresponding Gauss point is the element centroid. For any order of interpolation (or for the bi-linear quadrilateral), the derivatives of the field are obtained at Gauss integration points.

For a mesh of linear triangle (TRI3), the flux (1.1) can thus be obtained at the centroid of each element. Plotting the results may not be very easy for very unstructured mesh - it would be better to have the flux expressed at the nodes in order to use built-in plotting function (of Matlab) to display the component of the flux.

1.4.1 Projection

It is often useful for post-processing to estimate the gradient at the nodes from the knowledge of the field solution at the nodes - for example to ease plotting, estimate flux at boundaries etc. This can be done via finite element - a procedure often denoted projection. The term projection refers to the fact that we project the gradient for the original FE space which is typically C0 only between elements (i.e. where the gradient is discontinuous between elements) to a space where the gradient are continuous between element. We discuss here how to do that for the case of Darcy flow - restricting to 2D for clarity. The same can be for stresses in mechanics for example.

Assume that we know the value of the piezometric head h at all nodes in the mesh (we will solve this class of problem next week). We want to estimate the flux vector q_i at all the nodes of the mesh. We assume for simplicity that the hydraulic conductivity tensor K_{ij} is uniform (this can be easily relaxed). We thus need to solve Darcy's law on the whole domain Ω :

$$q_i(x_i) = -K_{ij} \times h_{,j} \quad \text{for all } x_i \text{ in } \Omega \quad i, j = 1, 2$$

from the knowledge of the head value h at all nodes of the mesh. Let's re-write this as two sub-problem for each component of the flux vector:

$$\begin{aligned} q_1 \equiv q_x(x_i) &= -K_{xj} \times h_{,j} & \text{for all } x_i \text{ in } \Omega & \quad j = 1, 2 \\ q_2 \equiv q_y(x_i) &= -K_{yj} \times h_{,j} & \text{for all } x_i \text{ in } \Omega & \quad j = 1, 2 \end{aligned}$$

We can do this by building a weak form. The same procedure apply for both equation therefore, I write the derivation for q_x (it is strictly the same for q_y pending using K_{yj} instead of K_{xj}). Multiplying the previous equation by a test vectorial field v (with the usual continuity requirement), after integration we obtain

$$\int_{\Omega} v q_x \, dV = \int_{\Omega} v \times (-K_{xj} h_{,j}) \, dV \quad (1.2)$$

which we can obtain by combining the integration over all the elements in the mesh: i.e. $\int_{\Omega} = \sum_e \int_{\Omega_e}$. We now switch to the index-free notation. We write

$$\int_{\Omega} v(\mathbf{x}) \cdot q_x(\mathbf{x}) \, dV = \sum_e \int_{\Omega_e} v(\mathbf{x}) \cdot q_x(\mathbf{x}) \, dV = \sum_e \int_{\Omega_e} v(\mathbf{x}) (-\mathbf{K} \cdot \nabla \mathbf{h}(\mathbf{x}))_x \, dV$$

where $\nabla \mathbf{h}$ is the gradient (vector) of h . Using a interpolation over the linear triangle of the same order for v , q_x and h , for $\mathbf{x}$ in element e

$$\begin{aligned} q_x(\mathbf{x}) &= \sum_{a=1}^3 N_a(\xi(\mathbf{x})) q_{xa}^e \\ v(\mathbf{x}) &= \sum_{a=1}^3 N_a(\xi(\mathbf{x})) v_a^e \\ h(\mathbf{x}) &= \sum_{a=1}^3 N_a(\xi(\mathbf{x})) h_a^e \end{aligned}$$

we can rewrite in a vector/matrix form

$$\begin{aligned} q_x(\mathbf{x}) &= \mathbf{N} \cdot \mathbf{q}_x^e \\ \mathbf{N} &= [N_1(\mathbf{x}) \quad N_2(\mathbf{x}) \quad N_3(\mathbf{x})] \\ \mathbf{q}_x &= \begin{bmatrix} q_{x1}^e \\ q_{x2}^e \\ q_{x3}^e \end{bmatrix} \end{aligned}$$

Similarly, the derivatives of h are

$$h_{,\xi} = \sum_{a=1}^3 N_{a,\xi}(\xi(\mathbf{x})) h_a^e$$

and in matrix form we write

$$\nabla \mathbf{h} = \underbrace{\begin{bmatrix} N_{1,1} & N_{2,1} & N_{3,1} \\ N_{1,2} & N_{2,2} & N_{3,2} \end{bmatrix}}_{\nabla \mathbf{N}} \cdot \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix}$$

where $N_{1,2} = \partial_{x_2} N_1$ etc.

We can therefore rewrite the weak form as

$$\sum_e \mathbf{v}^{eT} \cdot \underbrace{\left(\int_{\Omega_e} \mathbf{N}^T \cdot \mathbf{N} \, dV \right)}_{\mathbf{M}^e} \cdot \mathbf{q}_x^e = \sum_e \mathbf{v}^{eT} \cdot \underbrace{\left(\int_{\Omega_e} \mathbf{N}^T \cdot (-\mathbf{K} \cdot \nabla \mathbf{N} \cdot \mathbf{h}^e)_x \, dV \right)}_{\mathbf{f}_{\mathbf{q}_x}^e}$$

This can be further re-arranged as nodes are shared between neighbour elements. We can define a global vector $\mathbf{q}$ of length $n_d \times n_n - n_d = 2$ in 2D (there are 2 components of the flux vector at each nodes of the mesh), and a global vector $\mathbf{h}$ of length n_n (there is one value of h at each nodes of the mesh).

Then the contribution of the mass matrix of each element can be added at the different nodes (by looping over the element and mapping the local and global numbering of the equations). Such a procedure is called **assembly**, and we will schematically write it as

$$\dagger \mathbf{M}^e = \mathbf{M}$$

We finally obtain in matrix form, the following

$$\mathbf{v}^T \cdot \mathbf{M} \cdot \mathbf{q}_x = -\mathbf{v}^T \cdot \mathbf{f}_{\mathbf{q}_x}$$

This valid for any value of the test function $\mathbf{v}$, such that we obtain a linear system for $\mathbf{q}_x$

$$\mathbf{M} \cdot \mathbf{q}_x = \mathbf{f}_{\mathbf{q}_x}$$

The same will results for $\mathbf{q}_y$

$$\mathbf{M} \cdot \mathbf{q}_y = \mathbf{f}_{\mathbf{q}_y}$$

note that only the right hand side will differ.

Chapter 2

Confined Steady-state groundwater flow

In this course, a porous media will be defined by its porosity ϕ (also sometimes noted n in soil mechanics) and intrinsic permeability k (dimension $[L^2]$). In this chapter, we focus on steady-state fluid flow in porous media. In shallow geotechnical applications, the filling fluid is typically water, but the theory extend directly to other filling fluids (gas, oil, etc.). We will restrict to the case of single phase flow. In other words, only a single phase will be assumed to be filling the pores (liquid water or oil only etc.). Moreover, in this chapter, we will tackle problems where the phreatic surface location is known a-priori (so-called confined flow problems).

2.1 Mass conservation

For an unit representative volume element (RVE) of porous media, the conservation of fluid mass filling the pores of the porous media is expressed in differential form as

$$\frac{\partial \rho_f \phi}{\partial t} + (\rho_f q_i)_{,i} = \rho_f \gamma \quad i = 1, 2 (\text{in 2D}), 3 (\text{in 3D}) \quad (2.1)$$

where ρ_f is the fluid density, ϕ the porous medium porosity, q_i is the fluid filtration vector and γ a possibly existing source/sink term (for example due to an injection).

If you wonder how to obtain this equation - have a look at your groundwater flow course or fluid mechanics course.

- The first term $\frac{\partial \rho_f \phi}{\partial t}$ corresponds to the time variation of density and porosity in the unit RVE
- The divergence term $(\rho_f q_i)_{,i}$ express the difference between the flux of fluid mass exiting and entering the RVE

q_i is the filtration vector (also called specific discharge or fluid flux). It describes the fluid motion with respect to the solid, it's the phase-slip velocity weighted by the fluid volume in the RVE (i.e the porosity):

$$q_i = \phi \left(v_i^f - v_i^s \right)$$

where v_i^f and v_i^s are the fluid and solid velocity respectively.

2.2 Darcy's law

The filtration vector describes the fluid motion. In porous media, such flow is typically non-inertial (laminar), thus we can draw an analogy with Stokes flow -forgetting for a moment the complex pore structure and assuming for simplicity e.g. a simple pipe. For a Newtonian fluid of μ_f and density ρ_f , the fluid balance of momentum -neglecting inertia - simplifies to:

$$-\rho_f g \delta_{i3} - p_{,i} + \mu_f \left(v_{i,j}^f + v_{j,i}^f \right)_{,j} = 0$$

where the gravity is positive downward and aligned along $\mathbf{e}_3$ (e_3 is positive upward). δ_{ij} is the Kronecker delta (i.e. $\delta_{i3} = 1$ if $i = 3$, 0 otherwise). We therefore see - by dimensional arguments - that the velocity is proportional to $p + \rho_f g x_3$ (i.e. remember Bernoulli without inertia), fluid velocity thus scales as:

$$v_i^f \propto -\frac{C[L^2]}{\mu_f} (\rho_f g \delta_{i3} + \nabla p)$$

where C is a constant of dimension L^2 . In the case of a flow in a pipe, the solution of Stokes equation allows to obtain an expression for C which scales as a^2 where a is the pipe radius. Taking the argument, one step further, one can “lump” the complex pore network connectivity as a network of “pipes” and write the following linear relationship between the filtration vector and the driving force for the flow:

$$q_i = -\frac{k}{\mu_f} (p + \rho_f g x_3)_{,i} \quad (2.2)$$

The previous equation is known as Darcy's law, and k is denoted the intrinsic permeability (dimension $[L^2]$). The intrinsic permeability depends on: the pore-network geometry, porosity and “a pore diameter scale” a . It scales with respect to such a lengthscale as $k \propto a^2 \delta(\phi)$. A number of expression have been put forward for $\delta(\phi)$. A popular one, working reasonably well for an unconsolidated well-sorted sand is the Kozeny-Carman law: $k = a^2 \frac{\phi^3}{1-\phi^2}$. In typical porous media $a \leq 10^{-6}$ m, hence $k \leq 10^{-12} \text{m}^2$. The usual unit for the intrinsic permeability is the Darcy - equal to 10^{-12}m^2 for all practical purposes¹.

¹9.869233 $\times 10^{-13} \text{ m}^2$ to be precise.

In practice, the mobility coefficient $\kappa = k/\mu_f$ is sometimes used, as well as the hydraulic conductivity $K = k\rho_f g/\mu_f = k\gamma_w/\mu_f$ (in m/s).

2.2.1 Piezometric head / Expression for 2D cases

In hydrogeology / and applications near the surface where the fluid filling the pore is water, the piezometric head h is typically used. In a 3D coordinate system where gravity is positive downward and aligned along $\mathbf{e}_3$ (e_3 is positive upward), the piezometric head is defined as:

$$h = \frac{p}{\rho_f g} + x_3 \equiv \frac{p}{\rho_f g} + z$$

and $\rho_f g = \gamma_w$ is the fluid weight. In 2D, in a coordinate system where the gravity is positive downward and aligned along $\mathbf{e}_2$, the piezometric head is defined as

$$h = \frac{p}{\rho_f g} + x_2 \equiv \frac{p}{\rho_f g} + y$$

Darcy's law can be simply rewritten as

$$q_i = -\frac{k\rho_f g}{\mu_f} h_{,i} = -K h_{,i}$$

2.2.2 Permeability tensor

Darcy's law can be readily extended to account for anisotropy of permeability, by introducing an permeability tensor k_{ij} or hydraulic permeability coefficient tensor K_{ij} , i.e.

$$q_i = -K_{ij} h_{,j}$$

2.3 Steady-state flow problem

Under steady state conditions, the term involving time derivative disappears from the fluid mass conservation (2.1) which reduce to

$$\frac{1}{\rho_f} (\rho_f q_i)_{,i} = \gamma$$

Moreover for a liquid, compressibility is rather small $\rho_f \approx \rho_f^o(1 + \beta(p - p^o))$, the spatial variation of density can thus be safely neglected, and the fluid mass conservation reduces to the following steady-state volume conservation

$$q_{i,i} = \gamma$$

Using Darcy's law, we obtain the following PDE for the head - in the general case of an anisotropic permeability:

$$-(K_{ij}h_{,j})_{,i} = \gamma \quad (2.3)$$

at all points of the flow domain Ω of boundary Γ .

2.3.1 Boundary conditions

In steady-state flow problem, the following type of boundary conditions can be specified:

$$\begin{aligned} -q_i n_i &= q^g && \text{on the part of the boundary } \Gamma_q \\ h &= h^g \text{ (or } p = p^g) && \text{on the part of the boundary } \Gamma_p \end{aligned} \quad (2.4)$$

where n_i denotes the local outward normal to the boundary, q^g is the prescribed flux, and h^g the prescribed piezo-metric head. The first type of boundary conditions consist in specifying the flux discharge vector and is referred to as a Neumann boundary condition, while when the second type of boundary conditions (specification of the head which is the main unknown) is known as a Dirichlet boundary condition. In order for the problem to have an unique solution, the part of the boundaries where the head and flux are specified should not overlap, and their union must of course make the whole boundary of the domain e.g.

$$\Gamma_q \cap \Gamma_h = \emptyset \quad \Gamma_q \cup \Gamma_h = \Gamma$$

Other type of conditions will emerge in so-called unconfined flow problems as we shall see in the next chapter - e.g. line of seepage will be present when solving unconfined problems numerically (and will be a part of the solution).

2.4 Finite element discretization

We now detail briefly how to solve steady state flow problems (eq.(2.3)) using the finite element method. We will discretize the domain of interest Ω with finite elements. We first derive the weak form of the problem, combining the partial differential equation (2.3) with the boundary conditions (2.4).

2.4.1 Weak form

Let us denote as v a scalar test function with the usual continuity requirement (similar to the solution for the piezometric head h) such that $v = 0$ on Γ_p the part of the boundary where Dirichlet boundary conditions are specified. After multiplying (2.3) by v , integration over the domain gives

$$\int_{\Omega} v \times \left(-(K_{ij}h_{,j})_{,i} \right) dV = \int_{\Omega} v \times \gamma dV$$

using Green's identity (and because $v = 0$ on Γ_p -i.e. the test function is zero on the part of the boundary where Dirichlet boundary conditions are applied)

$$\int_{\Omega} v_{,i} K_{ij} h_{,j} dV = \int_{\Gamma_q} v q^q dS + \int_{\Omega} v \times \gamma dV \quad (2.5)$$

We note that if the imposed flux on Γ_q are zero (i.e. impermeable boundary conditions), the corresponding integral disappears - this is why zero flux / Neumann boundary conditions are denoted as “natural” boundary conditions (we do not need to do anything within the Finite Element Method: zero flux are imposed “naturally”). Note that we have written the Neumann condition as $-q_i n_i = q^q$ with n_i the outward normal, e.g. so q^q is positive for an entering flux.

2.4.2 Finite element discretization

We discretize Ω with a mesh $\mathcal{T}$. For sake of clarity, I specify here the discussion for **the linear triangle element**, but the same apply to any other type of isoparametric element. Using a interpolation over the linear triangle, we have the following interpolation for a point $\mathbf{x}$ within the element e :

$$v(\mathbf{x}) = \sum_{a=1}^3 N_a(\xi(\mathbf{x})) v_a^e$$

$$h(\mathbf{x}) = \sum_{a=1}^3 N_a(\xi(\mathbf{x})) h_a^e$$

where N_a are the corresponding shape functions (See previous chapter) and h_a^e the value of the head at the node a of element e . We will now switch to the index-free notation. We can rewrite in a vector/matrix form

$$h(\mathbf{x}) = \mathbf{N} \cdot \mathbf{h}^e$$

$$\mathbf{N} = [N_1(\mathbf{x}) \quad N_2(\mathbf{x}) \quad N_3(\mathbf{x})]$$

$$\mathbf{h}^e = \begin{bmatrix} h_1^e \\ h_2^e \\ h_3^e \end{bmatrix}$$

where $N_1, N_2, \dots$ are the shape function of the first, second... nodes of the element. We see that the gradient of h (or v) can be thus directly express as follows over element e

$$h_{,i}(\mathbf{x}) = \sum_{a=1}^3 N_{a,i}(\xi(\mathbf{x})) h_a^e$$

or in matrix vector form as

$$h_{,i}(\mathbf{x}) = \nabla \mathbf{N} \cdot \mathbf{h}^e$$

where for a linear triangle (in 2D) , the matrix $\nabla \mathbf{N}$ is of size 2 by 3 and is defined as

$$\nabla \mathbf{N} = \begin{bmatrix} \partial_x N_1(\mathbf{x}) & \partial_x N_2(\mathbf{x}) & \partial_x N_3(\mathbf{x}) \\ \partial_y N_1(\mathbf{x}) & \partial_y N_2(\mathbf{x}) & \partial_y N_3(\mathbf{x}) \end{bmatrix}$$

We can re-write the different terms of the weak form (2.5) as sum of the integral over all elements of the mesh, i.e.

$$\int_{\Omega} \cdot dV \equiv \sum_e \int_{\Omega^e} \cdot dV$$

The details of the element integrals are respectively

- The volume term on the left hand side

$$\int_{\Omega^e} \mathbf{v}^{eT} \cdot (\nabla \mathbf{N})^T \cdot \mathbf{K} \cdot \nabla \mathbf{N} \cdot \mathbf{h}^e dV = \mathbf{v}^{eT} \cdot \left(\int_{\Omega^e} (\nabla \mathbf{N})^T \cdot \mathbf{K} \cdot \nabla \mathbf{N} dV \right) \cdot \mathbf{h}^e$$

the middle matrix is usually denoted the element conductivity matrix $\mathbf{C}^e$:

$$\mathbf{C}^e = \int_{\Omega^e} (\nabla \mathbf{N})^T \cdot \mathbf{K} \cdot \nabla \mathbf{N} dV$$

- the source/sink integral on the right hand side

$$\int_{\Omega^e} v(\mathbf{x}) \gamma dV = \mathbf{v}^{eT} \cdot \left(\int_{\Omega^e} \mathbf{N}^T \cdot \gamma dV \right) \equiv \mathbf{v}^{eT} \cdot \mathbf{f}_{\gamma}^e$$

- the boundary integral (if one edge of the element lies on the boundary Γ_q , else the integral is zero)

$$\int_{\Gamma_q^e} v q^g dS = \mathbf{v}^{eT} \cdot \left(\int_{\Gamma_q^e} \mathbf{N}^T \cdot q^g dV \right) \equiv \mathbf{v}^{eT} \cdot \mathbf{f}_q^e$$

2.4.2.1 Final system

As usual nodes are shared between neighbouring element, the final weak form can be written as

$$\sum_e \mathbf{v}^{eT} \cdot \mathbf{C}^e \cdot \mathbf{h}^e = \sum_e \mathbf{v}^{eT} \cdot \mathbf{f}_{\gamma}^e + \sum_e \mathbf{v}^{eT} \cdot \mathbf{f}_q^e$$

From the element conductivity matrix, we can assemble a global conductivity matrix as - this is done by proper re-ordering of the equations to add the contribution of each element

to the entry of a global degree of freedom associated with node i . We schematically write such an assembly procedure as

$$\mathbf{C} = \dagger \mathbf{C}^e$$

You can have a look at e.g. function `AssembleConductivityMatrix` to see the step of such assembly procedure. Similarly, one can assemble the right hand side vectors

$$\dagger \mathbf{f}_q^e = \mathbf{f}_q \quad \dagger \mathbf{f}_\gamma^e = \mathbf{f}_\gamma$$

and as the weak form is valid for any test function v , we obtain the following system

$$\mathbf{C} \cdot \mathbf{h} = \underbrace{\mathbf{f}_q + \mathbf{f}_\gamma}_{\mathbf{f}}$$

2.4.2.2 Application of Dirichlet boundary conditions

In order to apply dirichlet boundary conditions (where the head is fixed), we first need to tag the necessary nodes. Let us denote this set of nodes as f (f for fixed) and u the complementary set of nodes where the head is to be solved (and is unknown before the solution of the problem). As we only need to solve for the unknown nodes, we can write the previous system only on the subset of unknown nodes:

$$\mathbf{C}^{uu} \cdot \mathbf{h}^u + \mathbf{C}^{uf} \cdot \mathbf{h}^f = \mathbf{f}^u$$

moving to the right hand side the contribution of the known fixed head nodes, we obtain the following system of equations to solve

$$\mathbf{C}^{uu} \cdot \mathbf{h}^u = -\mathbf{C}^{uf} \cdot \mathbf{h}^f + \mathbf{f}^u$$

which can be solved for $\mathbf{h}^u$ to finally obtain the vector of head at all nodes $\mathbf{h}$. Note also that instead of the head, the fluid pressure can be applied, i.e. recall that $h = p/\rho_f g + y$ so a boundary condition in pressure can be easily translated into its value for the head.

2.4.2.3 Evaluation of the flux

As discussed in the previous chapter, once the head is known at all nodes of the mesh, the gradient and flux discharge vector can be obtained at the Gauss integration point of each element. An additional projection problem can then be solved to obtain the values of the component of the fluid discharge q_i at all the nodes of the mesh (usually easier for plotting).

2.5 Extensions

The previous finite element procedure can be readily extended to the case where the hydraulic conductivity (permeability coefficient tensor) is spatially varying. Indeed, as can be seen from the discretized weak form, nothing prevent us to specify a different value of permeability in each element. Programatically, this is typically done by “tagging” elements belonging to the same material (therefore with the same physical properties) - such a vector is often denoted as the “matID” vector. Then a list of the different permeability (one value for each material) can be input to the assembly routine and prior to computing the corresponding element conductivity matrix, the proper element permeability is taken.

Chapter 3

Unconfined steady-state flow

In the previous chapter, we have dealt with problems where fluid flow was occurring over the whole domain, and pore-pressure was always positive or null (i.e. at atmospheric pressure at boundaries). However, there is a number of cases where the location of the phreatic surface is a priori unknown - this is the case for the flow across an earth dam, flow along an inclined surface (for stability problem) etc. The line of seepage is the boundary between the part of the soil which is fully saturated and fully dry. Of course, there is a “transition zone” denoted as the vadose zone whose extent/thickness is directly function of the fineness of the porous material (few centimeters or less for permeable / coarse grain soil, and up to dozen of meters for very fine / impermeable clay). Flow in the vadose zone is therefore intrinsically multiphase (liquid to gas), i.e. the saturation in liquid S_l falls from 1 (fully saturated) to zero (fully dry) - and $S_g = 1 - S_l$ is the vapor saturation.

For geotechnical applications, it is of utmost importance to obtain the location of the line of seepage/phreatic surface - and as a first approximation for most engineering application, one can neglect the thickness of the vadose zone and treated it as infinitely small. The type of problems where the location of upper boundary of the flow (line of seepage) is a-priori unknown are denoted as unconfined flow problems: such type of problems are non-linear as a

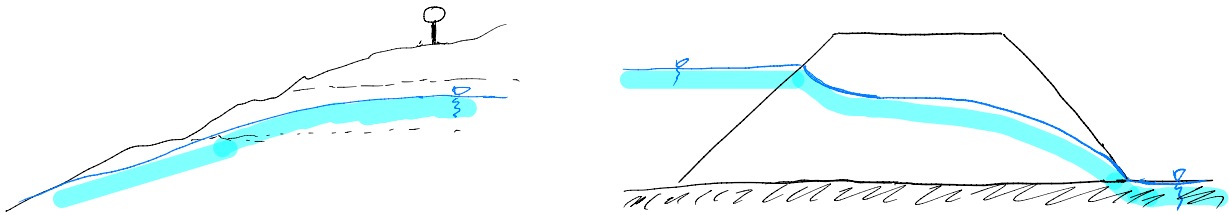


Figure 3.1: Examples of unconfined flow: natural seepage along a slope (left), simple flow through an earth dam (right). In these problems, the location of the phreatic surface is part of the solution.

result. A first approximation for such type of flow is the so-called Dupuit approximation (see your course on groundwater flow): analytical solutions (with different level of approximations) have been obtained for a number of problems (see notably the book of Harr (1962)). It is important to understand that because the location of the phreatic surface is part of the solution of the problem, the problem is *non-linear*.

Two families of numerical methods are in use for unconfined flow. Numerically, if the location of the line of seepage is known, the problem can then be solved with a strictly similar method than for confined flow - in that case, one of the boundary of the mesh conform to the line of seepage. Therefore, the first family of method iterates on the location of the boundary of the mesh coinciding with the line of seepage - this can be quite cumbersome as the domain needs to be re-meshed at each iterations until the line of seepage is properly resolved. The second family do not iterate on the mesh: the full domain is meshed (i.e. both the saturated and unsaturated part). The flow is resolved using a relative permeability function to capture the line of seepage / vadose zone, and an iterative procedure takes into account the different type of boundary conditions as function of the phreatic surface and domain boundaries location. This method is used in most in commercial codes, and the one we are going to follow here¹.

3.1 Capturing unconfined flow with a relative permeability function

In multiphase flow, the use of a relative permeability $k_r(S_l)$ function of saturation is a well established concept (see e.g. your geomechanics course). Such a relative permeability is equal to 1 for fully saturated conditions ($S_l = 1$) and zero for the fully dry case ($S_l = 0$). More precisely, Darcy's law can be rewritten for the fluid and vapour (gas) fluxes:

$$q_i^l = -\frac{k \times k_r(S_l)}{\mu_f} \times (p^l + \rho_f g x_2)_{,i}$$

$$q_i^v = -\frac{k \times (1 - k_r(S_l))}{\mu_v} \times (p^v + \rho_v g x_2)_{,i}$$

[with x_2 in direction of gravity here - e.g. 2D case]. A number of relation exist between relative permeability and saturation (Brooks & Corey, van Genuchten etc.).

Moreover, the saturation S_l is also related to the capillary pressure $p^l - p^v$, typically with a constitutive relation of the form

$$p^l - p^v = \frac{\sigma}{r} \times F(S_l, \dots)$$

¹In addition, automatic mesh refinement according to the gradient of the solution is often performed to improve the resolution of the seepage surface

where σ is the surface tension of air-vapor and r captures a characteristic pore-size (via its radius) - and we therefore see that the soil fineness will reflect on the size of the unsaturated vadose zone. [Note that irreversibility between drainage and imbibition often occurs - so things get more intricate].

3.1.1 Simplifications

However, for practical problems, one can do the following approximations:

1. we neglect vapor transport - because the vapor viscosity is much smaller than the fluid and as a result we can assume that vapor pressure in the pore space is equal to the atmospheric pressure: $p^v \sim 0$,
2. we can further adopt directly a direct relation between saturation and relative permeability liquid pressure $p^l = p$, the saturation and relative permeability (simplifying the non-linearity associated with capillary pressure)

$$k_r(S_l) = S_l$$

3. and adopt a simple approximation of the relation between liquid pressure and saturation in order to capture the transition from a fully saturated soil to a dry / completely unsaturated one.

We note in passing that de-saturation will physically start when the liquid pressure falls below zero (water can sustain a certain degree of tensile stresses due to capillary effect). In a “hard” approximation, one can write simply

$$\begin{aligned} S_l = k_r = 1 & \quad p > 0 \\ S_l = k_r = 0 & \quad p < 0 \end{aligned}$$

Of course, such an abrupt discontinuity is neither physically correct, neither numerically “friendly”. One thus typically smooth such an abrupt change. For example, the following piece-wise Heaviside function

$$\begin{aligned} S_l = k_r = 1 & \quad p > 0 \\ S_l = k_r = 1 + (1 - \epsilon)p/p_* & \quad -p_* < p < 0 \\ S_l = k_r = \epsilon & \quad p < -p_* \end{aligned}$$

where ϵ is a small number (avoiding 0 perm) and $p_* > 0$ a characteristic pressure. The previous approximation is piece-wise (derivative is not continuous). Another smoother function is

$$k_r(p) = \frac{1}{2} (1 + \tanh(p/p^*)) + \epsilon_M$$

where $p^* > 0$ controls the sharpness of the discontinuity, the pressure at which the relative permeability starts to fall below 1. ϵ_M is a small number (small but larger than machine precision) to ensure non-zero permeability (and avoid ill-conditioning of the resulting matrix after discretization). Indeed, for element with a permeability strictly equal to zero, all the entries of the element conductivity matrix will be zero. As a result, the global conductivity matrix may become singular! Numerically, this can become nasty - and it is reasonable not to be too drastic in the value of ϵ_M - i.e it should not be too small (e.g. $\epsilon_M = 10^{-10}$).

3.1.2 Pore pressure in the soil when $S_l < 1$

It is important to bear in mind that in the unsaturated part of the soil - where $S_l < 1$ - the liquid pore pressure will be negative $p^l < 0$. However, the actual average fluid pressure $\bar{p}$ is an average of the liquid and vapor pressure

$$\bar{p} = S_l p^l + (1 - S_l) p^v (\approx S_l p \text{ here for } p^v \approx 0)$$

Note that things are actually a bit more complicated and the average pressure is often written as

$$\bar{p} = \chi(S_l) p^l + (1 - \chi(S_l)) p^v$$

where $\chi(S_l)$ is a porous material dependent function (satisfying $\chi(1) = 1$, $\chi(S_l < 1) < 1$).

3.2 Boundary conditions on the potential seepage boundary

Another very important point related to unconfined flow problem is that the line of seepage is unknown, and some parts of the boundary may be fully dry ($S_l = 0$) and some fully wet ($S_l = 1$) with an outgoing fluid flux (i.e. seeping boundaries). Take for example the case of a earth dam as depicted in Fig. 3.2. Let's recapitulate the boundary conditions for the different segments of the boundary of the whole domain for that example (which covers all possible cases along the boundary of the domain).

- Impermeable base:

$$q_i n_i = q_n = 0 \quad \text{along } BC$$

- Hydrostatic pressure on the side of the dam (upstream):

$$h = h_A \quad \text{along } AB$$

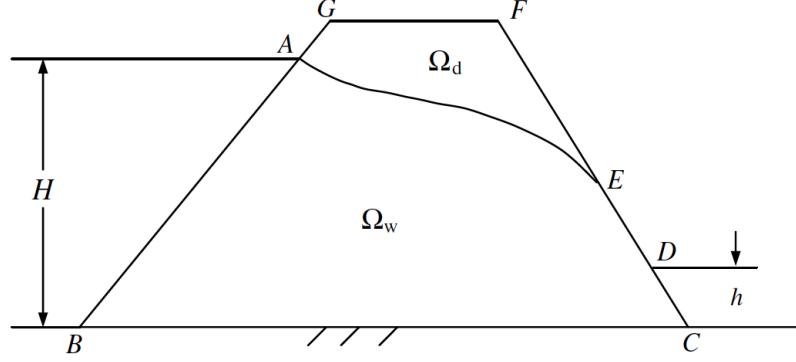


Figure 3.2: A simple earth-dam: Ω_w is the saturated part of the domain, and Ω_d the dry part. The line AE is unknown and is ED.

- Hydrostatic pressure downstream of the dam along DC:

$$h = h_D \quad \text{along } DC$$

- Seepage along ED - such that here the fluid pressure is equal to 0 (atmospheric)

$$h = y \quad \text{along } ED$$

- No outgoing liquid water along the external boundary of the dry zone:

$$q_n = 0 \quad \text{along } AG - GF - FE$$

The problem is that beside the internal boundary AE, we do not know a-priori the location of point E. It will be part of the solution. Let's consider the boundaries of the dam which are susceptible to seepage AG-GF-FD, we can distinguish between two cases:

1. seepage where the porous media is fully saturated $S_l = 1$ and $q_n < 0$ (we loose fluid) and the liquid must thus be at atmospheric pressure: $p = 0$ i.e. $h = y$. (this is the case of segment ED in the example discussed above)
2. dry part of the boundary where $S_l \sim 0$, then the liquid pressure is negative $p < 0$, i.e. $(h - y) < 0$ but its flux (of liquid water) is zero (we do not loose liquid from that face), i.e. $q_n = 0$. (this is the case of segments AGFE in the example discussed above)

We see that 1 and 2 above are corresponding to opposite cases. We can thus rewrite the conditions along a *potential* seepage boundary (AGFD in the example above) as

$$\begin{aligned} q_n &\leq 0 \ \& \ (h - y) \leq 0 \\ q_n(h - y) &= 0 \end{aligned}$$

We see that if $q_n = 0$ then $(h - y) \leq 0$, and if $q_n < 0$ we must have $(h - y) = 0$ (and vice versa). The condition $q_n(h - y) = 0$ is often called a complementary boundary conditions. As a result, we see that during the computations, we will have to iterate to specify the proper boundary conditions on different parts of the potential seepage boundary.

3.3 Solution of the corresponding non-linear problem

For the approach describe above, we solve on the complete domain Ω (say the complete geometrical dimension of an earth dam) the following PDE:

$$\begin{aligned} -(K_{ij} \times k_r(p) \times h_{,j})_{,i} &= \gamma \\ h &= \frac{p}{\rho_f g} + x_2 \equiv \frac{p}{\rho_f g} + y \\ k_r(p) &= \frac{1}{2} (1 + \tanh(p/p^*)) + \epsilon_M \end{aligned}$$

with in addition proper Dirichlet (head, pressure) and Neumann (flux etc.) boundary conditions (see the discussion above). We now see that the problem is now non-linear because of the relative permeability function (which is function of pore-pressure) .

The weak form - for a test function v (with $v = 0$ on the Γ_p the part of the boundary with Dirichlet BC), we have (see previous chapter)

$$\int_{\Omega} v_{,i} K_{ij} h_{,j} \times k_r(p) dV = \int_{\Gamma_q} v q^g dS + \int_{\Omega} v \times \gamma dV$$

Exactly, the same finite element discretization follow with the exception that now we have the effect of $k_r(p)$ in the element conductivity matrix expression: i.e.

$$\mathbf{C}^e(\mathbf{p}^e) = \int_{\Omega^e} (\nabla \mathbf{N})^T \cdot \mathbf{K} \cdot \nabla \mathbf{N} k_r(\mathbf{N} \mathbf{p}^e) dV$$

We have seen that when $k_r = 1$, for a linear 3 nodes triangle, a single gauss point is necessary to compute the corresponding integral. Now - if $k_r(p)$ is a linear function - the use of 3 gauss points would give exactly the same results than the use of a single Gauss point [think and derive why]. On the other hand if k_r is highly non-linear, one may want to use more than 1 Gauss point for better accuracy / convergence with respect to element size. Whatever, the element integration choice, upon assembly, we obtain a final linear system as

$$\mathbf{C}(\mathbf{h}) \cdot \mathbf{h} = \mathbf{f}$$

as the relative permeability depends on pressure (/head). More-precisely, upon application of the dirichlet boundary conditions we have

$$\mathbf{C}^{uu}(\mathbf{h}) \cdot \mathbf{h}^u = -\mathbf{C}^{uf}(\mathbf{h}) \cdot \mathbf{h}^f + \mathbf{f}^u$$

3.3.1 Solution of the non-linear system via fixed point iterations

Let's now look how we can solve such a non-linear system of the schematic form

$$\mathbf{A}(\mathbf{x}) \cdot \mathbf{x} = \mathbf{b}(\mathbf{x})$$

We have to bear in mind that **we do not know a-priori all the nodes along the potential seepage surface** where the liquid pressure will have to be set to atmospheric (i.e. 0) conditions.

3.3.1.1 Fixed point iterations scheme - known boundary conditions

We first assume in that sub-section that we know exactly the nodes where the liquid pressure have to be set. We will then devise in sub-section 3.3.1.3 a complete solver including an iteration on boundary conditions.

The simplest scheme to solve a non-linear system is to use “fixed point” iterations. Such a iterative scheme is very simple and easy to program.

1. Start from an estimate $\mathbf{x}^0 = 1$ such that $k_{rel} = 1$ for all elements - set $k = 0$ (where k will be the iterations count)
2. Repeat until convergence, i.e. $|\mathbf{x}^{k+1} - \mathbf{x}^k| < \epsilon |\mathbf{x}^{k+1}|$ (or until the maximum number of iterations has been reached) - e.g. While

$$k = k + 1$$

$$\mathbf{A}(\mathbf{x}^k) \cdot \mathbf{x}^{k+1} = \mathbf{b}(\mathbf{x}^k)$$

This algorithm requires of-course to recompute the conductivity matrix at each iteration k as the relative permeability is re-computed using the new estimate for pressure/head.

Although simple, the previous algorithm does not necessarily converge. In particular, it is prone to oscillations. A simple way to “help” convergence is to perform a so-called under-relaxation. Upon solution of the system

$$\mathbf{A}(\mathbf{x}^k) \cdot \mathbf{x}^{k+1} = \mathbf{b}(\mathbf{x}^k)$$

The new estimate is set to

$$\mathbf{x}^{k+1} = \beta \mathbf{x}^{k+1} + (1 - \beta) \mathbf{x}^k$$

with $\beta \in]0, 1]$. A value of $\beta = 1$ corresponds to the fixed point scheme without relaxation, whereas a small value of β corresponds to a very under-relaxed case. Such an under-relaxation of course drastically slows down convergence - but it also help stabilising the scheme. Typically, one will start using $\beta = 1$ and if oscillations / non-convergence are observed, the value of β is lowered. Here a value of $\beta = 0.75$ has proven adequate for most examples I have tried.

3.3.1.2 Newton-Raphson scheme - known boundary conditions

[Interlude - not really needed here and not used during the exercices]

The previous fixed-point scheme is very simple to code and sometimes - amazingly produce very good results (this is the case here ;)). Another scheme of choice is a generalization of Newton's method (a lot of variant exist). The non-linear system is re-casted into finding the root of the following system of equations

$$\mathcal{R}(\mathbf{x}) = \mathbf{A}(\mathbf{x}) \cdot \mathbf{x} - \mathbf{b}(\mathbf{x}) = 0$$

Writing a first-order Taylor expansion around iterate k

$$\mathcal{R}(\mathbf{x}) = \mathcal{R}(\mathbf{x}^k) + \nabla \mathcal{R}(\mathbf{x}^k) \cdot \Delta \mathbf{x} + O(\Delta \mathbf{x}^2)$$

enforcing $\mathcal{R}(\mathbf{x}^{k+1}) = 0$, one obtains the following linear system to solve for the increment $\Delta \mathbf{x}^{k+1} = \mathbf{x}^{k+1} - \mathbf{x}^k$

$$0 = \mathcal{R}(\mathbf{x}^k) + \nabla \mathcal{R}(\mathbf{x}^k) \cdot \Delta \mathbf{x}^{k+1}$$

i.e. writing the update explicitly

$$\mathbf{x}^{k+1} = \mathbf{x}^k - (\nabla \mathcal{R}(\mathbf{x}^k))^{-1} \cdot \mathcal{R}(\mathbf{x}^k)$$

where the Jacobian matrix $\nabla \mathcal{R}(\mathbf{x}^k)$ is defined as

$$\frac{\partial \mathcal{R}_i}{\partial x_j}$$

Important points

- The Newton-Raphson will exhibit quadratic convergence - if the functional is convex it will always converge. In the general case it will only converge if the first guess is not “too far” from the solution (think of a function with multiple minimum etc.).
- As one can see, it requires the solution of a linear system at each iterations (like the fixed point scheme) but it also requires the construction of the Jacobian matrix. If the algorithm always uses only the first estimate of the Jacobian matrix (for $\mathbf{x}^0$), it is known as the secant method (and does not converge as fast).
- Computing the Jacobian matrix explicitly is sometimes difficult (and requires additional programming). It can in some cases be done numerically via finite difference but this is neither very efficient nor robust. It is important to consistently obtained the so-called “tangent” operator. This is done by differentiating the weak form at the element level, and assemble the tangent matrix. We will rediscuss this technique for the solution of elastoplastic problem later in the course.
- A large number of variants exist, I refer you to your numerical analysis & numerical optimization classes.

3.3.1.3 Complete fixed-point solver including the location of the seeping nodes

We recall that we do not know a-priori which part of the boundary will exhibit seepage (where $q_n < 0$ and $h = y$) and which one has zero flux. We thus need to incorporate in the iterative scheme a test on the constraints along the nodes of the possible seepage face. This can be easily done within the fixed point algorithm.

In the following, we denote as f^o the set of nodes with known dirichlet boundary conditions - that do not change during iterations -e.g. the nodes on the upstream side of a dam for example (segment AB in Fig. 3.2). We denote ps the set of nodes on the potential seepage boundary, s the set of nodes with seepage (for which $q_n < 0$) and us the set of nodes on the unsaturated boundary (where the head is unknown and $q_n = 0$). Of course

$$ps = s \cup us \quad s \cap us = \emptyset$$

During the iterations, we will look for the nodes s where we need to set the dirichlet boundary condition $h = y$. A natural Neumann boundary conditions is specified on us (zero flux).

The overall iterative algorithm reads.

1. Start from an estimate $\mathbf{x}^o = 1$ such that $S_l = k_{rel} = 1$ for all elements - set $k = 0$ (k being the iterations count) and start with $s^o = \emptyset$ and $us^o = ps$ (this means that all nodes on the boundary of the potential seepage boundary have a zero flux boundary conditions,
2. Repeat until convergence, i.e. $|\mathbf{x}^{k+1} - \mathbf{x}^k| < \epsilon |\mathbf{x}^{k+1}|$ (or until the maximum number of iterations has been reached) - e.g. While

(a)

$$k = k + 1,$$

adjust the nodes where the head is to be fixed, i.e the set of nodes $f^k = f^o \cup s^k$

(b) Solve

$$\mathbf{C}^{uu}(\mathbf{h}^k) \cdot \mathbf{h}^u = -\mathbf{C}^{uf^k}(\mathbf{h}^k) \cdot \mathbf{h}^{f^k} + \mathbf{f}^u$$

(c) Perform an under-relaxation $\mathbf{h}^{k+1} = \beta[\mathbf{h}^u, \mathbf{h}^{f^k}] + 1(1 - \beta)\mathbf{h}^k$,

(d) Check the constraints - find the set of nodes s^{k+1} as the set of nodes on the potential failure surface (set ps) for which $h - y > 0$ and $q_n < 0$ (saturated and with a outgoing flux):

- i. For all nodes in us^k (with zero flux), if $h - y > 0$, move this nodes to s^{k+1}
- ii. For all nodes in s^k (seepage boundary), if $q_n > 0$ (influx) move back this nodes to us^{k+1} (unsaturated part of the boundary).

[Note that in principle, nothing ensure convergence of such a scheme ! the set of nodes s^k may oscillate and never converge - this has to be checked very carefully. More robust scheme from the field of constrained optimization may then be used.]

Note that in order to check the constraint $q_n < 0$ vs $q_n > 0$, we can simply check the sign of the nodal fluxes / “forces”, i.e. $\mathbf{C} \cdot \mathbf{h}$. [check the weak form derived in the previous chapter to be convinced].

Chapter 4

Transient flow

We now move to problems where the pore-pressure / head will vary in *both* space and time. This class of problem are important when fluids are injected or retrieved from wellbores at depth, when a dam is slowly filled etc. In other words when transient phenomena can not be neglected. I restrict the notes to isotropic permeability for simplicity. The extension to a permeability tensor is straightforward.

4.1 Governing equations

Let us restart by recalling the fluid mass conservation in a unit RVE of a porous media:

$$\frac{\partial \rho_f \phi}{\partial t} + (\rho_f q_i)_{,i} = \rho_f \gamma \quad i = 1, 2 (\text{in 2D}), 3 (\text{in 3D}) \quad (4.1)$$

where ρ_f is the fluid density, ϕ the porous medium porosity, q_i the fluid flux vector and γ a possibly existing source/sink term (for example due to an injection). We recall that q_i is the filtration vector (also called specific discharge or fluid flux). It describes the fluid motion with respect to the solid, it's the phase-slip velocity weighted by the fluid volume in the RVE (i.e the porosity):

$$q_i = \phi \left(v_i^f - v_i^s \right)$$

where v_i^f and v_i^s are the fluid and solid velocity respectively. It is related to the gradient of pore-pressure (and buoyancy fluid forces) via Darcy's law (laminar flow in porous media)

$$q_i = -\frac{k}{\mu_f} (p + \rho_f g x_3)_{,i} \quad (4.2)$$

4.1.1 Storage coefficient / diffusion equation

For the steady state case, we directly dropped the time-derivative. Let's now restart from (4.1) and under the assumption of a slightly compressible liquid for which $\rho_f \approx \rho_f^o(1 + \beta_f(p - p^o))$, we obtain

$$\frac{\partial \phi}{\partial t} + \phi \beta_f \frac{\partial p}{\partial t} + (q_i)_{,i} = \gamma \quad i = 1, 2 (\text{in 2D}), 3 (\text{in 3D})$$

Now, the porosity change as one can guess will be related to the mechanical deformation of the porous medium. Taking **for now** the simple view that the variation of porosity is solely dependent on the variation of fluid pore pressure p , we can write a linear “equation of state” for porosity as

$$\begin{aligned} \phi &= \phi_o(1 + \beta_\phi(p - p^o)) \\ d\phi &= \phi_o \beta_\phi dp \end{aligned}$$

where β_ϕ is the pore-compressibility. In doing so, we have basically neglected any contribution coming from the volumetric strain of the porous medium: imagine the case where a mechanical load is applied to a porous medium but the fluid is let to drain such that it's pore-fluid pressure does not change. With the previous relation, the porosity would not change ! We will come back to that when studying poroelasticity. For now, we keep such a simplification, and can define a storage coefficient (of dimension inverse of pressure) :

$$S = \phi_o \beta_\phi + \phi_o \beta_f$$

and the fluid conservation becomes (neglecting the second order term $\phi_o \beta_f \beta_\phi (p - p^o) \ll \phi_o \beta_f$), we obtain

$$S \frac{\partial p}{\partial t} + q_{i,i} = \gamma$$

and introducing Darcy (restricting to 2D with gravity along $x_2 = y$ and isotropy)

$$S \frac{\partial p}{\partial t} - \left(\frac{k}{\mu_f} (p + \rho_f g x_2)_{,i} \right)_{,i} = \gamma$$

We can re-write it directly in term of piezometric head as

$$S \frac{\partial h}{\partial t} - \left(\frac{k}{\mu_f} h_{,i} \right)_{,i} = \gamma$$

Note also, that in a large number of cases interesting to solve for change in pressure with respect to an initial state $p = p^o$ which satisfy the steady-state flow equations, i.e we can then re-write the previous equation as

$$S \frac{\partial p}{\partial t} - \left(\frac{k}{\mu_f} (p - p^o)_{,i} \right)_{,i} = \gamma$$

For a medium with uniform properties (S and k), it further reduces to

$$\frac{\partial p}{\partial t} - c(p - p^o)_{,ii} = \gamma/S$$

with the following diffusivity coefficient (m^2/s)

$$c = \frac{k}{\mu_f S}$$

The possible boundary conditions are essentially the same than for the steady-state flow case (prescribed pressure/head, or flux) but can vary in time.

The form of the PDE obtained belong to the class of second order parabolic PDEs, and is often refereed to as a diffusion equation, i.e. here it is not a specie that diffuses but fluid pressure. A number of analytical solutions for simple geometries and boundary conditions do exist for this class of boundary value problems - see the textbooks of Carslaw & Jaeger (1959), Crank (1979).

4.2 Numerical Solution by Finite difference

4.2.1 A one-dimensional problem

Let's first look at the following 1D dimensional diffusion equation arising for example in the 1D consolidation problem. Here initially a segment of length $2L$ is under uniform pore-pressure p_o . It is then let to drain from $t = 0$ at both ends. The problem can be re-written from 0 to L using a condition of symmetry at $x = L$. The problem reads as follow:

$$\begin{aligned} \frac{\partial p}{\partial t} - c \frac{\partial^2 p}{\partial x^2} &= 0 \\ p(x = 0, t^+) &= 0 \\ \frac{\partial p}{\partial x}(x = L, t) &= 0 \\ p(x, t = 0) &= p_o \end{aligned}$$

where p_o is the initial pore-fluid pressure before the fluid is let to drain from the top surface at $t = 0^+$. Note that this equation can be written generally in conservative form

$$\frac{\partial p}{\partial t} + \frac{\partial v}{\partial x} = 0 \quad v = -c \frac{\partial p}{\partial x}$$

v is a “flux”. The first equation is for example the mass conservation, and the second Darcy's law (after dividing it by the storage coefficient).

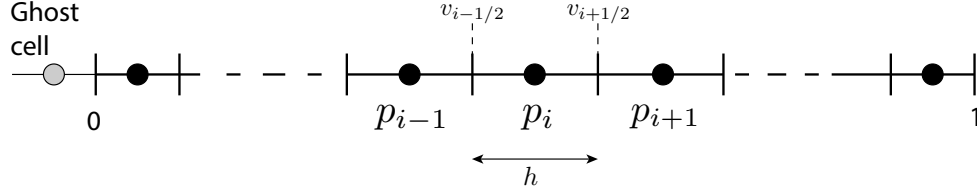


Figure 4.1: 1D finite difference (/ finite volume) grid - constant element size - pressure uniform in each element (piece-wise constant interpolation).

4.2.2 Spatial discretization

Let's discretize the spatial interval $[0, L]$ with N control volume of size h (see Fig. 4.1). Within a control volume (here a segment in 1D), we will assume the pressure is uniform. Such an approximation is often denoted piece-wise constant (or P0). It is the lowest approximation order possible. The pressure value is located at the mid-point of that control volume. Let's first integrate the conservation equation over the control volume i , we have (Green's theorem):

$$h \frac{\partial p_i}{\partial t} + v_{i+1/2} - v_{i-1/2} = 0$$

Moreover the flux $v_{i+1/2}$ between the control volume i and $i + 1$ can be approximated by a first order finite difference,

$$\begin{aligned} v_{i+1/2} &= -c \frac{p_{i+1} - p_i}{h} \\ v_{i-1/2} &= -c \frac{p_i - p_{i-1}}{h} \end{aligned}$$

We then obtain

$$h \frac{\partial p_i}{\partial t} - \frac{c}{h} (p_{i-1} - 2p_i + p_{i+1}) = 0$$

This exactly what would have been obtained using directly a second order finite difference for the second order derivative for pressure. However, the steps lay out above can be applied easily to non-uniform mesh (control volume of different size) and generalized to 2D and 3D with mesh which are non-necessarily Cartesian.

Specification of the initial and boundary conditions Note that to account for the boundary condition at $x = 0$, we impose in a so-called “ghost” control volume $i = -1$ the pressure directly to zero. In Finite difference such a “ghost-cell” are typically used to specify Dirichlet boundary conditions. Here it is zero, so the equation for the first cell on the left (here $i = 1$) just becomes

$$h \frac{\partial p_1}{\partial t} - \frac{c}{h} (-2p_1 + p_2) = 0$$

i.e. we do not need to model the ghost cell per se (hence its name).

For the last control volume, we will make sure that its edge corresponds to $x = L$ such that the no-flow boundary conditions can be directly taken into account, i.e. for that last control volume, we have

$$h \frac{\partial p_N}{\partial t} - \frac{c}{h} (p_{N-1} - p_N) = 0$$

The initial condition is that for all control volume $i = 1 - N$, we have $p_i(t = 0) = p_o$ and for the first control volume, we have $p_1(t > 0) = 0$.

Matrix-vector form We can rewrite the previous spatial discretization in a format with a vector of unknowns $\mathbf{p}$ and a finite difference matrix $\mathbb{L}$ (tri-banded matrix), i.e. a set of equations (one for each control volume):

$$\frac{\partial \mathbf{p}}{\partial t} = c \mathbb{L} \mathbf{p}$$

this actually only a semi-discretization of the problem, as we have kept the time derivatives continuous. Let's now develop a time-integration scheme. Note that $\mathbb{L}$ has the form

$$\mathbb{L} = \frac{1}{h^2} \begin{pmatrix} -2 & 1 & & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & \ddots & \ddots & \ddots \\ & & & & 1 & -1 \end{pmatrix} = \frac{\mathbb{D}}{h^2}$$

From now on, we will use the matrix form $\mathbb{L} = \frac{\mathbb{D}}{h^2}$ to illustrate the effect of cell size h when moving on the time-integration part of the problem.

Note also the extension to account for a non-zero flux conditions or non-zero pressure boundary conditions are straightforward. For given flux, a “force” like vector will appear with the prescribed flux (with proper constant $1/h$) on the row corresponding to the control volume next to the boundary (this can also be generalized easily to internal sources). Fixed pressure boundary conditions are treated similarly than fixed displacement in elasticity.

Important point It is very important to note that the discretization approach that we have outlined above is based on the integration of the mass conservation in a strong form over a control volume. **It is therefore locally conservative.** In other words, the fluid mass is conserved up to machine precision in all the cells, and as a result in all the domain. This is a very nice properties which render this type of scheme popular in the fluid mechanics community and in reservoir / groundwater flow applications, where having a scheme which does not “lose/gain” mass locally is essential (i.e. proper estimation of flow rates from well etc.).

4.2.3 Time integration

A popular scheme for the time-integration of the previous equation is the so-called θ -method. In what follows we will denote as $\mathbf{p}^n$ the solution at time t^n (assumed to be known), and $\mathbf{p}^{n+1}$ the solution at time $t^{n+1} = t^n + \Delta t$ that we are seeking to obtain over one time step of size Δt . We approximate the previous equation using a finite difference in time for the LHS and use for the estimate of the unknowns in the RHS, a linear combination of the solution at t^n and t^{n+1} weighted by a scalar $\theta \in [0, 1]$, i.e. the systems of equations now becomes

$$(\mathbf{p}^{n+1} - \mathbf{p}^n) = \frac{\Delta t \times c}{h^2} \mathbb{D} (\theta \mathbf{p}^{n+1} + (1 - \theta) \mathbf{p}^n)$$

which we can rewrite:

$$\left(\mathbb{I} - \theta \frac{\Delta t \times c}{h^2} \mathbb{D} \right) \mathbf{p}^{n+1} = \left(\mathbb{I} + (1 - \theta) \frac{\Delta t \times c}{h^2} \mathbb{D} \right) \mathbf{p}^n$$

or alternatively, working in increment $\Delta \mathbf{p} = \mathbf{p}^{n+1} - \mathbf{p}^n$, i.e. :

$$\left(\mathbb{I} - \theta \frac{\Delta t \times c}{h^2} \mathbb{D} \right) \Delta \mathbf{p}^n = \frac{\Delta t \times c}{h^2} \mathbb{D} \mathbf{p}^n$$

If $\theta = 0$, we recover the forward Euler / explicit scheme, and if $\theta = 1$, the backward Euler / implicit scheme.

Explicit scheme For the explicit scheme ($\theta = 0$), we see that we do not need to solve any linear system at each time step as:

$$\mathbf{p}^{n+1} = \left(\mathbb{I} + \frac{\Delta t \times c}{h^2} \mathbb{D} \right) \mathbf{p}^n$$

or

$$\Delta \mathbf{p}^n = \frac{\Delta t \times c}{h^2} \mathbb{D} \mathbf{p}^n$$

i.e. only a matrix vector multiplication is necessary to advance in time. So we see that if we start from a solution $\mathbf{p}^o$, the solution at time t^k is simply (by recursion)

$$\mathbf{p}^k = \left(\mathbb{I} + \frac{\Delta t \times c}{h^2} \mathbb{D} \right)^k \mathbf{p}^o$$

if the boundary conditions are of drainage, from a constant initial solution, we physically understand that at large/infinite time, the pressure must converge to zero. This is only possible (analysis of a convergent matrix) if the spectral radius of $\mathbb{I} + \frac{\Delta t \times c}{h^2} \mathbb{D}$ (i.e. the

absolute of the maximum eigenvalue) is lower than one. The maximum eigenvalue of $\frac{\Delta t \times c}{h^2} \mathbb{D}$ is $-2c \frac{\Delta t}{h^2}$ such that the condition reads

$$\|1 - 2c \frac{\Delta t}{h^2}\| < 1$$

in other words, we arrive at the following condition:

$$2c \frac{\Delta t}{h^2} < 1 \Leftrightarrow \Delta t < \frac{h^2}{2c}$$

This is so-called CFL condition for the diffusion equation (CFL for Courant-Friederichs-Lewy). If the time-steps violates such a condition, the explicit scheme is unstable (i.e. it blows up). The forward Euler / explicit method is thus said to be conditionally stable.

Implicit Scheme For the implicit scheme ($\theta = 1$), we have a linear system to solve at each time-step

$$\left(\mathbb{I} - \frac{\Delta t \times c}{h^2} \mathbb{D} \right) \mathbf{p}^{n+1} = \mathbf{p}^n$$

or

$$\left(\mathbb{I} - \frac{\Delta t \times c}{h^2} \mathbb{D} \right) \Delta \mathbf{p}^n = \frac{\Delta t \times c}{h^2} \mathbb{D} \mathbf{p}^n$$

In that case, if we start from a solution $\mathbf{p}^o$, the solution at time t^k is simply (by recursion)

$$\mathbf{p}^k = \left[\left(\mathbb{I} - \frac{\Delta t \times c}{h^2} \mathbb{D} \right)^{-1} \right]^k \mathbf{p}^o$$

the stability condition is now (again the max eigenvalue of $\frac{\Delta t \times c}{h^2} \mathbb{D}$ is $-2c \frac{\Delta t}{h^2}$)

$$\|1 + 2c \frac{\Delta t}{h^2}\|^{-1} < 1$$

which is always satisfied as time step, mesh size and diffusivity are all positive numbers. The implicit scheme is therefore unconditionally stable.

θ -method It can be shown that the θ - *method* is unconditionally stable for $\theta > 1/2$, conditionally stable otherwise. As far as accuracy, it can be shown that the local truncation error is of the order $\Delta t + h^2$ if $\theta \neq 1/2$ and in $\Delta t^2 + h^2$ for $\theta = 1/2$ (Crank-Nicholson scheme: both unconditionally stable and second order accurate in time).

4.3 Finite Element

Let's come back to the general form of the diffusion equation obtained for transient flow (still neglecting the mechanical coupling), i.e. a Darcy type flow

$$S \frac{\partial p}{\partial t} + q_{i,i} = \gamma \quad q_i = -\kappa p_{,i}$$

where here $\kappa = k/\mu_f$ (restricting to isotopy).

4.3.1 Weak form

Now, let's introduce a "test" pressure field r having the same continuity properties that p (and such that $r = 0$ on the boundary with Dirichlet boundary conditions), multiply the previous equation and integrate over the domain of interest Ω (here still $[0, L]$) of boundary Γ . We thus obtain the weak form (using Green's theorem)

$$\int_{\Omega} r S \frac{\partial p}{\partial t} dV + \int_{\Gamma} r q_i n_i dS - \int_{\Omega} r_{,i} q_i d\Omega = \int_{\Omega} r \gamma d\Omega$$

where n_i is the outward normal. After replacing Darcy's law:

$$\int_{\Omega} r S \frac{\partial p}{\partial t} dV + \int_{\Omega} r_{,i} \kappa p_{,i} d\Omega = \int_{\Omega} r \gamma d\Omega - \int_{\Gamma_q} r q_i n_i dS$$

This for any r field with proper continuity requirements and satisfying $r = 0$ on the boundary. We can now use the Finite element Method on the previous weak form, dividing the domain in a non-overlapping set of elements. More specifically, we use a continuous Galerkin formulation, where the pressure unknowns are located at the vertex of each elements (i.e. at the node), and the shape functions are defined on each element and equal to 1 at the corresponding node and zero at the other nodes.

We can then perform the numerical integration of the different integrals over each element, and correctly map the unknowns degree of freedoms to obtain the global system of equations. We obtain

$$\mathbf{r}^T \mathbf{M} \partial_t \mathbf{p} + \mathbf{r}^T \mathbf{C} \mathbf{p} = \mathbf{r}^T \mathbf{f}$$

for any $\mathbf{r}$. In other words, we obtain the following system

$$\mathbf{M} \frac{\partial \mathbf{p}}{\partial t} + \mathbf{C} \mathbf{p} = \mathbf{f} \quad (4.3)$$

where the mass and conductivity matrices $\mathbf{M}$ and $\mathbf{C}$ is at the element level are respectively

$$\begin{aligned} \mathbf{M}^e &= \int_{\Omega_e} S \mathbf{N}^T \cdot \mathbf{N} dV \\ \mathbf{C}^e &= \int_{\Omega_e} (\nabla \mathbf{N})^T \cdot \kappa \cdot \nabla \mathbf{N} dV \end{aligned}$$

and

$$\mathbf{f} = \sum_e \int_{\Omega_e} \mathbf{N}^T \gamma \, dV - \sum_{\text{on boundary}} \int_{\Gamma_q^e} \mathbf{N}_{\mathbf{d}-1}^T(x) q_i n_i \, dS$$

where the notation $\mathbf{N}_{\mathbf{d}-1}$ denote the shape function with dimension $d - 1$ (i.e. 1D for a 2D problem, 2D for a 3D problem).

Dirichlet boundary conditions are imposed in the same way than for the steady-state case (but now can vary in time)

Only globally conservative

It is important to note that the previous continuous Galerkin finite element method satisfying the mass conservation in a weak form globally. Contrary to a finite volume method, mass conservation is not enforced over each elements. This FE scheme is not strictly locally mass conservative.

4.3.2 Time integration

For time integration, similar scheme than for the Finite Difference / Finite Volume method described in section 4.2 can be used with the same restrictions with regards to the critical time-step. The system (4.3) discretize using the θ -method becomes:

$$\mathbf{M} \frac{\mathbf{p}^{n+1} - \mathbf{p}^n}{\Delta t} = \theta (\mathbf{f}^{n+1} - \mathbf{Cp}^{n+1}) + (1 - \theta) (\mathbf{f}^n - \mathbf{Cp}^n)$$

which can be rewritten in terms of update $\mathbf{p}^{n+1} = \Delta \mathbf{p} + \mathbf{p}^n$

$$(\mathbf{M} + \theta \Delta t \mathbf{C}) \Delta \mathbf{p} = -\Delta t \mathbf{Cp}^n + \Delta t \theta \mathbf{f}^{n+1} + \Delta t (1 - \theta) \mathbf{f}^n$$

The restriction for the time step are the same for the finite difference approximation (here we recall that $c = \kappa/S$):

- For $\theta < 1/2$ ($\theta = 0$ explicit), the scheme is conditionally stable, i.e the time step must be smaller than the following C-F-L

$$\Delta t < \frac{h^2}{2c}$$

- For $\theta \in [1/2 - 1]$, the scheme is unconditionally stable, and is second order accurate in time for $\theta = 1/2$

We see that even for the explicit case $\theta = 0$, one needs to solve a linear system as the mass matrix $\mathbf{M}$ (although sparse) is not diagonal as a result of the finite element discretization (this also why finite volume methods are often preferred). Similarly than in the elasto-dynamics case, an approach called as “mass lumping” is often used: it consists in summing up all entries of a row of the mass matrix to the diagonal - i.e. moving the effect of the mass from element to the nodes. It is not strictly correct and care must be used in using this approach for diffusion (and even more so for advection-diffusion problem) - see Guermond & Pasquetti (2013) for additional details and adequate corrective measures.

Chapter 5

Quasi-static Poroelasticity I

5.1 Material Description

Let consider a REV of porous material of porosity ϕ filled with a fluid of compressibility c_f and density ρ_f . We put ourselves at the continuum scale and view the porous medium as the superposition of 2 continua (solid and fluid) following Biot's formalism (Rice & Cleary 1976, Detournay & Cheng 1993). The mechanical behaviour of the porous media is characterized by two stress quantities: the total stress tensor σ_{ij} (total stress acting on the REV) and the fluid pore pressure p (i.e. we neglect shear stress in the fluid). In parallel, two “strain” like quantity describe the 'kinematic' evolution of the porous media, namely the porous solid strain ϵ_{ij} and the variation of fluid content ζ both are defined with respect to an initial configuration (i.e. Lagrangian variables). We will assume small-strain throughout. The variation of fluid content from the initial configuration is defined as

$$\zeta = \frac{m_f - m_f^o}{\rho_f^o} \quad (5.1)$$

where the fluid mass per unit volume m_f is simply

$$m_f = \rho_f \phi$$

For sake of clarity in exposition, in this chapter, we will assume that the initial configuration is stress-free: i.e. $\sigma_{ij}^o = 0$, $p^o = 0$. If the material is not stress-free in the initial/reference configuration, the linear poroelastic constitutive law can always be rewritten in terms of variation with respect to this initial state, i.e. $\sigma_{ij} - \sigma_{ij}^o$ and $p - p^o$. This is what will be typically done in geomechanical applications.

Note that a positive ζ indicates a gain in fluid while a positive strain indicates extension. Stresses are assumed here positive in tension (negative in compression) and pore pressure has the usual convention (positive pressure is compressive). Note that in geomechanical applications, for clarity as the state of stress in the earth is compressive, practionners use the convention of positive stress in compression (negative strain in extension).

5.2 Constitutive law

Two mechanisms play a key role in the deformation of a poroelastic material: i) an increase of pore pressure induces dilation of the rock (if not restrained) and ii) compression of the rock causes a rise of pore-pressure if the fluid is prevented from escaping the pore network. Under the assumption of linear elasticity (reversibility), we see that both the total strain and variation of fluid content depend linearly on both the pore-pressure and stresses. We can thus write:

$$\begin{aligned}\epsilon_{ij} &= \frac{1+\nu}{E}\sigma_{ij} - \frac{\nu}{E}\sigma_{kk}\delta_{ij} + \frac{p}{3H'}\delta_{ij} \\ \zeta &= \frac{\sigma_{kk}}{3H''} + \frac{p}{R'}\end{aligned}$$

where E and ν are the elastic constant of the porous solid (skeleton), H' , H'' and R' are poroelastic constants characterizing the coupling between pore fluid and solid. The hypothesis of reversibility implies that the work increment (infinitesimal strain work density)

$$dW = \sigma_{ij}d\epsilon_{ij} + pd\zeta = \epsilon_{ij}d\sigma_{ij} + \zeta dp$$

is an exact differential. The Euler conditions therefore gives

$$\frac{\partial \epsilon_{ij}}{\partial p} = \frac{\partial \zeta}{\partial \sigma_{ij}}$$

which imply that $H'' = H'$.

It is interesting in practice to express the constitutive law in a volumetric and deviatoric part (i.e. the coupling only appears in the volumetric response). Let e_{ij} and s_{ij} denote the deviatoric strain and stress respectively. Note that the volumetric strain is $\epsilon^v = \epsilon_{kk}$ and the mean stress is $\sigma = \sigma_{kk}/3$ ($s_{ij} = \sigma_{ij} - \sigma_{kk}/3\delta_{ij}$, $e_{ij} = \epsilon_{ij} - \epsilon^v/3\delta_{ij}$). We obtain

$$\begin{aligned}\epsilon_{kk} &= \frac{3(1-2\nu)}{E}\sigma + \frac{p}{H'} & e_{ij} &= \frac{1+\nu}{E}s_{ij} \\ \sigma &= H' \left(\zeta - \frac{p}{R'} \right)\end{aligned}$$

We recognize that $\frac{E}{3(1-2\nu)} = K$ is the bulk modulus of the porous solid (case without fluid), and $\frac{E}{1+\nu} = 2G$ is the shear modulus. Let's now only focus on the volumetric part.

Let's try to re-write ζ and the mean stress as function of the volumetric strain and pore pressure. We obtain:

$$\begin{aligned}\zeta &= \frac{K}{H'}\epsilon_{kk} + \frac{H'^2 - KR'}{H'^2 R'}p \\ \sigma &= K\epsilon_{kk} - \frac{K}{H'}p\end{aligned}$$

It is useful to define the following constants which are mostly used in practice: the Biot's coefficient α and Biot's modulus M as follow

$$\alpha = \frac{K}{H'} \quad \frac{1}{M} = \frac{H'^2 - KR'}{H'^2 R'}$$

such that the poroelastic constitutive law can be expressed as:

$$\sigma = K\epsilon_{kk} - \alpha p \quad s_{ij} = 2Ge_{ij} \quad (5.2)$$

$$\zeta = \alpha\epsilon_{kk} + \frac{1}{M}p \quad (5.3)$$

where G is the elastic shear modulus $2G = E/(1 + \nu)$. This is the form that we will be using mostly. Note that the total stress tensor σ_{ij} can be expressed as:

$$\sigma_{ij} = 2G\epsilon_{ij} + (K - 2/3G)\epsilon_{kk}\delta_{ij} - \alpha p\delta_{ij}$$

5.2.0.1 Variation of porosity, factoring out the fluid type

Let's denote $\varphi = \phi - \phi_o$ the variation of porosity from the reference state. Now, let's recall the definition of the variation of fluid content Eq.(5.1), which we can re-write as.

$$\zeta = \frac{\rho_f \phi - \rho_f^o \phi_o}{\rho_f^o} = \left(1 + \frac{\Delta \rho_f}{\rho_f^o}\right) (\phi_o + \varphi) - \phi_o$$

Neglecting the second order term $\varphi \frac{\Delta \rho_f}{\rho_f^o}$, we obtain

$$\zeta = \varphi + \phi_o \frac{\Delta \rho_f}{\rho_f^o}$$

For a slightly compressible liquid, the fluid density changes relates to pressure as $\frac{\Delta \rho_f}{\rho_f^o} = c_f p$, thus $\zeta = \varphi + \phi_o c_f p$, and we can relate the variation of porosity to variation of volumetric strain and pore pressure:

$$\varphi = \alpha \epsilon^v + \underbrace{\left(\frac{1}{M} - \phi_o c_f\right)}_{1/N} p$$

N is a modulus which directly relates variation of pore-pressure to variation of porosity (under restrained condition). Contrary to the Biot's modulus M , it is independent of fluid compressibility (i.e. of the type of fluid): it is an intrinsic properties of the porous solid¹. Factoring out the fluid type is of great use when investigating more complex behaviour of

¹we shall denote N as the intrinsic Biot modulus for sake of clarity

porous solids (and using the thermodynamics of porous media), especially in unsaturated conditions (see the work of Coussy (2004, 2010, 2007)).

Intuitively, we see that $\alpha > 0$: under conditions where pore pressure does not change, the variation of porosity can not be negative if the volumetric strain is positive and similarly $N > 0$: a pore pressure increase implies an increase of porosity under zero volumetric strain (e.g. case of a restrained sample): hence $1/M > \phi_o c_f$.

5.2.1 Drained / undrained conditions

Two limiting conditions for the deformation of porous material are interesting to grasp. Let consider a porous material, and let's apply a total isotropic stress to it (a compressive stress $\sigma < 0$).

5.2.1.1 drained (or “slow” loading)

Let first assume that the fluid is allowed to escape (drained conditions) such that no excess of pore-pressure occur $p = 0$. In that case, the volumetric strain is simply

$$\epsilon^v = \sigma / K$$

and the variation of fluid content equals the variation of porosity:

$$\zeta = \varphi = \alpha \epsilon^v = \alpha \sigma / K$$

It is clear -under such loading - that the variation of porosity can not be larger than the volumetric change of the element, at most it is equal, thus $0 < \alpha < 1$.

5.2.1.2 undrained (or “fast” loading)

Let now assume that the fluid is not allowed to escape during the transformation such that the variation of volume content is null: $\zeta = 0$. This is an undrained condition. Therefore,

$$p = -\alpha M \epsilon^v$$

the porous material will compress (under compressive stress) $\epsilon^v < 0$ and the pore pressure increases $p > 0$ ($\alpha > 0$). We can relate the applied stress to deformation as:

$$\sigma = (K + \alpha^2 M) \epsilon^v = K_u \epsilon^v$$

and the pore pressure to the applied mean stress as:

$$p = -\frac{\alpha M}{K + \alpha^2 M} \sigma = -B \sigma$$

. It is useful to introduce an undrained bulk modulus $K_u = K + \alpha^2 M$ and the Skempton coefficient $B = \frac{\alpha M}{K + \alpha^2 M} = \frac{\alpha M}{K_u} = \frac{K_u - K}{\alpha K_u}$. The undrained response always appear stiffer (due to the fluid “stiffness”) compared to the drained response, $K_u \in [K, \infty[$. Note that nothing prevent B to be larger than unity.

5.2.2 Simple micromechanic relations

[For completeness - we will not cover this any further during the course - skip during a first read - continue to section 5.2.4]

Our aim is here to obtain simple expression of the poroelastic coupling parameters - namely the intrinsic Biot modulus $1/N$ and the Biot coefficient α - as function of the porosity of the porous solid and the bulk modulus k_s of the solid phase of the porous skeleton. Again, we restrict to the volumetric behaviour.

The total volumetric strain of the porous material at the REV scale is just the average of the solid and fluid phase volumetric strain:

$$\epsilon^v = (1 - \phi_o)\epsilon_s^v + \underbrace{(\phi - \phi_o)}_{\varphi}$$

Similarly, the total volumetric stress σ in the REV is just the volume average of the micro-scale stress in the solid and fluid phases:

$$\sigma = (1 - \phi_o)\sigma_s - \phi_o p$$

(the fluid phase is under uniform pressure p in the REV by definition of the poroelastic model). The constituent of the porous solid is linear elastic and the solid phase stress is directly to its strain: $\sigma_s = k_s \epsilon_s$ ($k_s \geq K$, as the bulk modulus of the porous solid K can only be equal or less to the one of its solid constituent). Therefore, we have

$$\sigma = k_s(\epsilon^v - \varphi) - \phi_o p$$

Introducing the macroscopic poroelastic relation for the variation of porosity, we obtain:

$$\sigma = k_s(1 - \alpha)\epsilon^v - \left(\frac{k_s}{N} + \phi_o\right)p$$

Comparing the previous equation with the macroscopic poroelastic expression for the total stress as function of strain and pore pressure, we obtain:

$$\begin{aligned} K &= k_s(1 - \alpha) \\ \alpha &= \frac{k_s}{N} + \phi_o \end{aligned}$$

which we can re-write as:

$$\alpha = 1 - \frac{K}{k_s}$$

$$\frac{1}{N} = \frac{\alpha - \phi_o}{k_s} = \frac{k_s(1 - \phi_o) - K}{k_s^2}$$

[Note we can see that $\alpha > \phi_o$, ($N > 0$) and we recover that $\alpha \leq 1$]. We can further obtain:

$$\frac{1}{M} = \frac{k_s(1 - \phi_o) - K}{k_s^2} + \phi_o c_f$$

$$K_u = K + \alpha^2 \frac{k_s^2}{k_s(1 - \phi_o) - K + k_s^2 \phi_o c_f}$$

Limiting cases

- Incompressible solid constituent ($k_s \gg K$): $k_s \rightarrow \infty$, we have $\alpha = 1$ and $1/N = 0$. Therefore $1/M = \phi_o c_f$. This is the limiting case of soils. It is also the limit of large RVE of rock mass (RVE with a length-scale 5 to 10 meters) where a large number of defects (fractures at different scales) are such that the bulk modulus K of the porous medium at the scale of the RVE is much lower than the bulk modulus of its constituent.
- Incompressible solid and fluid: $k_s \rightarrow \infty$, $c_f \rightarrow 0$, we have $\alpha = 1$ and $1/N = 1/M = 0$, and $K_u \rightarrow \infty$ and $B = 1$. In that case, the variation of volume content is equal to the volumetric strain (change of volume of the porous solid), such that the material response is incompressible in the undrained limit.
- Highly compressible fluid $c_f \gg 1/K$, $1/M = \phi_o c_f$, $K_u = K + \frac{\alpha^2}{\phi_o c_f}$, $B = \frac{\alpha}{\phi_o c_f K}$. Moreover in the limit $c_f \rightarrow \infty$ (highly compressible fluid), the material behaves as an elastic material without fluid ($K_u = K$, $B = 0$).

Micromechanical models Micromechanical models to express the porous solid drained bulk modulus K as function of the constituents elastic properties (k_s but also their shear modulus g_s) and the porosity are needed to go a step further and obtain expressions functions solely of the microscale morphology and constituents properties. We will not do so in the remaining of this course. It is important to bare in mind, that typically more than one solid constituents are present in the REV and that the bulk modulus of the porous solid is *largely* dependent on the contacts between minerals, rendering any hope of predictive capabilities of micromechanical models in practice. However, micromechanics can help “guide”/constraint the choice of macroscopic constitutive relations (e.g. when non-linearities come into play).

5.2.3 Relations between different poroelastic constants

$$\begin{aligned} K_u &= K + \alpha^2 M \\ B &= \frac{\alpha M}{K + \alpha^2 M} = \frac{K_u - K}{\alpha K_u} \end{aligned}$$

Note $K_u \in [K, \infty[$. Note that the Skempton B coefficient can potentially be above 1.

We also have the usual elastic relations between K , G and ν , G (for example). We can therefore define a drained and undrained Poisson's ratio

$$\begin{aligned} \nu &= \frac{3K - 2G}{2(3K + G)} \\ \nu_u &= \frac{3K_u - 2G}{2(3K_u + G)} \end{aligned}$$

[we see that $\nu_u \in [\nu, 0.5]$, in the limit $K_u \rightarrow \infty$]

$$K = \frac{E}{3(1 - 2\nu)} = \frac{2G(1 + \nu)}{3(1 - 2\nu)}$$

we obtain

$$\begin{aligned} B &= \frac{3(\nu_u - \nu)}{\alpha(1 - 2\nu)(1 + \nu_u)} \\ \eta &= \frac{\alpha(1 - 2\nu)}{2(1 - \nu)} \\ M &= \frac{2G(\nu_u - \nu)}{\alpha^2(1 - 2\nu_u)(1 - 2\nu)} \end{aligned}$$

Note $\eta \in [0, 0.5]$ and $M \in [0, \infty[$

5.2.4 Non-linearities

It is important to bare in mind that most rocks will exhibit a linearly reversible elastic response only over a short range of stress / pore-pressure around their natural configuration (initial stress and pore pressure). The response of the material will exhibit non-linear effects relatively soon - either reversible (non linear poroelasticity, where one can always define tangent poroelastic properties), or irreversible (poroplastic response). In rocks, the effect of crack closure is non-linear and always prominent at low stress. Although we focus mostly on linear reversible poroelasticity, we should always keep in mind that the linear poroelastic model is valid for small stress/pore pressure variation.

5.3 Fluid flow in porous media

We now go back to the description of fluid flow in a RVE of a porous media - following previous chapter. The difference will now come from a rigorous treatment of the variation of mass. First, the conservation of fluid mass in the RVE is given by:

$$\frac{\partial m_f}{\partial t} + (\rho_f q_i)_{,i} = \rho_f \gamma$$

where q_i is the filtration vector (also called specific discharge) and γ denotes a fluid source. It describes the fluid motion with respect to the solid, it's the phase-slip velocity weighted by the fluid volume in the RVE (i.e the porosity):

$$q_i = \phi \left(v_i^f - v_i^s \right)$$

The term $\frac{\partial m_f}{\partial t} = \frac{\partial \rho_f \phi}{\partial t}$ can be re-written following our discussion in section 5.1:

$$\frac{\partial m_f}{\partial t} = \rho_f^o \frac{\partial \zeta}{\partial t}$$

Under the assumption of small strain - neglecting second order terms in density change - we can thus rewrite the mass conservation as:

$$\frac{\partial \zeta}{\partial t} + q_{i,i} = \gamma \quad (5.4)$$

and we have the usual Darcy's law relating the fluid discharge vector and the fluid pore-pressure (see the previous chapter)

$$q_i = -\frac{k}{\mu_f} (p + \rho_f g x_3)_{,i} = -K \times h_{,i} \quad (5.5)$$

Recall as usual, in practice, the mobility coefficient $\kappa = k/\mu_f$ is sometimes used, as well as the hydraulic conductivity $K = k\rho_f g/\mu_f$ (in m/s).

Rock	ϕ	k (milliDarcy)	G (GPa)	K (GPa)	K_u (GPa)	α	B
Berea Sandstone*	0.19	$1.9 \cdot 10^2$	6	8	16	0.79	0.62
Boise Sandstone*	0.26	$8 \cdot 10^2$	4.2	4.6	8.3	0.85	0.5
Ohio Sandstone*	0.19	$5.6 \cdot 10^0$	6.8	8.4	13	0.74	0.5
Pecos Sandstone*	0.2	$8.0 \cdot 10^{-1}$	5.9	6.7	14	0.83	0.61
Ruhr Sandstone*	0.02	$2.0 \cdot 10^{-1}$	13	13	30	0.65	0.88
Weber Sandstone*	0.06	$1 \cdot 10^0$	12	13	25	0.64	0.73
Tennessee marble*	0.02	$1.0 \cdot 10^{-4}$	24	40	44	0.19	0.51
Charcoal granite*	0.02	$1.0 \cdot 10^{-4}$	19	35	41	0.27	0.55
Westerly granite*	0.01	$4.0 \cdot 10^{-4}$	15	25	42	0.47	0.85
Mudstone **				2.13	10.1	0.95	0.83
Indiana Limestone**	0.13		12.1	21.2	31.2	0.71	0.46
Stiff clay***	-	$1.0 \cdot 10^{-5}$	2	4	14	1	0.71

Table 5.1: Examples of poroelastic properties for different rock types. * Taken from Detournay & Cheng (1993). ** taken from Wang (2000). *** plausible representative values.

5.4 Quasi-static poroelasticity

The mechanical evolution of a porous media in domain Ω is governed by the balance of momentum and fluid-mass conservation. Restricting to quasi-static conditions (dropping the inertial term $\rho \frac{\partial^2 u_i}{\partial t^2}$), the balance of momentum reduces to:

$$\sigma_{ij,j} + f_i = 0$$

where $\mathbf{f} = -\rho_b g \mathbf{e}_3$ is the body-weight force, σ is the total stress acting on the porous material. ρ_b is the bulk density: i.e. $\rho_b = \rho_f \phi_o + (1 - \phi_o) \rho_s$ (where ρ_s is the solid phase density). The fluid conservation of mass can be re-written (accounting for a source/sink term γ)

$$\frac{\partial \zeta}{\partial t} + q_{i,i} = \gamma$$

[Note that we have linearized fluid density variation and dropped second order terms - small strain / small displacement poroelasticity is a first order theory].

The constitutive linear poroelastic law is

$$\begin{aligned} \sigma_{ij} &= 2G\epsilon_{ij} + (K - 2/3G) \epsilon_{kk} \delta_{ij} - \alpha p \delta_{ij} \\ \zeta &= \alpha \epsilon_{kk} + \frac{1}{M} p \end{aligned}$$

with the usual definition of the strain tensor

$$\epsilon_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i})$$

The Darcy's law complete the formulation

$$q_i = -\frac{k}{\mu} (p_{,i} + \rho_f g \delta_{i3}).$$

We see that due to the time dependence associated with variation of fluid content, the problem is time-dependent. It is an initial boundary value problem. In addition to the previous equations, we need appropriate initial conditions and boundary conditions to close the system of equations. [We will see the details of both boundary and initial conditions next week]

5.5 Solution of undrained problems

The mechanical evolution of a porous media in domain Ω is governed by the balance of momentum and fluid-mass conservation. This week we will focus on so-called undrained problems - which corresponds to the instantaneous response of a porous media to sudden changes in loading where the fluid has no time to start “flowing” such that $q_i = 0$. The fluid mass conservation thus reads in that case in the absence of any body source terms (because the fluid has no time to start flowing such that $q_i = 0$)

$$\frac{\partial \zeta}{\partial t} = 0$$

the variation of fluid content is thus null: $\zeta = 0$ (because at time $t = 0$, $\zeta = 0$ i.e. $m_f = m_f^o$), which we can rewrite using the poroelastic constitutive relation

$$\alpha \epsilon^v + \frac{1}{M} p = 0 \quad (5.6)$$

On the other hand, restricting to quasi-static conditions, the balance of momentum reduces to:

$$\sigma_{ij,j} + f_i = 0 \quad (5.7)$$

where $\mathbf{f} = -\rho_b g \mathbf{e}_3$ is the body-weight force, σ_{ij} is the total stress acting on the porous material. ρ_b is the bulk density: i.e. $\rho_b = \rho_f \phi_o + (1 - \phi_o) \rho_s$ (where ρ_s is the solid phase density). The undrained problem has by nature no time-dependence - it corresponds to the instantaneous response of the porous media. Two ways can be approached for its solution: i) replacing (5.6) in the poroelastic constitutive law and solve for an equivalent undrained elastic problem, ii) solving (5.6) and (5.7) simultaneously. We start by discussing i) this week which reduce to an elasto-static problem. We will discuss ii) next week.

5.5.1 The undrained elasticity problem

As $\zeta = 0$ during the undrained response, we have $p = -\alpha M \epsilon^v$ and thus we can re-write the first poroelastic constitutive law as

$$\sigma_{ij} = 2G\epsilon_{ij} + (K_u - 2/3G)\epsilon^v \delta_{ij}$$

with $K_u = K + \alpha^2 M$. For simplicity for the weak form, we will use the fourth order stiffness tensor c_{ijkl} .

The problem therefore reduces to an elastic problem. Generalizing to anisotropy and using an undrained stiffness tensor as c_{ijkl} , the strong form reads in domain Ω

$$\begin{aligned} \sigma_{ij,j} + f_i &= 0 & \text{in } \Omega \\ \sigma_{ij} &= c_{ijkl}\epsilon_{kl} & \text{in } \Omega \\ t_i = \sigma_{ij}n_j &= t_i^g & \text{in } \Gamma_{t_i} \\ u_i &= u_i^g & \text{in } \Gamma_{u_i} \end{aligned}$$

The weak form of the problem can be obtained in the same way than for the case of confined flow problems. Introducing a test vector v_i with the same continuity requirement than the displacement u_i such that $v_i = 0$ on Γ_{u_i} , we have

$$\int_{\Omega} v_i \sigma_{ij,j} dV + \int_{\Omega} v_i f_i dV = 0$$

using the chain rule and Green's identity, we obtain

$$\int_{\Gamma_{t_i}} v_i t_i^g dS + \int_{\Omega} v_i f_i dV = \int_{\Omega} v_{i,j} c_{ijkl} u_{k,l} dV$$

(where we have also used the known symmetry of the stiffness tensor).

5.5.1.1 Finite element solution

The difference with the steady-state confined flow problem (Laplacian) is that the unknown is the displacement vector - i.e. the problem is vectorial and not simply scalar. That difference aside, the procedure are completely similar. The difference is that we now need a mapping between the equation number and the local degree of freedom.

We will use a similar unstructured mesh than in previous chapter. Now at every nodes of the mesh, in 2D, we have two unknowns u_x and u_y . We therefore introduce a so-called ID array which map the local degree of freedoms to the global equation number. For a mesh with n_n nodes, it is a matrix of size $(n_n, 2)$ containing integers. One row corresponds to one node, the first column to the global number of the u_x DOF for that node, and the second column the global number of the u_y DOF for that node. It is typically formatted as:

nodes # \ DOF	$u_1 = u_x$	$u_2 = u_y$
1	1	2
2	3	4
3	5	6
....		

We can also use a direct relation between global DOF and nodes number:

$$\text{DOF } u_1 \text{ node } i = 2 * (i - 1) + 1$$

$$\text{DOF } u_2 \text{ node } i = 2 * i$$

Isoparametric element We will continue to use isoparametric elements, but possibly quadratic. We write the displacement vector in element e : for $x \in \Omega_e$

$$\mathbf{u}(\mathbf{x}) = \sum_{a=1}^{n_n^e} N_a(\xi(\mathbf{x})) \mathbf{u}_a^e$$

where n_n^e is the number of nodes of element e - 3 for a linear triangle, 6 for a quadratic triangle. Similarly, the vectorial test function is interpolated as for $x \in \Omega_e$

$$\mathbf{v}(\mathbf{x}) = \sum_{a=1}^{n_n^e} N_a(\xi(\mathbf{x})) \mathbf{v}_a^e$$

The gradient of the displacement (and the test function) can thus be obtained for $x \in \Omega_e$ - restricting to the 2D case ($x_1 = x$, $x_2 = y$) where $\mathbf{u} = \begin{pmatrix} u_1 = u_x \\ u_2 = u_y \end{pmatrix}$ and to a linear triangle (constant strain)

$$\begin{pmatrix} u_{1,1} \\ u_{2,2} \\ u_{1,2} + u_{2,1} \end{pmatrix} = \begin{pmatrix} \partial_x u_x \\ \partial_y u_y \\ \partial_y u_x + \partial_x u_y \end{pmatrix} = \underbrace{\begin{bmatrix} \partial_x N_1 & 0 & \partial_x N_2 & 0 & \partial_x N_3 & 0 \\ 0 & \partial_y N_1 & 0 & \partial_y N_2 & 0 & \partial_y N_3 \\ \partial_y N_1 & \partial_x N_1 & \partial_y N_2 & \partial_x N_2 & \partial_y N_3 & \partial_x N_3 \end{bmatrix}}_{\mathbf{B}^e} \cdot \begin{pmatrix} u_x^1 \\ u_y^1 \\ u_x^2 \\ u_y^2 \\ u_x^3 \\ u_y^3 \end{pmatrix}$$

$$\epsilon^e(\mathbf{x}) = \mathbf{B}^e \cdot \mathbf{u}^e$$

where u_x^1 is the x -component of the displacement at node 1 of the corresponding reference element, and so on, while N_1 denotes the shape function associated with node 1 and so on. $\mathbf{B}^e$ is called the strain matrix. It allows to obtain the strain from the value of the displacement at the nodes of the element. We have introduce

$$\epsilon(\mathbf{x}) = \begin{pmatrix} \epsilon_{xx} \\ \epsilon_{yy} \\ 2\epsilon_{xy} \end{pmatrix}$$

as a vector containing the component of the strain tensor (in 2D). It is thus customary to introduce a stress vector containing the corresponding stress component (here in 2D)

$$\sigma = \begin{pmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{xy} \end{pmatrix}$$

such that the elastic stress-strain relation can be rewritten using a matrix vector format, with an undrained stiffness matrix $\mathbf{D}_u$

$$\sigma = \mathbf{D} \cdot \epsilon$$

with for the case of plane-strain and an isotropic material

$$\mathbf{D}_u = \begin{bmatrix} K_u + 4/3G & K_u - 2/3G & 0 \\ K_u - 2/3G & K_u + 4/3G & 0 \\ 0 & 0 & G \end{bmatrix} = \frac{E_u}{(1 + \nu_u)(1 - 2\nu_u)} \begin{bmatrix} (1 - \nu_u) & \nu_u & 0 \\ \nu_u & (1 - \nu_u) & 0 \\ 0 & 0 & \frac{1 - 2\nu_u}{2} \end{bmatrix}$$

[Note: in passing, we see that for an incompressible material $\nu = 0.5$, the stiffness matrix will become singular - and for that case, a different finite formulation must be devised altogether. In practice even in the undrained limit, for all geo-materials we have $\nu < 0.5$ - remembering that in our formulation water is slightly compressible].

Element stiffness matrix Focusing first on the integral

$$\int_{\Omega} v_{i,j} c_{ijkl} u_{k,l} dV$$

appearing on the weak form. It will be as usual split as a sum of integrals over all the elements, using the matrix vector notation just introduced, we obtain

$$\int_{\Omega^e} \mathbf{v}^{eT} \cdot \mathbf{B}^{eT} \cdot \mathbf{D} \cdot \mathbf{B}^e \cdot \mathbf{u}^e dV = \mathbf{v}^{eT} \cdot \underbrace{\left(\int_{\Omega^e} \mathbf{B}^{eT} \cdot \mathbf{D} \cdot \mathbf{B}^e dV \right)}_{\mathbf{k}^e} \cdot \mathbf{u}^e = \mathbf{v}^{eT} \cdot \mathbf{k}^e \cdot \mathbf{u}^e$$

where $\mathbf{k}^e$ is denoted as the stiffness matrix of element e . It is a 6×6 matrix for a linear 3 nodes triangle. For a quadratic 6 nodes triangle, it will be a 12×12 matrix.

Global stiffness matrix The assembly procedure to construct the global stiffness matrix follows the same procedure as for scalar problem but now as unknowns are vectorial, we must use the mapping between the global DOF number and the local DOF number, using the ID array discussed above. We schematically write this assembly procedure as

$$\mathbf{K} = \sum \mathbf{k}^e$$

Body force term

$$\int_{\Omega} v_i f_i \, dV$$

again will be obtained by summation of integrals over all the elements. In matrix form,

$$\int_{\Omega^e} \mathbf{v}^e \cdot \mathbf{N}^{eT} \cdot \mathbf{f} \, dV = \mathbf{v}^e \cdot \int_{\Omega^e} \mathbf{N}^{eT} \cdot \mathbf{f} \, dV$$

with for a linear triangle - for such a vectorial equation -

$$\mathbf{N}^e(\mathbf{x}) = \begin{bmatrix} N_1(\mathbf{x}) & 0 & N_2(\mathbf{x}) & 0 & N_3(\mathbf{x}) & 0 \\ 0 & N_1(\mathbf{x}) & 0 & N_2(\mathbf{x}) & 0 & N_3(\mathbf{x}) \end{bmatrix}$$

Applied tractions The last term appearing in the weak form is related to applied tractions and is thus a surface (/line in 2D) integral

$$\int_{\Gamma_{t_i}} v_i t_i^g \, dS$$

here again, we split it in as integral over the edge of the elements making the part of the boundary Γ_{t_i} . In 2D it reduces to a simple line integral, and for example for traction applied along direction 1

$$\int_{\Gamma_{t_1}^e} \mathbf{v}_1^{eT} \mathbf{N}^{eT} t_1^g \, dS$$

with here e.g. for a linear segment

$$\mathbf{N}^e = \begin{bmatrix} N_1 & N_2 \end{bmatrix}$$

with 1 and 2 the 2 nodes of the segment (for a linear interpolation) and t_1^g is the applied traction on the segment (possibly varying in space). In the case of a quadratic triangle, there are 3 nodes along one edge.

Chapter 6

Quasi-static Poroelasticity II

We now continue and solve the complete initial boundary value problem in poroelasticity using finite element in space and finite difference for time integration. We solve the problem using the displacement vector u_i and the fluid pore-pressure as the main unknowns.

6.1 The initial boundary value problem

Let's start by recalling the initial boundary value problem to solve.

The mechanical evolution of a porous media in domain Ω is governed by the balance of momentum and fluid-mass conservation. Restricting to quasi-static conditions, the balance of momentum reduces to:

$$\sigma_{ij,j} + f_i = 0$$

where $\mathbf{f} = -\rho_b g \mathbf{e}_3$ is the body-weight force, σ_{ij} is the total stress acting on the porous material. ρ_b is the bulk density: i.e. $\rho_b = \rho_f \phi_o + (1 - \phi_o) \rho_s$ (where ρ_s is the solid phase density). The fluid conservation of mass can be re-written (accounting for a source/sink term γ)

$$\frac{\partial \zeta}{\partial t} + q_{i,i} = \gamma$$

[Note that we have linearized fluid density variation and dropped second order terms - small strain / small displacement poroelasticity is a first order theory].

The constitutive linear poroelastic law is - generalizing to anisotropy and accounting for the possible existence of initial stresses and pore pressure:

$$\begin{aligned} \sigma_{ij} - \sigma_{ij}^o &= c_{ijkl} \epsilon_{kl} - \alpha_{ij} (p - p^o) \\ \zeta &= \alpha_{ij} \epsilon_{ij} + \frac{1}{M} (p - p^o) \end{aligned}$$

where the usual definition of the strain tensor

$$\epsilon_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i}).$$

The Darcy's law complete the formulation

$$q_i = -\kappa_{ij}(p_{,i} + \rho_f g \delta_{i3}).$$

The Biot coefficient tensor is notably symmetric α_{ij} , like the hydraulic mobility coefficient $\kappa_{ij} = k_{ij}/\mu_f$ (where μ_f is the pore-fluid viscosity).

We see that due to the time dependence associated with variation of fluid content, the problem is time-dependent. It is an initial boundary value problem. In addition to the previous equations, we need appropriate initial conditions and boundary conditions to close the system of equations.

6.1.1 Initial and boundary conditions

Initial conditions (e.g. stress/pore pressure free reference configuration) and appropriate (non-intersecting) boundary conditions (displacement-traction, pore-pressure - flow-rate) allows to obtain a well-posed problem. Let's denote Γ the boundary of Ω . The boundary conditions reads:

- Imposed displacement on part of the boundary Γ_{u_i} (Dirichlet type)

$$u_i = u_i^g \quad \text{on } \Gamma_{u_i}$$

- Imposed pore-pressure on part of the boundary Γ_p (Dirichlet type)

$$p = p^g \quad \text{on } \Gamma_p$$

- Imposed tractions on part of the boundary Γ_{t_i} (Neumann type)

$$t_i = \sigma_{ij}n_j = t_i^g \quad \text{on } \Gamma_{t_i}$$

- Imposed fluid flux on part of the boundary Γ_q (Neumann type)

$$q_n = q_i n_i = q^g \quad \text{on } \Gamma_q$$

In order for the problem to be well-posed, the parts of the boundary where Neumann and Dirichlet BC for mechanics and fluid flow must not intersect:

$$\Gamma_{u_i} \cup \Gamma_{T_i} = \Gamma \quad \Gamma_{u_i} \cap \Gamma_{T_i} = \emptyset \quad \text{for } i = 1, 3$$

and

$$\Gamma_p \cup \Gamma_q = \Gamma \quad \Gamma_p \cap \Gamma_q = \emptyset$$

Note also that an exclusively Neumann problem is ill-posed (up to a rigid body motion and a pore pressure). It is also worthwhile to note that the boundary conditions are possibly time-dependent (and parts of the boundary where e.g. tractions are applied can possibly change with time).

Initial conditions Remember that the initial conditions can consist of an initial stress σ_{ij}^o and pore-pressure p^o fields (This will be the case in most geotechnical problem):

$$\begin{aligned} \sigma_{ij}(t=0, \mathbf{x}) &= \sigma_{ij}^o(\mathbf{x}) & \text{in } \Omega \\ p(t=0, \mathbf{x}) &= p^o(\mathbf{x}) & \text{in } \Omega \\ u_i(t=0, \mathbf{x}) &= 0 & \text{in } \Omega \\ \zeta(t=0, \mathbf{x}) &= 0 & \text{in } \Omega \end{aligned}$$

The initial stress and pore-pressure field must of course satisfy balance of momentum and continuity equations under steady-state condition:

$$\begin{aligned} \sigma_{ij,j}^o + f_i &= 0 \\ -(\kappa_{ij}(p^o + \rho_f g x_3 \delta_{i3})_{,j})_{,i} &= 0 \end{aligned}$$

with the boundary conditions

$$t_i^o = \sigma_{ij}^o n_j \quad \text{on } \Gamma_{t_i}$$

Solving for perturbations from an initial state We thus see that we can rewrite the field equation of poroelasticity (balance of momentum of the porous media and fluid continuity equation) as

$$(\sigma_{ij} - \sigma_{ij}^o)_{,j} = 0 \quad \text{in } \Omega \quad (6.1)$$

$$\frac{\partial \zeta}{\partial t} - \left(\kappa_{ij} (p - p^o)_{,j} \right)_{,i} = \gamma \quad \text{in } \Omega \quad (6.2)$$

We can thus solve for the perturbed state. In that way, we can thus put the effect of the solid and fluid body forces into an initial stress and pore-pressure field.

6.2 Weak form of the problem

The first toward the solution of the quasi-static poroelastic problem with the finite element method is to obtain a weak form. We will use the continuous Galerkin method and focus

on a formulation where the primary unknowns will be the displacement vector u_i and pore-pressure p (one vectorial unknown, one scalar). As a result we will introduce two test/virtual functions: one vectorial - virtual displacement v_i , one scalar - virtual pore-pressure r . We assume the same continuity requirements (C_o) for these tests function as for the unknowns. Moreover, we enforce these test functions to have zero value on the part of the boundary where Dirichlet boundary conditions are imposed.

We follow exactly the same route than for the elastic case. Here, we use equation (6.1) for the balance of momentum and obtain

$$\int_{\Omega} \epsilon_{ij}(\mathbf{v}) \sigma_{ij}(\mathbf{u}) \, dV = \int_{\Gamma_{t_i}} t_i^g v_i \, dS$$

Similarly, using the same route than for transient flow (using here (6.2) for the fluid continuity, we obtain

$$\int_{\Omega} r \frac{\partial \zeta}{\partial t} \, dV + \int_{\Omega} r_{,i} \kappa_{ij} (p + \rho_f g x_3 \delta_{j3})_{,j} = - \int_{\Gamma_q} q_i n_i r \, dS + \int_{\Omega} r \gamma \, dV$$

6.2.1 Introducing the poroelastic constitutive equation

Introducing the poroelastic constitutive relation, the weak form of the balance of momentum becomes

$$\int_{\Omega} \epsilon_{ij}(\mathbf{v}) c_{ijkl} \epsilon_{kl}(\mathbf{u}) \, dV - \int_{\Omega} \epsilon_{ij}(\mathbf{v}) \alpha_{ij} p \, dV = - \int_{\Omega} \epsilon_{ij}(\mathbf{v}) \sigma_{ij}^o \, dV - \int_{\Omega} \epsilon_{ij}(\mathbf{v}) \alpha_{ij} p^o \, dV + \int_{\Gamma_{t_i}} t_i^g v_i \, dS$$

and for the fluid continuity

$$\int_{\Omega} r \alpha_{ij} \frac{\partial \epsilon_{ij}(\mathbf{u})}{\partial t} \, dV + \int_{\Omega} r \frac{1}{M} \frac{\partial p}{\partial t} \, dV + \int_{\Omega} r_{,i} \kappa_{ij} (p + \rho_f g x_3 \delta_{j3})_{,i} \, dV = - \int_{\Gamma_q} q_i n_i r \, dS + \int_{\Omega} r \gamma \, dV$$

6.2.2 Final weak form

We can re-write the weak forms using the following bilinear operators

$$\begin{aligned} \mathcal{E}(v_i, u_k) &= \int_{\Omega} \epsilon_{ij}(\mathbf{v}) c_{ijkl} \epsilon_{kl}(\mathbf{u}) \, dV \\ \mathcal{A}(v_i, p) &= \int_{\Omega} \epsilon_{ij}(\mathbf{v}) \alpha_{ij} p \, dV \\ \mathcal{S}(r, p) &= \int_{\Omega} r \frac{1}{M} \frac{\partial p}{\partial t} \, dV \\ \mathcal{C}(r, p) &= \int_{\Omega} r_{,i} \kappa_{ij} p_{,j} \, dV \end{aligned}$$

such that the weak form of the static linear balance of momentum is

$$\mathcal{E}(v_i, u_l) - \mathcal{A}(v_i, p) = \int_{\Gamma_{t_i}} t_i^g v_i \, dS - \int_{\Omega} \epsilon_{ij}(\mathbf{v}) \sigma_{ij}^o dV + \mathcal{A}(v_i, p^o)$$

and the weak form of the fluid continuity equation is

$$\mathcal{A}\left(\frac{\partial u_i}{\partial t}, r\right) + \mathcal{S}\left(r, \frac{\partial p}{\partial t}\right) + \mathcal{C}(r, p) = - \int_{\Gamma_q} q_i n_i r \, dS + \int_{\Omega} r \gamma dV - \int_{\Omega} r_{,i} \kappa_{i3} \rho_f g dV$$

where we see that the operator $\mathcal{A}(v_i, p)$ captures the coupling between displacement and pressure.

Note that in order to obtain a final system that is “symmetric”, we can take the opposite (multiply by -1) the weak form of the fluid continuity equation.

6.2.3 Discretized final system of ODEs

We see that each bi-linear weak form will corresponds to a block of a finite element matrix after discretization by finite element:

$$\begin{aligned} \mathcal{E}(v_i, v_l) &\approx \mathbf{v}^T \mathbf{K} \mathbf{u} \\ \mathcal{A}(v_i, p) &\approx \mathbf{v}^T \mathbf{A} \mathbf{p} \\ \mathcal{A}\left(\frac{\partial u_i}{\partial t}, r\right) &\approx \dot{\mathbf{u}}^T \mathbf{A} \mathbf{r} = \mathbf{r}^T \mathbf{A}^T \dot{\mathbf{u}} \\ \mathcal{S}\left(r, \frac{\partial p}{\partial t}\right) &\approx \mathbf{r}^T \mathbf{S} \dot{\mathbf{p}} \\ \mathcal{C}(r, p) &\approx \mathbf{r}^T \mathbf{C} \mathbf{p} \end{aligned}$$

where we have switched to the matrix notation with unknowns located at the nodes of the finite element mesh. The matrices are denoted respectively as the stiffness ($\mathbf{K}$), hydro-mechanical coupling ($\mathbf{A}$), storage ($\mathbf{S}$), and conductivity ($\mathbf{C}$).

The weak form can thus be re-written as:

$$\begin{aligned} \mathbf{v}^T \mathbf{K} \mathbf{u} - \mathbf{v}^T \mathbf{A} \mathbf{p} &= \mathbf{v}^T \mathbf{f}_m \\ -\mathbf{r}^T \mathbf{A}^T \dot{\mathbf{u}} - \mathbf{r}^T \mathbf{S} \dot{\mathbf{p}} - \mathbf{r}^T \mathbf{C} \mathbf{p} &= \mathbf{r}^T \mathbf{f}_f \end{aligned}$$

where $\mathbf{f}_m$ and $\mathbf{f}_f$ are the forcing terms associated with imposed tractions and initial stress and pore-pressure (mechanics) and imposed flux + fluid sources (flow - consistent with the proper sign convention chosen). We can re-write the system as:

$$\begin{pmatrix} \mathbf{v} \\ \mathbf{r} \end{pmatrix}^T \begin{bmatrix} \mathbf{K} & -\mathbf{A} \\ \mathbf{0} & -\mathbf{C} \end{bmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{p} \end{pmatrix} + \begin{pmatrix} \mathbf{v} \\ \mathbf{r} \end{pmatrix}^T \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ -\mathbf{A}^T & -\mathbf{S} \end{bmatrix} \begin{pmatrix} \dot{\mathbf{u}} \\ \dot{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} \mathbf{v} \\ \mathbf{r} \end{pmatrix}^T \begin{pmatrix} \mathbf{f}_t \\ \mathbf{f}_q \end{pmatrix}$$

which is valid for any couple of test functions $(\mathbf{v}, \mathbf{r})$ such that we obtain the following system of ODEs (Ordinary Differential Equation)

$$\begin{bmatrix} \mathbf{K} & -\mathbf{A} \\ \mathbf{0} & -\mathbf{C} \end{bmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{p} \end{pmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ -\mathbf{A}^T & -\mathbf{S} \end{bmatrix} \begin{pmatrix} \dot{\mathbf{u}} \\ \dot{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} \mathbf{f}_t \\ \mathbf{f}_q \end{pmatrix}$$

Now we could have from the onset taken the time-derivative of the momentum equation. It is the same (Convince your self by deriving it) to derivate the first “row” block of the previous system. We thus obtain the following system of ODEs:

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{C} \end{bmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{p} \end{pmatrix} + \begin{bmatrix} \mathbf{K} & -\mathbf{A} \\ -\mathbf{A}^T & -\mathbf{S} \end{bmatrix} \begin{pmatrix} \dot{\mathbf{u}} \\ \dot{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} \dot{\mathbf{f}}_t \\ \mathbf{f}_q \end{pmatrix}$$

6.2.4 Solving for variation from the initial state

It is also possible to solve for variation from the initial state (this is possible only for the linear case). We will write abusively

$$\begin{aligned} \tilde{\sigma}_{ij} &= \sigma_{ij} - \sigma_{ij}^o \\ \tilde{p} &= p - p^o \end{aligned}$$

note that as a result, the traction boundary conditions becomes on Γ_{t_i}

$$\tilde{\sigma}_{ij} n_j = t_i^g - \sigma_{ij}^o n_j$$

similarly for the dirichlet boundary conditions on pressure if any become on Γ_p

$$\tilde{p} = p^g - p^o$$

The initial stress and pore pressure do not enter the system of equation to solve anymore

Introducing the poroelastic constitutive equation Introducing the poroelastic constitutive relation, the weak form of the balance of momentum becomes

$$\int_{\Omega} \epsilon_{ij}(\mathbf{v}) c_{ijkl} \epsilon_{kl}(\mathbf{u}) \, dV - \int_{\Omega} \epsilon_{ij}(\mathbf{v}) \alpha_{ij} \tilde{p} \, dV = \int_{\Gamma_{t_i}} (t_i^g - \sigma_{ij}^o n_j) v_i \, dS$$

and for the fluid continuity

$$\int_{\Omega} r \alpha_{ij} \frac{\partial \epsilon_{ij}(\mathbf{u})}{\partial t} \, dV + \int_{\Omega} r \frac{1}{M} \frac{\partial \tilde{p}}{\partial t} \, dV + \int_{\Omega} r_{,i} \kappa_{ij} \tilde{p}_{,j} = - \int_{\Gamma_q} q_i n_i r \, dS + \int_{\Omega} r \gamma \, dV$$

This is of course strictly equivalent than solving for p and σ_{ij} ... [do the back-substitution to convince yourself] the rest follow from there.

6.3 Spatial discretization

So far, we have not specified the order of interpolation for the primary unknowns: pore-pressure and displacement. One could think for example to use linear finite element for both. However, we can directly see from the poroelastic stress-strain relation that this may not be the best choice. Indeed, using a linear interpolation for displacement imply that for example for a triangular finite element, the strain are uniform in the element - and so is $c_{ijkl}\epsilon_{kl}$. In poro-elasticity to estimate the total stress, one then needs to add the effect of pore-pressure. If the pore-pressure is also interpolated linearly, we end up summing a constant term with a linear one inside the element to estimate the total stress. We can “feel” that this is not optimal may lead to numerical problem. In fact, it can be mathematically proven, that such choice of interpolation (linear displacement - linear pore-pressure) is unstable. It may work for some problems, but typically can exhibit oscillations.

To ensure that we add term of similar order in the constitutive poroelastic relation, we can choose to use a quadratic interpolation for displacement (and thus a linear variation of strain) and a linear interpolation for pore-pressure.

6.4 Time Integration

We see that an explicit (forward Euler) scheme is not very appropriate here as i) the time step must be smaller than the critical time-step of the flow / diffusion equation and that ii) we would have to invert a rather large matrix at each (small) time step. We will thus stick to an implicit scheme - which is both unconditionnally stable and first order accurate in time¹.

We thus write

$$\Delta t \begin{pmatrix} \dot{\mathbf{u}} \\ \dot{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} \Delta \mathbf{u} \\ \Delta \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{u}^{n+1} \\ \mathbf{p}^{n+1} \end{pmatrix} - \begin{pmatrix} \mathbf{u}^n \\ \mathbf{p}^n \end{pmatrix}$$

and express the ODE system at time t_{n+1} (Implicit scheme / backward Euler) such that we can re-write it as

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\Delta t \mathbf{C} \end{bmatrix} \begin{pmatrix} \mathbf{u}^{n+1} \\ \mathbf{p}^{n+1} \end{pmatrix} + \begin{bmatrix} \mathbf{K} & -\mathbf{A} \\ -\mathbf{A}^T & -\mathbf{S} \end{bmatrix} \begin{pmatrix} \Delta \mathbf{u} \\ \Delta \mathbf{p} \end{pmatrix} = \Delta t \begin{pmatrix} \dot{\mathbf{f}}_t \\ \dot{\mathbf{f}}_q \end{pmatrix}$$

or re-arranging in terms of increments:

$$\begin{bmatrix} \mathbf{K} & -\mathbf{A} \\ -\mathbf{A}^T & -\mathbf{S} - \Delta t \mathbf{C} \end{bmatrix} \begin{pmatrix} \Delta \mathbf{u} \\ \Delta \mathbf{p} \end{pmatrix} = \Delta t \begin{pmatrix} \dot{\mathbf{f}}_t \\ \dot{\mathbf{f}}_q \end{pmatrix} - \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\Delta t \mathbf{C} \end{bmatrix} \begin{pmatrix} \mathbf{u}^n \\ \mathbf{p}^n \end{pmatrix}$$

¹One can also devise a θ -method - and use a Crank-Nicholson scheme $\theta = 1/2$ which is second order accurate in time.

6.4.1 Undrained problem

Upon sudden application of a loading at time $t = 0^+$, as $\Delta t \times \dot{\mathbf{f}}_t \approx \mathbf{f}_t(t = 0^+)$ is finite (i.e. it is the load suddenly applied), we thus obtain the following “undrained” system:

$$\begin{bmatrix} \mathbf{K} & -\mathbf{A} \\ -\mathbf{A}^T & -\mathbf{S} \end{bmatrix} \begin{pmatrix} \Delta \mathbf{u} \\ \Delta \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{f}_t \\ 0 \end{pmatrix}$$

which we can simplify. Indeed, the second row gives:

$$\Delta \mathbf{p} = -\mathbf{S}^{-1} \mathbf{A}^T \Delta \mathbf{u}$$

and back-substituting in the first row, we obtain:

$$(\mathbf{K} + \mathbf{A} \mathbf{S}^{-1} \mathbf{A}^T) \Delta \mathbf{u} = \mathbf{f}_t$$

Note that $\mathbf{A}$ scales as α and $\mathbf{S}$ as $1/M$, and $\mathbf{K}$ (as K and G), it is akin (but not strictly equivalent because of numerical round-off and inversion of $\mathbf{S}$) to taking the undrained bulk modulus of the poroelastic material and solve for an equivalent problem, i.e.

$$\mathbf{K}_u \Delta \mathbf{u} = \mathbf{f}_t$$

Chapter 7

Extension to Thermoporoelasticity

these set of notes will not be reviewed during class - and would need some tidy up.

We will assume local thermal equilibrium within the Representative Element of Volume: i.e. the fluid and solid phase are at the same temperature T inside the REV. In what follows, we restrict to small strain linear behavior & isotropy for clarity.

7.1 Constitutive equations

7.1.1 The solid porous skeleton

Assuming reversibility, the solid Helmholtz free energy of the solid $\Psi_s = u_s - TS_s$ is a total differential

$$d\Psi_s = \sigma_{ij}d\epsilon_{ij} + pd\varphi - S_s dT$$

where S_s is the porous solid entropy per unit of volume (not per unit of mass) and u_s the porous solid internal energy (also per unit volume).

Assuming isotropy, we will write the constitutive equations from an initial state $(\sigma_{ij}^o, p^o, T^o)$, assuming small strain and a linear behavior. Using Maxwell's relation (i.e. the solid free energy is a total differential), the definition of the Gibbs potential ($G_s = \Psi_s - p\varphi$), we obtain the following constitutive relations (for the porous solid):

$$\begin{aligned}\sigma_{ij} - \sigma_{ij}^o &= 2G\epsilon_{ij} + (K - 2/3G)\epsilon_{kk}\delta_{ij} - \alpha(p - p^o)\delta_{ij} - \beta K(T - T^o)\delta_{ij} \\ \varphi - \varphi^o &= \alpha\epsilon_{kk} + \frac{p - p^o}{N} - \beta_\phi(T - T^o) \\ S_s - S_s^o &= \beta K\epsilon_{kk} - \beta_\phi(p - p^o) + C^v \frac{(T - T^o)}{T}\end{aligned}$$

β is the volumetric thermal expansion coefficient of the solid skeleton (in K^{-1} K for Kelvin), β_ϕ is ther volumetric thermal expansion associated with the porosity. C^v is the skeleton

tangent *volumetric* heat capacity ($\text{J.K}^{-1}/\text{L}^3$) (the specific heat capacity is per unit mass, i.e. ($\text{J.K}^{-1}/\text{M}$) - the two are related via density).

Similarly than for the poroelastic case, the response of the porous material can be splitted in a volumetric and deviatoric part: the effect of pore-pressure and temperature changes will only affect the volumetric deformation in the isotropic case

$$\begin{aligned}\sigma - \sigma^o &= K\epsilon_{kk} - \alpha(p - p^o) - \beta K(T - T_o) \\ s_{ij} &= 2Ge_{ij}\end{aligned}$$

7.1.2 The saturating fluid

The previous porous skeleton constitutive relation are independent of the saturating fluid (only contact force, i.e. pore pressure exerted on the solid matrix are considered). Similarly, the constitutive equation of the fluid are independent of the solid. In differential form, we can write the fluid constitutive relation in a linearized form (e.g. for a liquid) as:

$$\frac{d\rho_f}{\rho_f} = \frac{dp}{K_f} - \beta_f dT \quad ds_f = -\beta_f \frac{dp}{\rho_f} + C_p \frac{dT}{T}$$

where K_f is the fluid bulk moduli, β_f its volumetric expansion coefficient and C_p the fluid *specific* heat capacity at constant pressure (note all this coefficient are 'tangent' ones). Note that here s_f has the dimension of J/K/M .

Note also that for an ideal gas, or if we want to account for phase changes, we need to use more advanced equation of state (EoS) for the saturating fluid.

Remember also that here C_p is the specific heat capacity (not the volumetric heat capacity).

7.1.3 The porous material

Like in the poroelastic case, we can define a variation of volume content as

$$\zeta = \frac{m_f - m_f^o}{\rho_f^o} \quad (7.1)$$

where the fluid mass per unit volume m_f is simply

$$m_f = \rho_f \phi$$

Similarly than in poroelasticity, neglecting second order term, we can write

$$\zeta = \varphi + \phi_o \frac{\Delta\rho_f}{\rho_f^o}$$

(i.e. $\zeta = dm_f/\rho_f^o$) and thus combining the constitutive description of the solid skeleton and fluid, we obtain

$$\zeta = \alpha\epsilon_{kk} + \left(\frac{1}{N} + \frac{\phi_o}{K_f}\right)(p - p^o) - (\beta_\phi + \phi_o\beta_f)(T - T^o)$$

or using the definition of the Biot Modulus M and $\beta_m = \beta_\phi + \phi_o\beta_f$ is the hydro-thermal expansion coefficient

$$\zeta = \alpha\epsilon_{kk} + \frac{p - p^o}{M} - \beta_m(T - T^o)$$

Moreover, neglecting second order terms,

$$d(m_f s_f) = s_f dm_f - \phi_o \beta_f dp + m_f \frac{C_p}{T} dT$$

Similarly the total entropy per unit of volume for the porous material in the REV is $S = m_f s_f + S_s$. Such that one has (neglecting second order terms),

$$dS = S - S^o = s_f \rho_f^o \zeta + \beta K \epsilon_{kk} - \beta_m(p - p^o) + C_d^v \frac{T - T^o}{T}$$

where $C_d^v = C^v + m_f C_p = C^v + \rho_f^o \phi_o C_p$ (at first order). Note that C_D is a volumetric heat capacity (while C_p is the fluid *specific* heat). Note that it is more usual to use the specific heat, so we can define for the porous material, the following specific heat

$$C = \frac{C^v + \rho_f^o \phi_o C_p}{\rho}$$

where $\rho = \rho_f^o \phi_o + (1 - \phi_o)\rho_s$.

7.1.4 Recapitulation of the linear thermoporoelastic constitutive laws

$$\sigma - \sigma^o = K \epsilon_{kk} - \alpha(p - p^o) - \beta K(T - T^o) \quad s_{ij} = 2G \epsilon_{ij}$$

$$\zeta = \alpha\epsilon_{kk} + \frac{p - p^o}{M} - \beta_m(T - T^o)$$

$$S - S^o = s_f \rho_f^o \zeta + \beta K \epsilon_{kk} - \beta_m(p - p^o) + C_d \frac{T - T^o}{T}$$

We can also re-write it as (focusing on the volumetric part) only

$$\sigma - \sigma^o = K_u \epsilon_{kk} - \alpha M \zeta - \beta_u K_u (T - T^o)$$

$$\zeta = \frac{\alpha}{K} \left((\sigma - \sigma^o) + \frac{p - p^o}{B} \right) - (\beta_m - \alpha\beta)(T - T^o)$$

where $K_u = K + \alpha^2 M$ is the undrained bulk moduli and $\beta_u = \beta \frac{K}{K_u} + \frac{\alpha M}{K_u} \beta_m$, $B = \alpha M / K_u$ is the Skempton coefficient etc.

7.2 Conservation laws

Balance of momentum

$$\sigma_{ij,j} + f_i = \rho \frac{\partial^2 u_i}{\partial t^2}$$

Mass conservation (under the assumption of small density changes) can be written as

$$\frac{\partial \zeta}{\partial t} + q_{i,i} = \gamma$$

where γ is a fluid source / sink term (e.g. injection from a well).

Energy conservation We have now to consider the first and the second law of thermodynamics (see e.g. your continuum and fluid mechanics classes)! One can then writes the following local entropy balance in the absence of irreversible mechanical dissipation (plasticity)

$$T \left(\frac{\partial S}{\partial t} + (\rho_f s_f q_i)_{,i} \right) = -Q_{i,i} + \rho r$$

where ρr is a thermal volumetric source / sink term.

Using the mass conservation, and the constitutive relation for the porous material entropy

$$T \left(\beta K \frac{\partial \epsilon_{kk}}{\partial t} - \beta_m \frac{\partial p}{\partial t} \right) + C_d \frac{\partial T}{\partial t} + T q_i \times (\rho_f s_f)_{,i} = -Q_{i,i} + \rho r$$

Linearizing around the initial temperature, noting that $\rho_f ds_f = -\beta_f dp + \rho_f C_p \frac{dT}{T}$

$$T^o \left(\beta K \frac{\partial \epsilon_{kk}}{\partial t} - \beta_m \frac{\partial p}{\partial t} \right) + \rho C \frac{\partial T}{\partial t} + q_i \times (\rho_f C_p (T - T^o) - \beta_f T^o (p - p^o))_{,i} = -Q_{i,i} + \rho r$$

Moreover the latent heat due to deformation and pore pressure changes are always much smaller that the one associated with compare to the heat capacity, i.e $\beta K \ll C_d$ and $\beta_m dp \ll \rho C$, similarly $\beta_f dp \ll \rho_f C_p$ such that we can write the heat equation as:

$$\rho C \frac{\partial T}{\partial t} + [\rho_f C_p (T - T^o)]_{,i} q_i + Q_{i,i} = \rho r$$

If convection is negligible - which the cases in a large number of applications (the noticeable exception being geothermal energy) we have

$$\rho C \frac{\partial T}{\partial t} + Q_{i,i} = \rho r$$

Fourier's law relates heat flux to temperature gradient

$$Q_i = -\lambda T_{,i}$$

where λ is the thermal heat conductivity.

	λ
	W/m/K
Granite	2.4
Carbonate	2.1-2.9
Chalk	0.9
Marble	3.15
Schist	0.9-1.6
Salt rock	11.3
Dolomite	3.15

Table 7.1: Some value for the thermal conductivity of rock (note that it vary much less than the intrinsic permeability). Specific heat does not vary much between rocks - 0.5-1 kJ/kg/K.

7.3 Summary

In summary, we have the following conservation laws:

$$\sigma_{ij,j} + f_i = 0$$

$$\frac{\partial \zeta}{\partial t} + q_{i,i} = \gamma$$

$$\rho C \frac{\partial T}{\partial t} + Q_{i,i} = \rho r$$

Constitutive relations

$$\begin{aligned} \sigma - \sigma^o &= K \epsilon_{kk} - \alpha(p - p^o) - \beta K(T - T^o) & s_{ij} &= 2G \epsilon_{ij} \\ \zeta &= \alpha \epsilon_{kk} + \frac{p - p^o}{M} - \beta_m(T - T^o) \end{aligned}$$

Transport laws: Darcy & Fourier

$$q_i = -\frac{k}{\mu_f} (p + \rho_f g x_3)_{,i}$$

$$Q_i = -\lambda T_{,i}$$

Initial and boundary conditions Same as before for in-situ stresses and pore-pressure. For temperature, similarly we will solve for variation from an initial temperature field. (See next paragraph).

7.4 The geothermal gradient

The distribution of initial Temperature at depth - like the initial pore-pressure profile - is increasing with depth although deviation occurs (Heat anomalies etc.). In continental Europe in average (apart from anomalies), the Temperature increases roughly of $3^{\circ}C$ per 100 meters, i.e. $0.03^{\circ}C/m$. Thus, one needs to drill to $\sim 4.5km$ to reach $150C$ (adding a average surface temperature of $15C$). Note that counter-example abounds, for example a temperature of $200^{\circ}C$ is reached at 300 meters depth in the geothermal field of Lardarello (Tuscany, Italy)!

Chapter 8

Failure of geomaterials - recaps

The notes here are a brief summary - and do not replace your geomechanics course notes.

8.1 Geomaterials fail in term of Terzaghi's effective stress

It is extremely important to know that geomaterials fail in terms of Terzaghi effective stress: in our convention $\sigma_{ij} + p\delta_{ij}$, i.e. experimentally we can define a yield surface as $f(\sigma_{ij} + p\delta_{ij})$ (see Fig.8.1 for an experimental validation of Terzaghi's effective stress under tri-axial conditions). This is because for rocks (and soils), most -if not all- of the volumetric part of the plastic deformation corresponds to plastic variation of porosity. The (irreversible) plastic work increment δW^p is of the form

$$\delta W^p = \sigma_{ij} d\epsilon_{ij}^p + p d\varphi^p$$

For plastically incompressible solid constituents ($\epsilon_s^p = 0$), i.e. all the volumetric irreversible strain of the porous material are due to porosity change (compaction or dilatancy) such that $\epsilon_{kk}^p \approx \varphi^p$. In that case, we see that driving force for plastic deformation is $\sigma_{ij} + p\delta_{ij}$, i.e. $\delta W^p = (\sigma_{ij} + p\delta_{ij}) d\epsilon_{ij}^p$. The driving force for poro-plastic deformation is Terzaghi's effective stress

$$\sigma'_{ij} = \sigma_{ij} + p\delta_{ij}$$

Recall that the driving force for poro-elastic deformation is the Biot's effective stress : $\sigma_{ij} + \alpha p\delta_{ij}$. The two are generally different for rocks (not for soils for which $\alpha = 1$ in most case) - this important difference has brought much confusion in the literature.

Important Remarks

- In rock/soil mechanics, the convention used by most practitioner is of positive stress in compression (positive strain in reduction), such that using such convention the linear

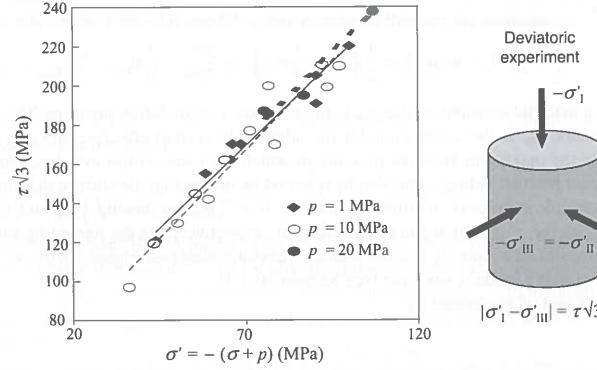


Figure 10.9 Investigation, through a deviatoric experiment, of the relevance of extension (Equation (10.70), of the loading function in order to account for a nonzero pore pressure (data from Vincké, O., Boutéca, M., Piau, J-M. and Fourmaintraux D. (1998). Study of the effective stress at failure *Poromechanics, A tribute to M.A. Biot, Proceedings of the First Biot conference*, eds. Thymus et al., Balkema)

Figure 8.1: Validity of Terzaghi's effective stress at failure. Deviatoric triaxial experiments - Study of the effective stress at failure for different pore-pressure. It clearly demonstrate the validity of Terzaghi's effective stress to check the stress at failure for a porous material. Taken from Coussy (2010).

poroelastic constitutive relation is:

$$\sigma_{ij} - \sigma_{ij}^o = c_{ijkl}\epsilon_{kk} + \alpha(p - p^o)\delta_{ij}$$

[as a pressure is >0 in compression by convention] and the Biot's effective stress tensor is thus $\sigma_{ij} - \alpha p \delta_{ij}$. and the Terzaghi's effective stress tensor (driving plastic flow) is $\sigma'_{ij} = \sigma_{ij} - p \delta_{ij}$.

- In this course, we stick to the continuum mechanics convention, but when plotting e.g. Mohr Circle etc. , we plot in terms of $-\sigma'_{ij}$

8.2 Elasto-plastic constitutive law

Fig. 8.2 displays a schematic of a failure envelope $f(\sigma_{ij} + p\delta_{ij}, \chi)$ in the $q - P'$ median plane (i.e. writing f as function of the first two invariant of the effective stress tensor), where $q = \sqrt{3J_2}$ (J_2 is the second invariant of the stress deviator), and $P' = \sigma'_{kk}/3$ the effective mean stress. In the elastic regime, the stress within the porous material remains inside the yield surface ($f < 0$). As the yield surface is reached ($f = 0$), irreversible strain occurs (plastic strain) and the state of stress remains on the yield surface $f = 0$ (i.e values of $f > 0$ are not admissible).

The *small* strain and porosity variation can be splitted in an elastic and a plastic part:

$$\epsilon_{ij} = \epsilon_{ij}^e + \epsilon_{ij}^p \quad \varphi = \varphi^e + \varphi^p = \varphi = \varphi^e + \epsilon_{kk}^p$$

or in terms of increments using a dot for the (possibly fictitious) time derivative

$$\dot{\epsilon}_{ij} = \dot{\epsilon}_{ij}^e + \dot{\epsilon}_{ij}^p$$

The constitutive law thus becomes in incremental form:

$$\begin{aligned} \dot{\sigma} &= K(\dot{\epsilon}_{kk} - \dot{\epsilon}_{kk}^p) - \alpha \dot{p} & \dot{s}_{ij} &= 2G(\dot{\epsilon}_{ij} - \dot{\epsilon}_{ij}^p) \\ \dot{\varphi} - \dot{\varphi}^p &= \alpha(\dot{\epsilon}_{kk} - \dot{\epsilon}_{kk}^p) + \frac{\dot{p}}{N} \\ \dot{\zeta} &= \dot{\varphi} + c_f \phi_o \dot{p} \end{aligned}$$

for the elastic part, while plasticity, with a possible hardening/softening described via a variable χ (e.g. the volumetric plastic strain), is mathematically formulated by:

1. the fact that the stress must remain limited by the yield surface

$$f(\sigma_{ij} + p\delta_{ij}, \chi) \leq 0$$

2. the plastic strain derives from a potential as function of the stress state

$$\dot{\epsilon}_{ij}^p = \lambda \frac{\partial g}{\partial (\sigma_{ij} + p\delta_{ij})}$$

with g is denoted as the plastic potential and the plastic multiplier λ is necessary positive or null

$$\lambda \geq 0$$

3. and of course plastic flow occurs only if the stress state is such that it is exactly on the yield function (a condition often called as the complementary condition):

$$f(\sigma_{ij} + p\delta_{ij}, \chi) \lambda = 0$$

4. a law for the evolution of the hardening variable

$$\dot{\chi} = \lambda H(\sigma_{ij} + p\delta_{ij}, \chi)$$

Note that in addition, a consistency condition stipulates that if the state of stress is on the yield surface, it must remain on the yield surface if plastic flow occurs ($\lambda > 0$), which translates in:

$$\dot{f} \times \lambda = 0$$

where $\dot{f}$ is the possible variation of the yield surface due to hardening / softening

$$\dot{f} = \frac{\partial f}{\partial \sigma'_{ij}} \dot{\sigma}'_{ij} + \frac{\partial f}{\partial \chi} \dot{\chi}$$

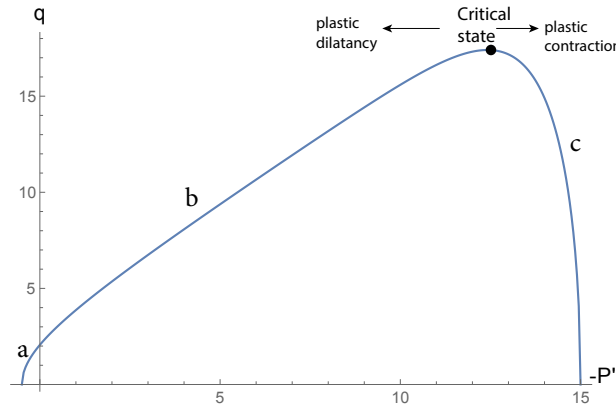


Figure 8.2: Yield surface in the $q/ - P'$ plane. $q = \sqrt{3J_2} = \sqrt{\frac{1}{2}((\sigma_I - \sigma_{II})^2 + (\sigma_{II} - \sigma_{III})^2 + (\sigma_I - \sigma_{III})^2)}$, $P' = (\sigma + p)$ in the convention of positive stress in tension.

8.3 Typical failure envelopes for geomaterials

Three typical “zones” can be roughly distinguished for the failure of geomaterials depending on the mean compressive stress. Referring to Fig.8.2: a) tensile / low confining stress non-linearity, b) frictional response (medium confinement) and c) compaction / pore-collapse (large confining stress). In zones a and b, the material dilates plastically, while in c), it contracts. The point where plastic flow occurs without volume changes is often denoted as the critical state.

Remarks

- The effect of the intermediate principal stress, although often neglected, can have an influence....
- Geomaterials are typically “non-associative”. It means that the function governing plastic flow rule g is different than the yield surface f , $\left(d\epsilon_{ij}^p = d\lambda \frac{\partial g}{\partial (\sigma_{ij} + p\delta_{ij})}\right)$ - or in others words the plastic flow does not occur in a direction perpendicular to the yield function in the stress space.
- hardening/softening is more or less complex - it is always better to do material characterization for the type of loading that will be encountered (instead of trying to be very general - which ends up in deriving models with zillions of parameters)

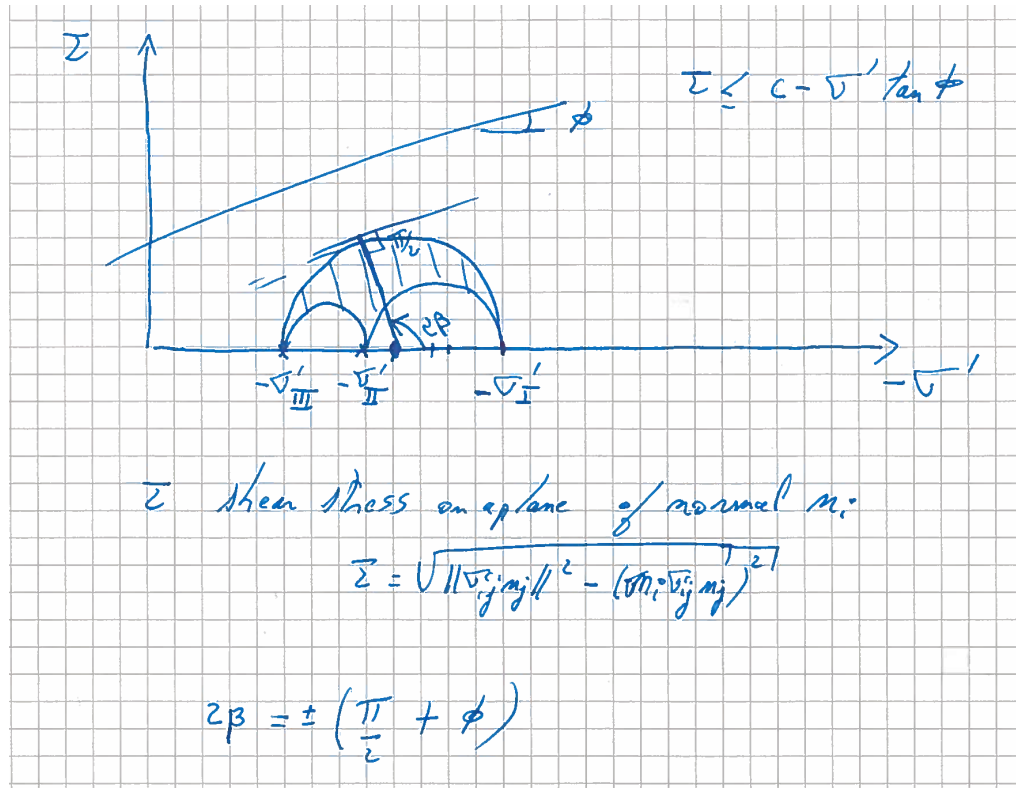


Figure 8.3: Mohr Diagram & Mohr-Coulomb failure envelope.

8.3.1 Mohr-Coulomb

Note that the frictional response is pervasive over a large range of confinement. It is usual to plot the failure envelope in the Mohr diagram (see Fig.8.3). The Mohr-Coulomb yield surface is written as (with stresses >0 in traction) :

$$f(\sigma_{ij} + p \delta_{ij}) = \tau - (C - (\sigma + p) \tan \phi) \leq 0$$

with τ the maximum shear stress (e.g. $\|\sigma_3 - \sigma_1\|/2$) and $-(\sigma + p)$ the effective mean confining stresses. C is the cohesion and ϕ the friction angle.

For Mohr-Coulomb, the maximum shear and corresponding effective normal stress τ_m and σ'_m are related (i.e. coordinates of the point touching the failure envelope) by:

$$\tau_m = (-\sigma'_m) \sin \phi + C \cos \phi$$

Note also that $C \cos \phi = UCS \times (1 - \sin \phi)/2$.

Remember that at low confining stress, Mohr-Coulomb is not adequate - although a “tangent” one can always be written with a friction angle increasing with decreasing confinement. A tensile strength σ_t (under uniaxial, zero confinement) is typically used to characterize tensile failure - from a strength perspective.

In term of principal (effective) stresses We can re-write the Mohr-Coulomb envelope in term of principal effective stress - (again convention positive stress in extension/ compressive stress <0 in compression) - $\sigma_I < \sigma_{III} < 0$

$$f(\sigma'_I, \sigma'_{III}) = (\sigma'_{III} - \sigma'_I) + (\sigma'_I + \sigma'_{III}) \sin \phi - 2C \cos \phi$$

or

$$f(\sigma'_I, \sigma'_{III}) = -\sigma'_I(1 - \sin \phi) + \sigma'_{III}(1 + \sin \phi) - 2C \cos \phi$$

Note - undrained response In that case, for soil ($\alpha \sim 1$ and typically soils exhibit a nearly incompressible response under undraind loading - i.e. $\nu_u \sim 1/2$), the mean effective stress remains constant, such that the deviatoric load at failure is

$$|\sigma'_{III} - \sigma'_I| = 2C \cos \phi + |\sigma'_I + \sigma'_{III}| \sin \phi$$

i.e. in a geotechnical setting, $|\sigma'_I + \sigma'_{III}| = (K_o + 1)\sigma_v'^o$, and one has the following expression for the undrained shear strength c_u

$$c_u \approx C \cos \phi + \frac{1}{2}(K_o + 1)\sigma_v'^o \sin \phi$$

Such an approximation is grossly over-estimated in the case of normally consolidated soil - **and should not be used!** Moreover it is an under-estimation for very over-consolidated soil. If sufficient experimental data is available, it is recommended to use a modified Cam-Clay model to properly capture the undrained behavior of soils. At minima, if using Mohr-Coulomb, a non-associated flow rule should be used to minimize dilatancy (we will discuss that in more details in the coming weeks).

8.3.1.1 Drucker-Prager equivalent

The Drucker Prager is a smooth/continuous version of the Mohr-coulomb criteria (without corner points) written in terms of $q = \sqrt{3J_2} = \sqrt{\frac{1}{2}((\sigma_I - \sigma_{II})^2 + (\sigma_{II} - \sigma_{III})^2 + (\sigma_I - \sigma_{III})^2)}$ and the mean stress $-P$ (again here in the convention of positive stress in tension)

$$f(q, -P') = q - \xi(-P') - k$$

which with values

$$\xi = 6 \frac{\sin \phi}{3 - \sin \phi} \quad k = 6C \frac{\sin \phi}{3 - \sin \phi}$$

fit the outer Mohr-Coulomb in the π -plane. Note that it is only an approximated equivalence - and the experimental data appears to be closer to the original Mohr-Coulomb criteria.

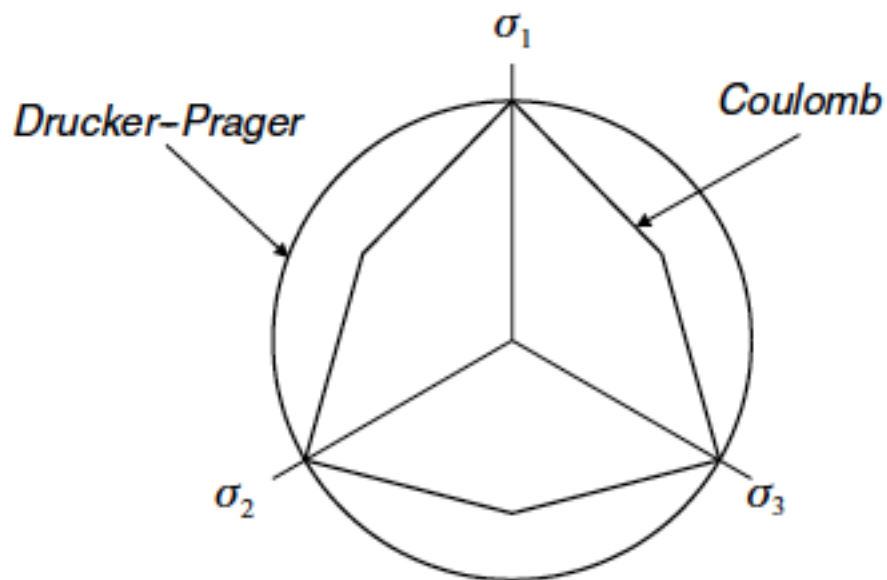


Figure 3.14. The Drucker–Prager and Coulomb yield surfaces.

Figure 8.4: Equivalence between Drucker-Prager and the Mohr Coulomb criterium on the so-called π -plane. Taken from Davis & Selvadurai (2005).

8.3.2 Tresca

It is a subset of Mohr-Coulomb, with $\phi = 0$. A criteria valid for soils under undrained loading (where $C = c_u$ or sometimes written also s_u in the soil mechanics literature).

8.3.3 Modified Cam-Clay

I refer you to your Geomechanics class notes - see also Muir Wood (1990). The Modified Cam-Clay is a volumetric hardening/softening model. In addition it includes the non-linear dependence of the volumetric stress-strain relation even in the elastic reversible regime.

Here writing in this section $p = -P'$ as the mean effective stress (be careful p is not the pore-pressure here) - positive in compression - to keep in line with classical expression of the model. The modified Cam-Clay model reads as

$$f(p, q, \chi) = q^2 - M^2 p(\chi - p)$$

where χ is a stress hardening variable (i.e. χ is the pre-consolidation pressure). The yield surface has the form of an ellipse in the $p-q$ plane with the critical state line $q = Mp$ passing through the top point. Plastic flow is taken associated, and the hardening rule is

$$\dot{\chi} = \chi \frac{v}{c^p - c^e} \dot{\epsilon}_{kk}^p$$

where c^p is the compression index, c^e the recompression index and $v = 1 + e$ is the soil specific volume (e is void ratio).

In addition the reversible elastic response of the soil is non-linear in its volumetric part- the elastic part of the strain (i.e. the recompression line) is related to (effective) stress as:

$$v = v_1 - c^e \ln p$$

where c^e is the recompression index and v_1 the soil specific volume at a reference state- note $dv/v = -d\epsilon_{kk}$ (soil mechanics convention of strain >0 in compression used in this subsection here).

8.3.4 Failure of faults / pre-existing fractures

In the case of pre-existing fracture / fault of normal n_i , a Coulomb criteria can be directly written relating the effective normal stresses ($\sigma'_n = n_i \sigma_{ij} n_j + p$) to the plane to the maximum shear stresses $\tau = \sqrt{n_i \sigma_{ij} \sigma_{jk} n_k - (n_i \sigma_{ij} n_j)^2}$ on the plane. The Coulomb yield surface is

$$\tau \leq C - \mu \times \sigma'_n$$

(again convention of positive stress in tension), where $\mu = \tan \phi$ is the friction coefficient. Usually, cohesion is negligible compared to the shear stress at depth. Compilation of data suggest $\mu \approx 0.6 - 0.8$ (Byerlee law - see Fig. 8.5).

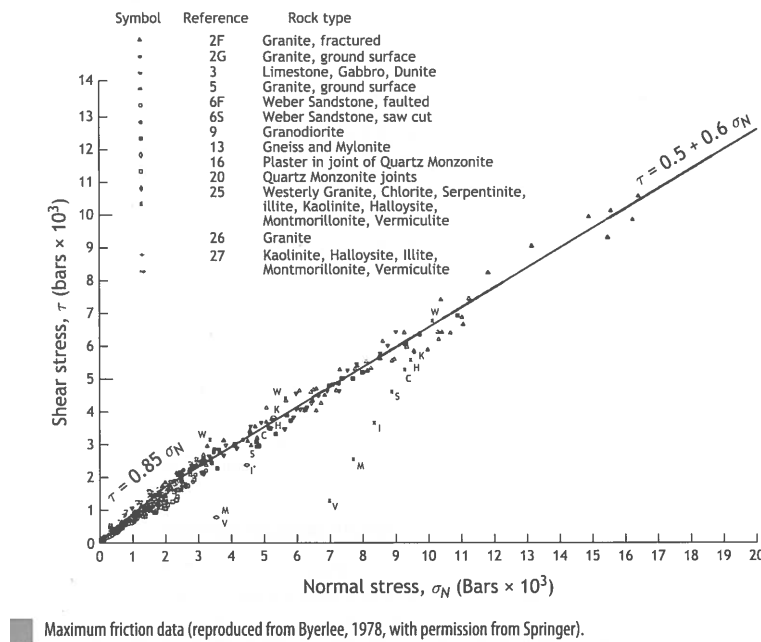


Figure 8.5: Joint/fracture friction. Taken from Cornet (2015).

Note that faults can be either frictionally weakening or hardening with slip: i.e. their friction coefficient decays (weakening) / increase (hardening) with accumulated slip until a residual value is reached. This is a very important feature as an initially stable failure can lead to a dynamic instability in the weakening case (i.e. an earthquake occurs) ! Frictional properties are also typically rate dependent.

Faults are also modeled with so-called rate and state friction models. In these models, an internal state variable is able to capture the recovery of frictional properties (healing) as well as weakening/hardening. These models are very popular in seismology as they allow to capture the complete earthquake cycle (note that they assume that a fault “always” slips).

Chapter 9

Elasto-plasticity I

Here, we leave hydro-mechanical couplings aside for simplicity and focus on the solution of small-strain elasto-plastic problems - with an emphasis on constitutive models relevant for geomaterials (frictional, non-associated etc.). Leaving hydro-mechanical coupling aside, we thus write the yield surface as function of the stress tensor σ_{ij} for simplicity. When tackling poroplastic problem, we 'just' need to switch to Terzaghi's effective stress $\sigma'_{ij} = \sigma_{ij} + p\delta_{ij}$ - which is the stress driving plastic deformation - note also that $\varphi^p = \epsilon^p_{kk}$.

In what follows, we denote the time-derivative with a dot. These notes do not replace a good textbook on elasto-plasticity, for example Lubliner, J. (1990), Davis & Selvadurai (2005). For numerical methods in elasto-plasticity, I recommend de Souza Neto et al. (2011) for a good description of practical algorithms, Simo & Hughes (1998) for more advanced/mathematical details.

9.1 Elasto-plastic mathematical description - recaps

Without particularizing to a given failure model, the elasto-plastic constitutive law accounting for hardening/softening is

- strain rate decomposition

$$\dot{\epsilon}_{ij} = \dot{\epsilon}^e_{ij} + \dot{\epsilon}^p_{ij}$$

- Elastic relations between increment of stress and elastic part of the strain rate $\dot{\epsilon}^e_{ij} = \dot{\epsilon}_{ij} - \dot{\epsilon}^p_{ij}$:

$$\dot{\sigma}_{ij} = c_{ijkl}(\dot{\epsilon}_{ij} - \dot{\epsilon}^p_{ij})$$

- Yield function

$$f(\sigma_{ij}, \chi) \leq 0$$

where χ represents hardening variable (for simplicity, we will assume there is a single hardening variable)

- Plastic flow rule (possibly non-associated if $f \neq g$), and evolution of the hardening variable

$$\begin{aligned}\dot{\epsilon}_{ij}^p &= \lambda \frac{\partial g}{\partial \sigma_{ij}} \quad \lambda \geq 0 \\ \dot{\chi} &= \lambda H(\sigma_{ij}, \chi)\end{aligned}$$

λ is often denoted the plastic multiplier.

- Complementaty condition

$$\lambda f(\sigma_{ij}, \chi) = 0$$

which implies that (with the previous conditions on yield and plastic multiplier):

$$\begin{aligned}\lambda &> 0 \text{ if and only if } f = 0 \\ \lambda &= 0 \text{ if and only if } f < 0\end{aligned}$$

- Consistency condition

$$\lambda \dot{f} = 0$$

which implies that (with the previous conditions on yield and plastic multiplier):

$$\begin{aligned}\lambda &> 0 \text{ if and only if } \dot{f} = 0 \\ \lambda &= 0 \text{ if and only if } \dot{f} < 0\end{aligned}$$

in other words if the material flow plastically, the stress (and hardening variables) evolves such that the condition $f = 0$ remains satisfied. However if the state of stress or hardening is such that we move away from yield - no plastic flow occur.

Note that the conditions

$$f \leq 0 \quad \lambda \geq 0 \quad \lambda f = 0$$

are often referred to mathematically as the Karush-Kuhn-Tucker conditions (they are encountered in problems involving inequalities constraints - we have already encountered those when tackling unconfined flow).

9.2 Tri-axial test of a Mohr-Coulomb material

Let's now focus on a Mohr-Coulomb material in order to understand how one can solve the elasto-plastic constitutive relations in an incremental manner. We investigate the response of the material in a triaxial experiment where the stress and strain in field are uniform - such that the balance of momentum is trivially satisfied. This corresponds to a 0D problem

- which is very similar to what one will have to solve at each Gauss integration points when solving a elasto-plastic problem using the finite element method.

We re-write the Mohr-Coulomb enveloppe (again convention positive stress in extension/ stress <0 in compression): $\sigma_I < \sigma_{III} < 0$

$$f(\sigma_I, \sigma_{III}) = (\sigma_{III} - \sigma_I) + (\sigma_I + \sigma_{III}) \sin \phi - 2C \cos \phi$$

or as

$$f(\sigma_I, \sigma_{III}) = -\sigma_I(1 - \sin \phi) + \sigma_{III}(1 + \sin \phi) - 2C \cos \phi$$

Let's consider the case of a conventional tri-axial test (in-fact biaxial) under vertical displacement control - i.e. we will control the total vertical strain rate. The specimen is a core: cylinder of radius R and length L . In such a test, the stresses are homogeneous, and in cylindrical coordinates, we have

$$\begin{aligned} \sigma_{zz} &= \sigma_I \\ \sigma_{rr} &= \sigma_{\theta\theta} = \sigma_{III} \quad \sigma_{r\theta} = \sigma_{rz} = 0 \end{aligned}$$

[no shear in the absence of friction at the loading platten/specimen interface which is never truly the case]. The axis of the core is the direction of major principal stress.

Under homogeneous deformation, the displacement field takes the form

$$[u_r(r), u_\theta = 0, u_z(z)]$$

and the only non-zero strain components are:

$$\epsilon_{rr} = \frac{\partial u_r}{\partial r} \quad \epsilon_{\theta\theta} = \frac{u_r}{r} \quad \epsilon_{zz} = \frac{\partial u_z}{\partial z}$$

The elastic part of the constitutive relation reduces thus to

$$\begin{aligned} \epsilon_z^e &= (1 + \nu) \frac{\sigma_{zz}}{E} - \frac{\nu}{E} (\sigma_{zz} + 2\sigma_{rr}) \\ \epsilon_{rr}^e &= (1 + \nu) \frac{\sigma_{rr}}{E} - \frac{\nu}{E} (\sigma_{zz} + 2\sigma_{rr}) \\ \epsilon_{\theta\theta}^e &= (1 + \nu) \frac{\sigma_{rr}}{E} - \frac{\nu}{E} (\sigma_{zz} + 2\sigma_{rr}) \end{aligned}$$

9.2.1 Hydrostatic confinement stage

The first stage of the test consist in the application of a hydrostatic confining pressure (this is a hydrostatic compression) - no failure occurs during that stage for a mohr-Coulomb material, and the state of stress at the end of this stage is:

$$\sigma_I = \sigma_{III} = -\sigma_c$$

and the axial and volumetric strain (purely elastic) are simply

$$\begin{aligned} \epsilon_{zz} &= -(1 - 2\nu)\sigma_c/E \\ \epsilon_{kk} &= -3(1 - 2\nu)\sigma_c/E \end{aligned}$$

9.2.2 Deviatoric loading stage

We now keep the radial confining pressure constant, and increase the axial load (or more precisely in the test we control vertical total strain rate $\dot{\epsilon}_{zz}$).

It is easy to obtain the compressive axial stress $\sigma_I = -\sigma_a$ where plasticity will occur, i.e. when $f = 0$

$$\sigma_a = \sigma_c \frac{(1 + \sin \phi)}{(1 - \sin \phi)} + 2C \frac{\cos \phi}{(1 - \sin \phi)}$$

the total axial strain (when first reaching this point is yet the one given by elasticity):

$$\begin{aligned} -\epsilon_{zz} &= (1 + \nu) \frac{\sigma_a}{E} - \frac{\nu}{E} (\sigma_a + 2\sigma_c) = \frac{\sigma_a}{E} - \frac{2\nu}{E} \sigma_c \\ -\epsilon_{kk} &= \frac{(1 - 2\nu)}{E} (\sigma_a + 2\sigma_c) \end{aligned}$$

Let's now look at how the material deforms plastically when we keep increasing the load / total axial displacement. We will assume, for sake of discussion, that the flow rule may be non-associated and write the plastic potential as

$$g = -\sigma_I(1 - \sin \psi) + \sigma_{III}(1 + \sin \psi)$$

where ψ is a dilatancy angle ($\psi = \phi$ in case of associated flow). Moreover, we will investigate hardening / softening: assuming that the cohesion evolves linearly as function of accumulated volumetric plastic strain (for illustration - other evolution could be thought of)

$$C(\chi = \epsilon_{kk}^p) = C + H\epsilon_{kk}^p$$

The plastic flow rule is (moving the sign to the RHS - recall deformation in contraction are negative)

$$\begin{aligned} -\dot{\epsilon}_{zz}^p &= \lambda(1 - \sin \psi) \\ -\dot{\epsilon}_{rr}^p &= -\lambda(1 + \sin \psi) \\ -\dot{\epsilon}_{kk}^p &= \lambda((1 - \sin \psi) - 2(1 + \sin \psi)) = -\lambda(1 + 3 \sin \psi) \end{aligned}$$

we see that because $\lambda > 0$, the plastic volumetric strain is always positive - i.e. always dilatant ! This is a particular feature of the basic Mohr-Coulomb criteria - (one typical extension is to make the dilatancy angle decrease with accumulated strain to make it more realistic).

We now solve for the plastic multiplier λ The consistency condition ($\lambda \dot{f} = 0$) tells us that as long as the material satisfies the yield function and flows plastically ($\lambda > 0$), then the state of stress must remain on the yield surface : i.e. $\dot{f} = 0$, in other words

$$\dot{f} = \frac{\partial f}{\partial \sigma_{ij}} \dot{\sigma}_{ij} + \frac{\partial f}{\partial \chi} \dot{\chi} = 0$$

where $\chi = \epsilon_{kk}^p$ here. Particularizing for our case, we obtain

$$-(1 - \sin \phi) \dot{\sigma}_{zz} + (1 + \sin \phi) \dot{\sigma}_{rr} - 2H \cos \phi \dot{\epsilon}_{kk}^p = 0$$

Moreover here, the confining pressure is kept constant such that $\dot{\sigma}_{rr} = 0$. such that we obtain

$$-(1 - \sin \phi) \dot{\sigma}_{zz} - 2H \cos \phi (1 + 3 \sin \psi) \lambda = 0$$

and also (because $\dot{\sigma}_{rr} = 0$.)

$$\dot{\sigma}_{zz} = (\dot{\epsilon}_{zz} - \dot{\epsilon}_{zz}^p) E$$

such that we obtain

$$(1 - \sin \phi) (\dot{\epsilon}_{zz} - \dot{\epsilon}_{zz}^p) E + 2H \cos \phi (1 + 3 \sin \psi) \lambda = 0$$

from which we can solve for λ

$$\lambda = -\dot{\epsilon}_{zz} \frac{E(1 - \sin \phi)}{E(1 - \sin \phi)(1 - \sin \psi) + 2H \cos \phi (1 + 3 \sin \psi)}$$

Note that because it is a compressive test $\dot{\epsilon}_{zz} < 0$ such that we recover $\lambda > 0$.

Axial stress rate We finally obtain the axial stress rate as:

$$\dot{\sigma}_{zz} = \left(1 - \frac{E(1 - \sin \phi)(1 - \sin \psi)}{E(1 - \sin \phi)(1 - \sin \psi) + 2H \cos \phi (1 + 3 \sin \psi)} \right) E \dot{\epsilon}_{zz}$$

Perfectly plastic case ($H = 0$) In the perfectly plastic case ($H = 0$), we have

$$\lambda = -\dot{\epsilon}_{zz} \frac{1}{(1 - \sin \psi)}$$

and

$$\dot{\sigma}_{zz} = E(\dot{\epsilon}_{zz} + \lambda(1 - \sin \psi)) = \dot{\epsilon}_{zz} \times E \times \left(1 - \frac{(1 - \sin \psi)}{1 - \sin \psi} \right) = 0$$

we recover that in the perfectly plastic case, the stress remains constant during plastic flow $\dot{\sigma}_{zz} = 0$.

Total volumetric strain rate The total volumetric strain rate (note one can measure the total volumetric strain by adding the vertical and radial strain measurements) can be obtained as the sum of the elastic and plastic strain rate - expressing it as function of the applied vertical strain rate:

$$\begin{aligned}
 \dot{\epsilon}_{kk} &= \dot{\epsilon}_{kk}^e + \dot{\epsilon}_{kk}^p \\
 &= \frac{(1-2\nu)}{E} \dot{\sigma}_{zz} - \dot{\epsilon}_{zz} \frac{E(1-\sin\phi)(1+3\sin\psi)}{E(1-\sin\phi)(1-\sin\psi) + 2H \cos\phi(1+3\sin\psi)} \\
 &= \left((1-2\nu) \left(1 - \frac{E(1-\sin\phi)(1-\sin\psi)}{E(1-\sin\phi)(1-\sin\psi) + 2H \cos\phi(1+3\sin\psi)} \right) \right. \\
 &\quad \left. - \frac{E(1-\sin\phi)(1+3\sin\psi)}{E(1-\sin\phi)(1-\sin\psi) + 2H \cos\phi(1+3\sin\psi)} \right) \dot{\epsilon}_{zz}
 \end{aligned}$$

9.2.3 Full Stress-strain curve

Figure 9.1 displays some example of stress-strain curves (axial stress versus axial strain and volumetric strain) for a) a perfectly plastic associated material, b) a non-associated perfectly plastic material, c) a non-associated hardening material and d) a softening non-associated material where in the cases b) to d) we have $\psi < \phi$.

9.2.4 Summary

We have seen in the simple case where the stress field is homogeneous how to solve for material deformation. For any practical geometry, the stress/strain fields are not spatially homogeneous. We will need to couple the solution of the balance of momentum with the solution of the non-linear elastoplastic constitutive law. We will have to do this at every Gauss point where the stresses are estimated in a finite element setting. Before moving to such a problem, as a first step, to introduce numerical methods, we investigate the finite element solution of the non-linear elastic case first.

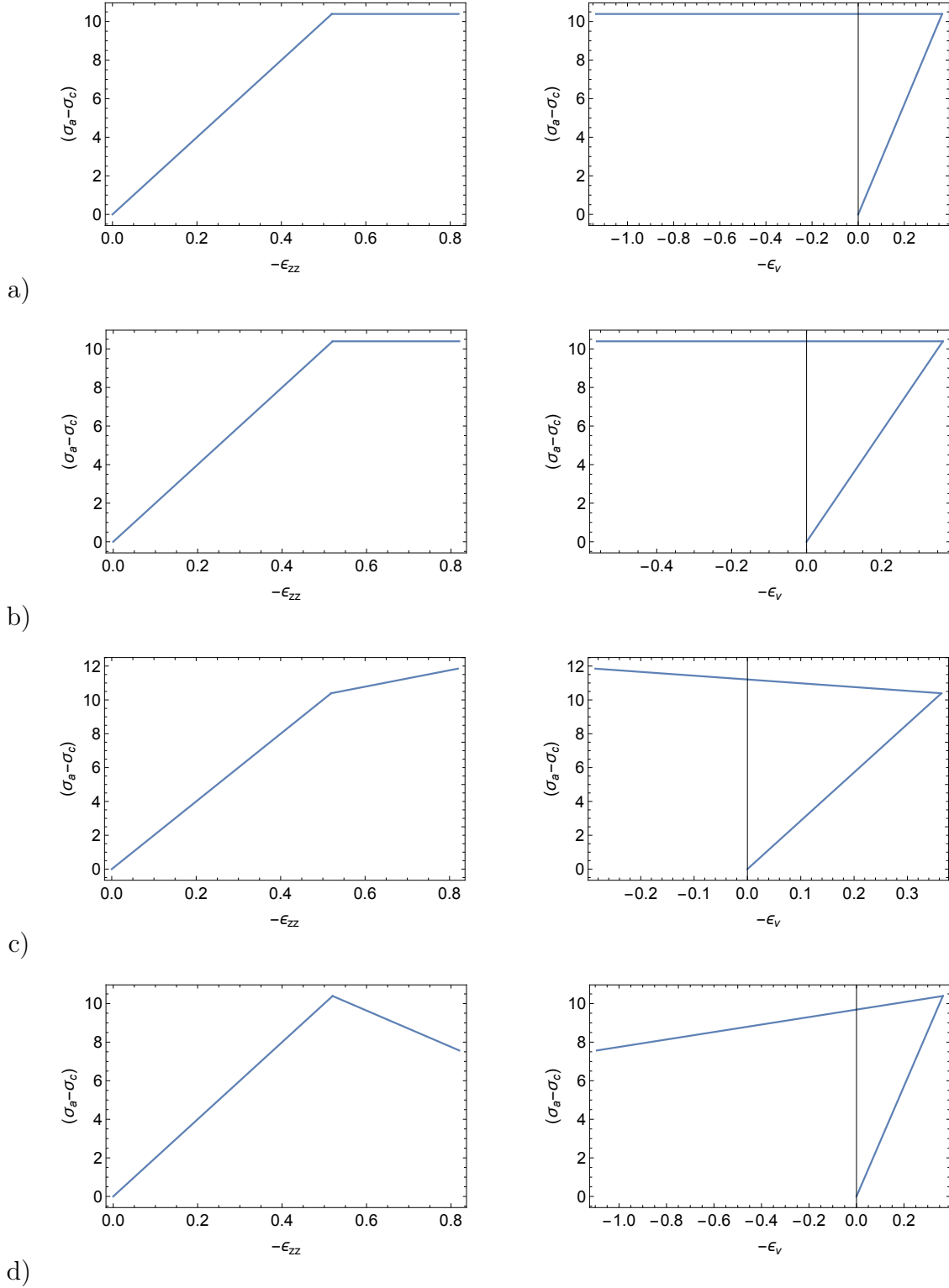


Figure 9.1: Example of stress-strain curve for a MC material during the deviatoric stage of a conventional triaxial test ($\sigma_c = 12$) - all stress/stiffness in e.g. MPa. The base materials parameters are: $E = 20$, $\nu = 0.15$, $\phi = 30^\circ$, $C = 3$

a) associated perfectly plastic ($\psi = \phi, H = 0$), b) non-associated perfectly plastic ($\psi = 2/3\phi, H = 0$), c) non-associated hardening ($\psi = 2/3\phi, H = 0.6$), d) non-associated softening ($\psi = 2/3\phi, H = -0.6$)

9.3 Toward the solution of elasto-plastic problems with FEM: non-linear small strain elasticity

Let's focus on a reversible non-linear elastic behavior. For a material in domain Ω , the boundary value problem reads

$$\begin{aligned}\sigma_{ij,j} + f_i &= 0 && \text{in } \Omega \\ t_i = \sigma_{ij}n_j &= t_i^g(t) && \text{on } \Gamma_{t_i} \\ u_i &= u_i^g(t) && \text{on } \Gamma_{u_i} \\ \sigma_{ij} &= c_{ijkl}(\epsilon_{ij})\epsilon_{ij} && \text{in } \Omega \\ \epsilon_{ij} &= \frac{1}{2}(u_{i,j} + u_{j,i})\end{aligned}$$

where the stiffness tensor is now taken as a smooth / continuous function of the current strain for example. The problem is now path dependent and as a result it is time-dependent. Let's write the evolution of the load via a multiplier, i.e.

$$t_i^g = \alpha T_i^g$$

and we step by increase α from 0 to 1 (imposed)- stepping in “artificial” time. We will thus solve for the balance of momentum and the constitutive relation at every step - following the deformation of the material as the load increases.

The weak-form / principle of virtual power reads in the absence of body forces $f_i = 0$:

$$\int_{\Omega} \epsilon_{ij}(v_i) \sigma_{ij}(\epsilon_{ij}(u_i)) \, dV = \alpha \int_{\Gamma_{t_i}} v_i T_i^g \, dS$$

The left hand side is often referred to as the internal forces and the RHS as the external forces. Upon discretization by finite element, we can write a residual at a given load step from t_n to t_{n+1} , i.e. from α_n to α_{n+1} . In a finite element method, our aim is to solve for the displacement at $\mathbf{u}_{n+1}$ (vector of unknowns at the nodes) . We build the residual of the weak form. Upon discretization via finite element

$$\mathbf{r}(\mathbf{u}_{n+1}) = \mathbf{f}_{int}(\mathbf{u}_{n+1}) - \alpha_{n+1} \mathbf{f}_{ext}$$

where at the element level

$$\begin{aligned}\mathbf{f}_{int}^e &= \int_{\Omega_e} \mathbf{B}^T \sigma(\mathbf{u}_{n+1}) \, dV \\ \mathbf{f}_{ext}^e &= \int_{\Gamma_{te}} \mathbf{N}^T T \, dS\end{aligned}$$

As the problem is non-linear, for a given state, we must minimize the residuals. Using the Newton-Raphson algorithm, we expand to first order the residuals - around $\mathbf{u}_{n+1}^k$, we write

$$\mathbf{r}(\mathbf{u}_{n+1}^{k+1}) = \mathbf{r}^k(\mathbf{u}_{n+1}^k) + \frac{\partial \mathbf{r}^k(\mathbf{u}_{n+1}^k)}{\partial \mathbf{u}_{n+1}^k} \delta \mathbf{u}_{n+1}^{k+1}$$

and enforcing that $\mathbf{r}(\mathbf{u}_{n+1}^{k+1}) = \mathbf{0}$, we obtain a tangent linear system for the increment

$$\frac{\partial \mathbf{r}^k(\mathbf{u}_{n+1}^k)}{\partial \mathbf{u}_{n+1}^k} \delta \mathbf{u}_{n+1}^{k+1} = -\mathbf{r}^k(\mathbf{u}_{n+1}^k)$$

upon its solution we write

$$\mathbf{u}_{n+1}^{k+1} = \mathbf{u}_{n+1}^k + \delta \mathbf{u}_{n+1}^{k+1}$$

$\frac{\partial \mathbf{r}^k(\mathbf{u}_{n+1}^k)}{\partial \mathbf{u}_{n+1}^k}$ is the so-called tangent stiffness matrix (e.g. Jacobian)

$$\frac{\partial \mathbf{r}^k(\mathbf{u}_{n+1}^k)}{\partial \mathbf{u}_{n+1}^k} = \mathbb{K}_{n+1}^k = \sum_e \int_{\Omega^e} \mathbf{B}^T \frac{\partial \sigma(\mathbf{u}_{n+1})}{\partial \mathbf{u}_{n+1}} \bigg|_{\mathbf{u}_{n+1}^k} dV = \sum_e \int_{\Omega^e} \mathbf{B}^T \frac{\partial \sigma(\mathbf{u}_{n+1})}{\partial \epsilon_{n+1}} \bigg|_{\epsilon_{n+1}^k} \mathbf{B} dV$$

The matrix $\mathbf{D} = \frac{\partial \sigma(\mathbf{u}_{n+1})}{\partial \epsilon_{n+1}} \bigg|_{\epsilon_{n+1}^k}$ is coined as the consistent tangent operator - linking stress - strain increment : it is dependent of the specific constitutive relation used. It results from a linearization of the incremental stress updating procedure¹. We see that we need to obtain the expression of $\mathbf{D}$ at every Gauss point in order to obtain the tangent stiffness matrix.

Summary of the FE algorithm for a path-dependent non-linear material

1. Assemble the global external force and set up the loading curve $\alpha(t)$
2. Initialize increment counter $i := 1$, start the loop on the load steps
3. Set the load factor for the step $[t_n, t_{n+1}]$, i.e α_{n+1}
4. Find the root of the residual

$$\mathbf{r}(\mathbf{u}_{n+1}) = \mathbf{f}_{int}(\mathbf{u}_{n+1}) - \alpha_{n+1} \mathbf{f}_{ext}$$

using a Newton-Raphson procedure:

¹Note using the *infinitesimal* tangential relation and linearizing the problem before its discretization results in poor accuracy and convergence. It is always better to linearize **after** discretization. Hence the term Consistent Tangent Operator.

- (a) Start a iterative loop (iterator k) - iterate upon convergence $\|\mathbf{r}(\mathbf{u}_{n+1}^{k+1})\| < \epsilon_{tol} \times (\alpha_{n+1} \|\mathbf{f}_{ext}\|)$
- (b) at iteration $k + 1$
Compute $\mathbf{r}^k$, assemble the tangent stiffness matrix $\mathbf{K}_{n+1}^k$, and solve for

$$\mathbf{K}_{n+1}^k \delta \mathbf{u}_{n+1}^{k+1} = -\mathbf{r}^k$$

9.4 Solution of elasto-plastic problem via FEM

The overall methodology is the same than the one described for the non-linear path dependent case above, but now the material constitutive behavior is more complicated.

Let us first recall, the time/load multiplier problem. In the absence of body forces ($f_i = 0$), the principle of virtual work reads :

$$\int_{\Omega} \epsilon_{ij}(v_i) \sigma_{ij}(\epsilon_{ij}(u_i)) \, dV = \alpha \int_{\Gamma_{t_i}} v_i T_i^g \, dS$$

Upon discretization by finite element, we can write a residual at a given load step from t_n to t_{n+1} , i.e. from α_n to α_{n+1} . In a finite element method, our aim is to solve for the displacement at $\mathbf{u}_{n+1}$ (vector of unknowns at the nodes). We build the residual of the weak form. Upon discretization via finite element

$$\mathbf{r}(\mathbf{u}_{n+1}) = \mathbf{f}_{int}(\mathbf{u}_{n+1}) - \alpha_{n+1} \mathbf{f}_{ext}$$

where at the element level

$$\begin{aligned} \mathbf{f}_{int}^e &= \int_{\Omega_e} \mathbf{B}^T \sigma(\mathbf{u}_{n+1}) \, dV \\ \mathbf{f}_{ext}^e &= \int_{\Gamma_{te}} \mathbf{N}^T T \, dS \end{aligned}$$

Now, let's re-write the displacement, stress and strain at $n + 1$ as

$$\begin{aligned} \mathbf{u}_{n+1} &= \mathbf{u}_n + \Delta \mathbf{u} \\ \sigma_{n+1} &= \sigma_n + \Delta \sigma \\ \epsilon_{n+1} &= \epsilon_n + \Delta \epsilon \end{aligned}$$

The aim is to find the root of the residuals $\mathbf{r}(\Delta \mathbf{u})$. Like in the non-linear elastic case, this is done using a Newton-Raphson algorithm. At every iterative steps of the Newton-Raphson algorithm, one needs to compute

1. the residual

$$\mathbf{r}(\Delta \mathbf{u}^k) = \mathbf{f}_{int}(\Delta \mathbf{u}^k) - \alpha_{n+1} \mathbf{f}_{ext}$$

in other words, to estimate the internal forces (assembly over all elements), one needs to estimate the stresses at the Gauss integration points for a given displacement field at the nodes $\mathbf{u}_{n+1}^k = \mathbf{u}_n + \Delta \mathbf{u}^k$. This is step of the algorithm is the constitutive integration steps / incremental stress updating procedure- also denoted the radial return mapping in elastoplasticity.

2. the tangent operator

$$\mathbb{K}^k = \frac{\partial \mathbf{r}^k(\mathbf{u}_{n+1}^k)}{\partial \mathbf{u}_{n+1}^k} = \sum_e \int_{\Omega^e} \mathbf{B}^T \frac{\partial \sigma(\mathbf{u}_{n+1})}{\partial \epsilon_{n+1}} \Big|_{\epsilon_{n+1}^k} \mathbf{B} \, dV$$

which is actually obtained from the linearization of the stress update procedure.

We now discuss these two steps.

9.4.1 Incremental stress update

The local stress update aims at obtaining the stress tensor σ_{n+1} given the strain increment $\Delta \epsilon$ over the load step (i.e. the displacement increment $\Delta \mathbf{u}$ gives the strain increment at the Gauss point of the element as $\mathbf{B} \Delta \mathbf{u}$ in finite element). Here I use the matrix notation throughout.

For an elasto-plastic material, the stress update is performed as follow.

1. Trial elastic state

First, during the load step, the deformation is supposed to occur elastically. In other words, $\Delta \epsilon^p = 0$, such that the elastic part of the strain is

$$\epsilon_{n+1}^e = \epsilon_n^e + \Delta \epsilon$$

and the trial stresses are just obtain as

$$\sigma_{n+1}^{trial} = \sigma_n + \mathbf{D}^e \Delta \epsilon$$

where $\mathbf{D}^e$ is the elastic stiffness matrix. Moreover, the hardening variable are assumed not to change (as no plastic deformation are assumed to occur):

$$\chi_{n+1}^{trial} = \chi_n$$

2. Plastic corrections

- (a) if $f(\sigma_{n+1}^{trial}, \chi_{n+1}^{trial}) < 0$, then the trial stresses are in the elastic range: no plastic deformation occurs for the given strain increment. The stress update is finished:

$$\sigma_{n+1} = \sigma_{n+1}^{trial} \quad \chi_{n+1}^{trial} = \chi_n$$

and

$$\frac{\partial \sigma(\mathbf{u}_{n+1})}{\partial \epsilon_{n+1}} = \mathbf{D}^e$$

[the elastic matrix is the consistent tangent operator]

- (b) if $f(\sigma_{n+1}^{trial}, \chi_{n+1}^{trial}) \geq 0$, then the trial stresses violate the yield function: plastic flow occur for the given strain increment. The stress state must be “returned” onto the yield function in order to satisfy it. This procedure is called the radial return mapping.

9.4.1.1 Plastic corrector step (Radial return mapping)

If the trial state violates the yield function, we need to account for plastic deformation and solve locally for the constitutive law. Let first, recall the hardening elastoplastic law

$$\Delta \epsilon^p = \lambda \frac{\partial g}{\partial \sigma} = \lambda \mathbf{N}(\sigma_{n+1}, \chi_{n+1})$$

where we have written $\frac{\partial g}{\partial \sigma} = \mathbf{N}$ for short. Moreover, the increment of hardening variable can be written as

$$\Delta \chi = \lambda \mathbf{H}(\sigma_{n+1}, \chi_{n+1})$$

where $\mathbf{H}$ is a hardening function (given / observed experimentally - recall the simple one we used for the triaxial test which was simply linear with a constant hardening modulus). Moreover, the stresses must satisfy the yield function, and the elastic relation

$$f(\sigma_{n+1}, \chi_{n+1}) = 0 \quad \sigma_{n+1} = \sigma_n + \underbrace{\mathbf{D}^e (\Delta \epsilon - \lambda \mathbf{N}(\sigma_{n+1}, \chi_{n+1}))}_{\Delta \sigma} = \sigma_{n+1}^{trial} - \lambda \mathbf{D}^e \mathbf{N}(\sigma_{n+1}, \chi_{n+1})$$

We thus must solve for λ , $\Delta \sigma$ and $\Delta \chi$ - to satisfy the following three equations (yield function, elastic relation, hardening / softening evolution)

$$\begin{aligned} f(\sigma_{n+1}, \chi_{n+1}) &= f(\Delta \sigma, \Delta \chi) = 0 \\ \Delta \sigma &= \mathbf{D}^e (\Delta \epsilon - \lambda \mathbf{N}(\Delta \sigma, \Delta \chi)) \\ \Delta \chi &= \lambda \mathbf{H}(\Delta \sigma, \Delta \chi) \end{aligned}$$

This of course depends on the constitutive model (expression for f , $\mathbf{N}$ and $\mathbf{H}$). Such a system can be solved via any root finding scheme (e.g. Newton-Raphson or others).

9.4.1.2 Consistent tangent operator

Because $\lambda > 0$ (plastic step), we must have

$$\Delta f = 0 = \left(\frac{\partial f}{\partial \sigma} \right)^T \Delta \sigma + \left(\frac{\partial f}{\partial \chi} \right)^T \Delta \chi$$

[recall the consistency condition: the stresses must remain on the yield surface during plastic deformation] - Writing $\frac{\partial f}{\partial \sigma} = \mathbf{N}_A$ and $\frac{\partial f}{\partial \chi} = \mathbf{X}$ for short, we obtain

$$\mathbf{N}_A(\Delta \sigma, \Delta \chi)^T \mathbf{D}^e (\Delta \epsilon - \lambda \mathbf{N}(\Delta \sigma, \Delta \chi)) + \lambda \mathbf{X}(\Delta \sigma, \Delta \chi)^T \mathbf{H}(\Delta \sigma, \Delta \chi) = 0$$

In the simple case where the yield function, plastic potential and hardening evolution are linear in term of stress, and hardening variable, then

$$\mathbf{H}(\Delta \sigma, \Delta \chi) = \mathbf{H}$$

etc. then we have

$$\begin{aligned} \mathbf{N}_A^T \mathbf{D}^e \Delta \epsilon &= \lambda (\mathbf{N}_A^T \mathbf{D}^e \mathbf{N} - \mathbf{X}^T \mathbf{H}) \\ \lambda &= \frac{\mathbf{N}_A^T \mathbf{D}^e \Delta \epsilon}{\mathbf{N}_A^T \mathbf{D}^e \mathbf{N} - \mathbf{X}^T \mathbf{H}} \end{aligned}$$

and we obtain the incremental stress-strain relation

$$\Delta \sigma = \left(\mathbf{D}^e - \frac{(\mathbf{D}^e \mathbf{N})(\mathbf{N}_A \mathbf{D}^e)^T}{\mathbf{N}_A^T \mathbf{D}^e \mathbf{N} - \mathbf{X}^T \mathbf{H}} \right) \Delta \epsilon$$

in other words the consistent tangent operator

$$\frac{\partial \sigma(\mathbf{u}_{n+1})}{\partial \epsilon_{n+1}} = \mathbf{D} = \left(\mathbf{D}^e - \frac{(\mathbf{D}^e \mathbf{N})(\mathbf{N}_A \mathbf{D}^e)^T}{\mathbf{N}_A^T \mathbf{D}^e \mathbf{N} - \mathbf{X}^T \mathbf{H}} \right)$$

More complicated constitutive relations For more complicated constitutive relation, the consistent tangent operator is more tedious to obtain. It is always the linearization of the local stress / total strain increment - which is obtained anyway when solving the plastic correction step using a Newton algorithm.

Note that the term radial return algorithm (or radial return mapping) is often used for the plastic corrector step - this terminology arises from a graphical representation of the scheme for the Von Mises criteria which is a simple circle in the principal stresses space: hence the term radial and return for the fact that we project back the stresses onto the circular yield function in that case.

9.5 Unstable equilibrium - The arc length method

In a number of cases, we are interested in determining the ultimate limit load of the structure - the load at collapse. The global (structure) load-displacement curve can exhibit instability: buckling, snap-through or snap-back (which can happen in the presence of softening). In these cases, the applied load must decrease to keep following an equilibrium path.

The idea behind the arc length method is to solve for both the load multiplier α_{n+1} in combination with $\mathbf{u}_{n+1}$. We write

$$\Delta\alpha = \alpha_{n+1} - \alpha_n$$

$$\Delta\mathbf{u} = \mathbf{u}_{n+1} - \mathbf{u}_n$$

and write the residuals as

$$\mathbf{r}(\Delta\mathbf{u}, \Delta\alpha) = \mathbf{f}_{int}(\mathbf{u}_{n+1}) - (\alpha_n + \Delta\alpha) \mathbf{f}_{ext}$$

The arc-length method needs an additional constraint to solve for the increment of the load multiplier. The most commonly use is to restrict the displacement increment $\Delta\mathbf{u}$ via a “cylindrical” constraint (a scalar equation):

$$\Delta\mathbf{u}^T \cdot \Delta\mathbf{u} = \ell^2$$

This ensure that the load increment will decrease if needed. It’s like switching to a “displacement” control whereas the boundary conditions are initially set in tractions.

It is possible to solve in a coupled fashion for the residual equations and the displacement increment constraint via a Newton-Raphson algorithm. This is the more robust method.

Other methods (partly coupled) attempt at keeping the usual Newton-Raphson update while updating for $\Delta\alpha$ in a staggered manner.

Chapter 10

Poro-plasticity

In this short chapter, we highlight important aspects of the impact of plasticity on the response of fluid saturated porous media. This set of notes recaps some of the points seen during the exercise session #10.

As discussed previously, porous material plastic response is driven by Terzaghi's effective stress:

$$\sigma'_{ij} = \sigma_{ij} + p\delta_{ij}$$

[Recall that this come from the fact plastic variation of porosity is directly equal to plastic bulk volumetric strain]. Note again in our convention stresses are positive in tension.

In porous media, the yield function and plastic potential is thus function of the effective stress tensor, and plastic flow is

$$\begin{aligned}\dot{\epsilon}_{ij}^p &= \lambda \frac{\partial g}{\partial \sigma'_{ij}} \\ \varphi^p &= \dot{\epsilon}_{kk}^p\end{aligned}$$

The variation of porosity $\varphi = \phi - \phi_o$ - just like the strain - is the sum of an elastic and a plastic part

$$\begin{aligned}\epsilon_{ij} &= \epsilon_{ij}^e + \epsilon_{ij}^p \\ \phi - \phi_o &= \varphi = \varphi^e + \varphi^p\end{aligned}$$

Before starting the discussion, it is important to point out that the variation of fluid content

$$\zeta = \frac{m_f - m_f^o}{\rho_f^o}$$

as discussed when introducing poroelasticity, we have seen that this variation of fluid content contains 2 parts: one related to the variation of porosity, one related to the fluid compressibility

$$\zeta = \varphi + \phi_o c_f (p - p_o)$$

where c_f is the fluid compressibility and p_o is the initial pore pressure.

The stress-strain poroelastic relation can thus be written accounting for plastic strain as [and for simplicity assuming zero initial stress and pore pressure]

$$\begin{aligned}\sigma_{ij} &= c_{ijkl}(\epsilon_{kl} - \epsilon_{kl}^p) - \alpha p \delta_{ij} \\ \varphi - \varphi^p &= \alpha(\epsilon_{kk} - \epsilon_{kk}^p) + \frac{p}{N}\end{aligned}$$

10.1 Undrained poroplastic response

Let's investigate the case of plasticity occurring under undrained conditions (e.g. sudden applications of a load). By definition in the undrained limit, the variation of fluid content is null, such that [for simplicity assuming zero initial stress and pore pressure]

$$\varphi = -\phi_o c_f p$$

i.e. we have

$$\varphi = \varphi^p + \alpha(\epsilon_{kk} - \epsilon_{kk}^p) + \frac{p}{N} \quad \varphi^p = \epsilon_{kk}^p$$

i.e.

$$\begin{aligned}\epsilon_{kk}^p + \alpha(\epsilon_{kk} - \epsilon_{kk}^p) + \frac{p}{M} &= 0 \\ p &= -M(\epsilon_{kk}^p + \alpha\epsilon_{kk}^e)\end{aligned}$$

Shear under constant total volumetric stress In that case, σ is constant [e.g. case of a test with ct confinement], but the effective stress is

$$\sigma' = \sigma - M(\epsilon_{kk}^p + \alpha\epsilon_{kk}^e)$$

we see that if under increasing shear stress the plastic response is dilatant, $\epsilon_{kk}^p > 0$, then the effective stresses decreases. In our convention, they become more compressive. So for a Mohr-Coulomb (MC) material, we move away from failure. This effect is refer to “dilatant hardening”.

We therefore see that in a structural context (where stress / strain are not uniform), an excess of dilatancy can lead to an over-estimation of the undrained collapse load. As already mentioned, an associated MC material leads to unbounded plastic dilatancy and thus should not be used under undrained loading. Non-associated MC is recommended, and if a degree of dilatancy is observed, it is always bounded - such that the dilatancy angle should eventually goes to zero after a given amount of inelastic strain.

If a Tresca model is used (with undrained cohesion), no dilatancy occur, and in all cases, the mean effective stress does not play a role on failure. A tresca model thus provides reasonable collapse load.

Of course, the complex response of a soil is better captured by a Modified Cam-Clay which as we have seen in the exercise provides good estimation of elasto-plastic collapse under undrained (and drained) conditions).

10.2 Drained poroplastic response

In the drained limit, all initial excess of pore-pressure have dissipated. The effective stresses reduces to the total stress. In that limit, a MC constitutive relation provides good estimated of collapse load.

Chapter 11

Limit analysis / rigid plasticity

We have seen (c.f. exercise week 10) that an elasto-plastic analysis allow to obtain the ultimate load above which unbounded plastic flow occur / in other word total collapse of the structure.

In this chapter, we briefly re-derive the lower and upper bound theorem of limit analysis - which allow to directly obtain the collapse load for rigid plastic material. We then briefly describe numerical techniques based on a combination of finite element and mathematical programming for the solutions of these class of problems.

Here for simplicity, we work in total stress and strain - the application to drained problems in fluid saturated geo-materials follows directly from the use of the effective stress where a Mohr-Coulomb is adequate. Undrained problems can be tackled in total stress with the use of a Tresca model.

11.1 The rigid plastic boundary value problem

When aiming at determining the collapse load of a (geo-)structure whose constituent behave in a ductile / elasto-plastic manner, neglecting elastic deformation is a reasonable assumption: at plastic collapse, most if not all the deformation will be irreversible. We therefore neglect any elastic strain and thus have:

$$\epsilon_{ij} = \epsilon_{ij}^p$$

When doing so, we will see that two useful theorems can be derived (for associated plastic flow) that allow to bound the ultimate limit load.

As usual, restricting to quasi-static, for a body Ω following an associated rigid plastic constitutive relation, the mechanical problem reads

$$\sigma_{ij,j} + f_i = 0 \quad \text{in } \Omega$$

in all points of Ω , the stress tensor is continuous and differentiable, but surfaces S_Σ (lines in 2D) where the displacement is discontinuous may exist (i.e. shear bands). Denoting n_i the normal to such surface of displacement discontinuity, the traction vector $t_i = \sigma_{ij}n_j$ must be continuous across S_Σ , in other words:

$$(\sigma_{ij}^- - \sigma_{ij}^+) n_j = [[\sigma_{ij}]] n_j = 0$$

where the $+$ and $-$ denotes the field above/below the surface of discontinuity. The notation $[[\cdot]]$ denotes the jump across the surface of displacement discontinuity.

The usual boundary conditions are:

$$\begin{aligned} \sigma_{ij}n_j &= t_i^g & \text{on } \Gamma_{t_i} \\ u_i &= u_i^g & \text{on } \Gamma_{u_i} \end{aligned}$$

with the usual non-overlapping restrictions on Γ_{u_i} and Γ_{t_i} . The rigid plastic constitutive relation for a yield function $f(\sigma_{ij})$ (no hardening) is

$$\begin{aligned} f(\sigma_{ij}) < 0 & \quad \dot{\epsilon}_{ij} = 0 & \quad \lambda = 0 \\ f(\sigma_{ij}) = 0 & \quad \dot{\epsilon}_{ij} = \lambda \frac{\partial f}{\partial \sigma_{ij}} & \quad \lambda > 0 \end{aligned} \tag{11.1}$$

$$\lambda f = 0 \tag{11.2}$$

where the strain rate tensor $\dot{\epsilon}_{ij}$ defined as

$$\dot{\epsilon}_{ij} = \frac{1}{2} (\dot{u}_{i,j} + \dot{u}_{j,i})$$

obeys the compatibility relations.

Important Note:

The lower bound and upper theorems are strictly valid only for associated plastic models (with a convex yield surface). However, one can 'violate' this mathematical proof, and perform limit analysis with a non-associated flow rule - a lower bound is no longer guarantee. However, it has been established ("experimentally") that a upper bound for an associated flow rule is also an upper bound for a non-associated flow rule.

11.1.1 Principle of virtual power

Denoting $\hat{u}_i$ a virtual velocity field, the principle of virtual power reads

$$\mathcal{P}_{ext}(\hat{u}_i) = \mathcal{P}_{int}(\hat{u}_i)$$

with the external power

$$\mathcal{P}_{ext}(\hat{u}_i) = \int_{\Omega} f_i \hat{u}_i \, dV + \int_{\Gamma_{t_i}} t_i^d \hat{u}_i \, dS$$

and the internal power - splitted in region of bulk plastic flow and along discontinuity of displacement

$$\mathcal{P}_{int}(\hat{u}_i) = \int_{\Omega/S_{\Sigma}} \sigma_{ij} \dot{\epsilon}_{ij}(\hat{u}_i) \, dV + \int_{S_{\Sigma}} \sigma_{ij} n_j \left[\left[\hat{u}_i \right] \right] \, dS$$

11.1.2 Load multiplier

It is usual to parametrize the load applied on a structure (or the gravity) as e.g. $t_i^g = \alpha t_i^c$ where t_i^c is the load at collapse - such that in that case $\alpha = 1$ at collapse. [Note that any other value than t_i^c can be used, -results can always be rescaled after].

Typically several loading parameters can be applied on the structure. The goal of limit analysis is to determine the domain K of “safe” load where collapse is prevented. This domain typically depends on:

1. the geometry of the structure
2. the loading
3. the type of failure criteria used.

but it does not depend on the loading path neither on the initial conditions.

In practice, K is approximated / bounded by static and kinematic approach.

- The static approach provides a lower-bound of K . It consists in finding a statically admissible stress field for a parametrized loading e.g. $(Q_1, Q_2) = \alpha \times (Q_1^*, Q_2^*)$ (where Q_1^*, Q_2^* are small values inside K) and then to maximise the loading multiplier while ensuring that the plastic criterium is satisfied in all points in the domain $f(\sigma_{ij}) \leq 0$.
- The kinematic approach provides an upper-bound of K . It consists in postulating a kinematically admissible displacement field describing the failure. Then the inner maximum dissipated power is computed while ensuring that the plastic criterium is satisfied. Using the principle of virtual work and minimising on the load multiplier, we obtain an upper bound of the collapse load.

11.2 Lower bound and upper bound theorems

We recall the principle of virtual power - when all the applied displacements are constant in time (for simplicity) - for any kinematically admissible displacement field

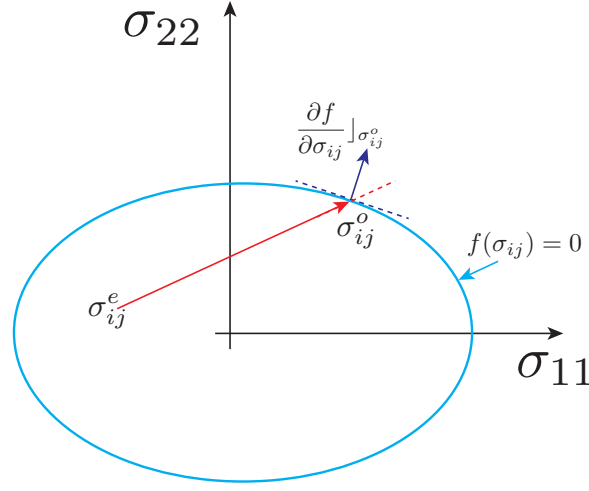


Figure 11.1: Example of yield surface in 2D ($f(\sigma_{ij}) = 0$). Illustration of the property of convexity $(\sigma_{ij}^o - \sigma_{ij}^e) \frac{\partial f}{\partial \sigma_{ij}} \big|_{\sigma_{ij}^o} > 0$.

$$\int_{\Omega/S_\Sigma} \sigma_{ij} \dot{\epsilon}_{ij}(\hat{u}_i) \, dV + \int_{S_\Sigma} \sigma_{ij} n_j \left[\left[\hat{u}_i \right] \right] \, dS = \int_{\Omega} f_i \hat{u}_i \, dV + \int_{\Gamma_t} t_i^d \hat{u}_i \, dS$$

For simplicity in the following, we assume that there are no shear bands $S_\Sigma = \emptyset$ - note that the same demonstration can be made relaxing that assumptions.

11.2.1 Lower bound

Let σ_{ij}^a be a statically admissible stress field corresponding to a load multiplier α^a satisfying the yield function $f(\sigma_{ij}^a) < 0$ in all points of the domain. The principle of virtual power for such a statically admissible stress field taking for the velocity field the exact solution $\dot{u}_i$

$$\int_{\Omega} \sigma_{ij}^a \dot{\epsilon}_{ij} \, dV = \alpha^a \int_{\Gamma_t} t_i^c \dot{u}_i \, dS + \int_{\Omega} (f_i) \dot{u}_i \, dV$$

where $\dot{u}_i$ is the solution of the velocity field and $\dot{\epsilon}_{ij}$ the corresponding strain rate.

For the exact stress field σ_{ij} at collapse, the corresponding value of the plastic multiplier at collapse $\alpha = 1$, we have

$$\int_{\Omega} \sigma_{ij} \dot{\epsilon}_{ij} \, dV = 1 \int_{\Gamma_t} t_i^c \dot{u}_i \, dS + \int_{\Omega} (f_i) \dot{u}_i \, dV$$

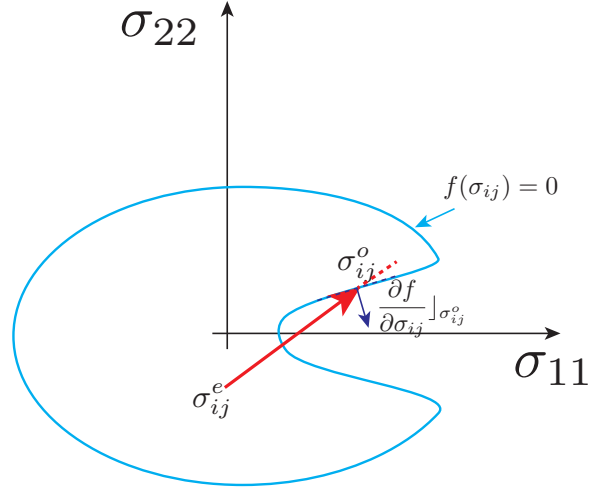


Figure 11.2: Example of a concave yield surface where one can have $(\sigma_{ij}^o - \sigma_{ij}^e) \frac{\partial f}{\partial \sigma_{ij}} \big|_{\sigma_{ij}^o} < 0$. Experimentally, (and physically), we observe that the yield surface are always convex. This can be intuitively understood.

Substracting the above expressions, we obtain

$$(1 - \alpha^a) \int_{\Gamma_t} t_i^c \dot{u}_i dS = \int_{\Omega} (\sigma_{ij} - \sigma_{ij}^a) \lambda \frac{\partial f}{\partial \sigma_{ij}} dV > 0$$

such that

$$\alpha^a < 1$$

because the external power is positive $t_i^c \dot{u}_i > 0$ (can also be used as a scaling factor). **The value α^a is thus a lower bound of the true collapse load** because

$$(\sigma_{ij} - \sigma_{ij}^a) \dot{\lambda} \frac{\partial f}{\partial \sigma_{ij}} > 0$$

This last inequality is related to the convexity of the yield surface (and the fact that σ_{ij}^a satisfy the yield criteria) - see Figures 11.1-11.2.

11.2.2 Upper bound

Let's take a kinematically admissible velocity field $\dot{u}_i^a$. For such a displacement field, corresponds a stress field σ_{ij}^b satisfying the yield criteria (but not necessarily satisfying equilibrium) such that $\dot{\epsilon}_{ij}^a = \lambda^b \frac{\partial f}{\partial \sigma_{ij}^b}$ where λ^b is the corresponding plastic multiplier. One can define the load multiplier

$$\alpha_{ab} \int_{\Gamma_t} t_i^c \dot{u}_i^a dS = \int_{\Omega} (\sigma_{ij}^b) \lambda^b \frac{\partial f}{\partial \sigma_{ij}^b} dV - \int_{\Omega} f_i \dot{u}_i^a dV$$

For the solution stress σ_{ij} (at plastic collapse) - which coincides with $\alpha = 1$ (due to our definition of α), the PPV for the test field $\dot{u}_i^a$ is:

$$1 \times \int_{\Gamma_t} t_i^c \dot{u}_i^a dS = \int_{\Omega} (\sigma_{ij}) \lambda^b \frac{\partial f}{\partial \sigma_{ij}^b} dV - \int_{\Omega} f_i \dot{u}_i^a dV$$

such that

$$(\alpha_{ab} - 1) \int_{\Gamma_t} t_i^c \dot{u}_i^a dS = \int_{\Omega} \lambda^b (\sigma_{ij}^b - \sigma_{ij}) \frac{\partial f}{\partial \sigma_{ij}^b} dV > 0$$

i.e. finally we thus have

$$\alpha_{ab} > 1$$

due to the property of convexity of the yield function (and the fact that we have used an associated plastic flow).

The displacement field $\dot{u}_i^a$ corresponding to the load multiplier α_{ab} thus gives an upper bound $\alpha_{ab} t_i^c$ of the true collapse load t_i^c .

11.3 Numerical limit analysis

The problem of numerical analysis can be formulated as an optimization problem with constraints: both equalities and inequalities constraints. Algorithms from numerical optimization can then be used - notably interior point methods which can tackle large amount of constraints.

11.3.1 Lower bound problem

The requirements of a lower-bound problem are

1. Equilibrium (everywhere in the domain)

$$\sigma_{ij,j} + f_i = 0$$

2. Tractions boundary conditions - with a multiplier term

$$\sigma_{ij} n_j = \alpha t_i$$

3. Yield condition (everywhere in the domain)

$$f(\sigma_{ij}) \leq 0$$

One can devise a lower-bound element where the component of the stress tensor are unknowns at the nodes, i.e. in matrix-form one can write

$$\sigma(\mathbf{x}) = \mathbf{N}_\sigma(\mathbf{x})\sigma$$

and the above continuous requirement translate after finite element discretization into the following discrete problem

$$\begin{aligned} & \text{Maximize } \alpha \\ & \text{subject to } \mathbf{B}_\nabla^T \sigma + \mathbf{f} = 0 \\ & \quad \mathbf{n}^T \sigma = \alpha \mathbf{t} \\ & \quad \mathbf{F}^T \sigma - \mathbf{k} + \mathbf{s} = \mathbf{0}, \quad \mathbf{s} \geq \mathbf{0} \end{aligned}$$

where $\mathbf{B}_\nabla^T$ provides the discretized divergence operator, and we have “linearized” the yield function and introduced a slave variable s . Note that the strain / displacement are given as

$$\nabla \dot{\mathbf{u}} = \mathbf{F}^T \lambda \quad \lambda \geq 0$$

and the KKT conditions are (point-wise)

$$s\lambda = 0 \quad \lambda \geq 0 \quad s \geq 0$$

Such a constrained optimization problem is solved via an interior point method in Optum.

In OptumG2 (2D), the lower bound triangular element has 9 stress unknowns and a constant displacement vector (2 unknowns). Triangular patch of zero thickness are introduced between all triangles to better capture stress discontinuity associated with shear banding. The yield function inequality is enforced at every nodes. Figure 11.3 sketch 2 neighbouring element with 2 zero-thickness patch. In total, there are $2 \times (3 \times 3) = 18$ stress unknowns + $2 \times 2 = 4$ for displacement (for the 2 elements - none for the patch) - 22 in total. There are $2 \times 4 = 8$ equilibrium constraints, 6 yield function inequality constraints, and 8 tractions continuity constraints (4 outside edges) - 22 equations in total.

11.3.2 Upper bound problem

The requirement for the upper bound rigid plastic problem can be written as

1. Principle of virtual work

$$\mathcal{P}_{ext}(\dot{u}_i, \alpha) = \mathcal{P}_{int}(\dot{u}_i)$$

2. Scaling of displacement at collapse via the external work (on all surfaces)

$$\int_{\Gamma} t_i^g \dot{u}_i \, dS = 1$$

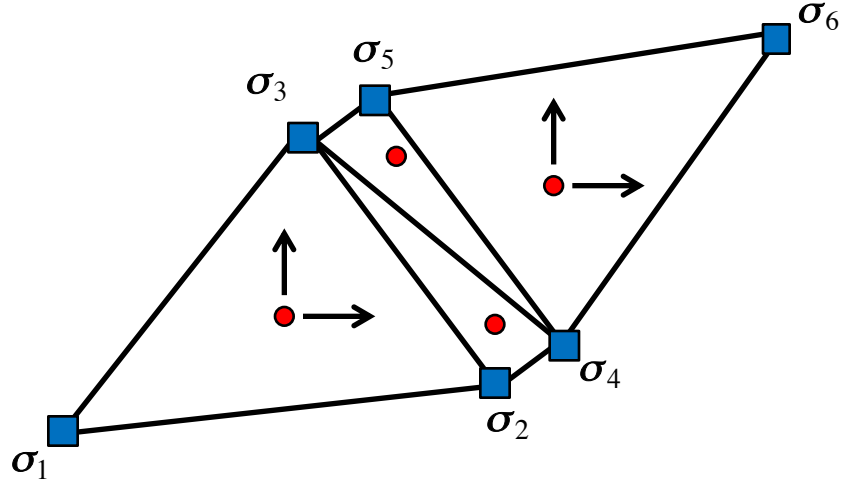


Figure 11.3: Lower bound stress element (Taken from Optum G2 documentation).

3. Yield function / Plastic criterium everywhere

$$f(\sigma_{ij}) \leq 0$$

4. Flow rule and complementary conditions

$$\dot{\epsilon}_{ij} = \lambda \frac{\partial f}{\partial \sigma_{ij}} \quad \lambda \geq 0 \quad f \lambda = 0$$

Note a slack variable s can be introduced such that the yield function inequality is replaced by

$$f(\sigma_{ij}) + s = 0 \quad s \geq 0$$

The above upper bound problem can be formally re-written as

$$\begin{aligned} & \text{Minimize } \mathcal{P}_{int}(\dot{u}_i) - \mathcal{P}_{ext}(\dot{u}_i, \alpha) = \\ & \text{subject to } \dot{\epsilon}_{ij} = \lambda \frac{\partial f}{\partial \sigma_{ij}} \\ & \lambda \geq 0 \\ & \int_{\Gamma} t_i^g \dot{u}_i \, dS = 1 \end{aligned}$$

where the internal dissipation (internal power) $\mathcal{P}_{int}(\dot{u}_i)$ is function of the type of yield criteria: it is either zero (when the yield criteria is strictly negative) or positive. It requires to obtain

an explicit expression of the internal dissipation function which depends on the type of yield criteria used.

Upper bound element have displacements at the nodes as primary unknowns. For 6-nodes Gauss, the displacement is quadratic, and stresses linear. Like for the lower bound element, patch of zero thickness can be introduced between elements to reproduce discontinuity of displacements (but the displacement are always continuous).

11.3.3 Mixed formulation

Here one solve for stresses and displacements. This is what is done in Optum - and depending on the choice of the type of element (lower-bound / upper bound). Let's re-write the problem in displacement / stresses. The variational form of the rigid plastic problem is (note the discontinuity are modelled via zero thickness element so their contributions in the internal power is left into a single integral for simplicity):

$$\min_{u_i} \max_{\alpha, \sigma_{ij}} \alpha + \int_{\Omega} \sigma_{ij} \dot{\epsilon}_{ij}(\dot{u}_i) dV - \int_{\Omega} f_i \dot{u}_i dV - \alpha \int_{\Gamma_{t_i}} t_i^g \dot{u}_i dS$$

with(at all points) $f(\sigma_{ij}) + s = 0, \quad s \geq 0$

Note that

$$\dot{\epsilon}_{ij}(\dot{u}_i) = \lambda \frac{\partial f}{\partial \sigma_{ij}} \quad \lambda \geq 0$$

Let's build the following Lagrangian

$$\mathcal{L} = \alpha + \int_{\Omega} \sigma_{ij} \dot{\epsilon}_{ij}(\dot{u}_i) dV - \int_{\Omega} f_i \dot{u}_i dV - \alpha \int_{\Gamma_{t_i}} t_i^g \dot{u}_i dS - \int_{\Omega} \lambda (f(\sigma_{ij}) + s) dV$$

with $s \geq 0$ and the complementary conditions $s\lambda = 0$ and $\lambda \geq 0$.

The optimality conditions (zero gradient at the optimal point) are

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \alpha} &= 1 - \int_{\Gamma_{t_i}} t_i^g \dot{u}_i dS = 0 \\ \frac{\partial \mathcal{L}}{\partial \dot{u}_i} &= \int_{\Omega} (\sigma_{ij,j} + f_i) d\Omega + \int_{\Gamma_{t_i}} (\sigma_{ij} n_j - \alpha t_i^g) dS = 0 \Rightarrow \sigma_{ij,j} + f_i = 0, \\ \frac{\partial \mathcal{L}}{\partial \sigma_{ij}} &= \int_{\Omega} (\dot{\epsilon}_{ij}(\dot{u}_i) - \lambda \frac{\partial f}{\partial \sigma_{ij}}) dV = 0 \Rightarrow \dot{\epsilon}_{ij}(\dot{u}_i) = \lambda \frac{\partial f}{\partial \sigma_{ij}} \\ \frac{\partial \mathcal{L}}{\partial \lambda} &= \int_{\Omega} (f(\sigma_{ij}) + s) dV = 0 \Rightarrow (f(\sigma_{ij}) + s) = 0 \end{aligned}$$

and the complementary conditions related to the inequality constraints are:

$$\begin{aligned}s &\geq 0 \\ \lambda &\geq 0 \\ s\lambda &= 0\end{aligned}$$

everywhere in the medium.

We see that from the optimality conditions, we recover: 1) the balance of momentum, 2) the tractions BC, 3) the scaling of the external work, 4) the plastic flow rule, 5) the inequality related to the yield function and 6) the complementary conditions.

Lower Bound Minimizing the Lagrangian with respect to displacement reduce the problem to the lower-bound case. The solution (for stress and displacement) can thus be obtained using the lower bound element previously described.

Upper bound The upper bound problem is actually better solved using the mixed formulation (solving for displacement and stress), as the internal dissipation function entering the principle of virtual work does not need to be explicitized (this is a nice practical advantage). In OptumG2 two upper bound elements are available: 1) a linear displacement / constant stress triangle, 2) a quadratic displacement / linear stress triangle. In both case, patch of zero thickness are added between all elements allowing to capture displacement discontinuity.

Chapter 12

Two-phases flow in porous media

This set of notes introduces the basic of the flow of 2 immiscibles fluid in porous media: typically a wetting and non-wetting phase. Here for clarity, indices 1, 2 will refer to the fluid phases. As a result, I will use bold faces for vector and greek letters (∇ etc.) for spatial differential operators.

12.1 Formulation

12.1.1 Saturation and relative permeability

The pore-space of the porous media of porosity ϕ is filled with 2 compressible immiscible fluids (phase 1 and 2) at the level of the RVE. We denote S_1 the saturation of the pore space by fluid 1, and the following closure relation holds:

$$S_1 + S_2 = 1$$

Typically one phase is wetting (e.g. 1) and the other non-wetting (e.g. 2). A capillary pressure p_c may develops such that

$$p_c = p_2 - p_1.$$

A relation between capillary pressure and saturation exist. This form may be intrcated, depend on residual saturation etc. Here, we write it schematically as

$$p_c(S_1) = p_2 - p_1$$

i.e. as function of the wetting phase saturation. Similarly, it is customary to introduce relative permeability for these 2 phases - non-linear function of S_1 :

$$\begin{aligned} k_{r1}(S_1) \\ k_{r2}(S_1) \end{aligned}$$

with $k_{r1}(S_1 = 1) = 1$, $k_{r1}(S_1 = 0) = 0$ and similarly $k_{r2}(S_1 = 0) = 1$.

12.1.1.1 Mass balances

The mass balance of each phase in the RVE reads

$$\frac{\partial \phi \rho_\alpha S_\alpha}{\partial t} + \nabla \cdot \rho_\alpha \mathbf{q}_\alpha = \rho_\alpha \gamma_\alpha \quad \alpha = 1, 2$$

where γ_α is a volumetric sink/source term for phase α .

12.1.1.2 Phases fluxes and total fluxes

Darcy;s law now reads

$$\mathbf{q}_\alpha = -\frac{k_{r\alpha}(S_\alpha)}{\mu_\alpha} \mathbb{K} (\nabla p_\alpha - \rho_\alpha \mathbf{g}) \quad \alpha = 1, 2$$

with $\mathbb{K}$ the porous media intrinsic permeability. It is usual to define the phase mobility as

$$\lambda_\alpha(S_\alpha) = \frac{k_{r\alpha}(S_\alpha)}{\mu_\alpha}$$

and the total mobility as

$$\lambda(S_1, p_2) = \sum_\alpha \rho_\alpha \lambda_\alpha = \rho_1 \lambda_1 + \rho_2 \lambda_2$$

the fractional flow functions

$$f_\alpha = \frac{\rho_\alpha \lambda_\alpha}{\lambda}$$

(note that $f_1 + f_2 = 1$) and the following non-linear functions:

- and 'average' density (which will appear in the buoyant term)

$$\rho = \frac{\lambda_1 \rho_1^2 + \lambda_2 \rho_2^2}{\lambda}$$

- and

$$\chi = \frac{\rho_1 \rho_2 \lambda_1 \lambda_2}{\lambda}$$

$$b = (\rho_1 - \rho_2) \chi \quad a = \chi \times \frac{dP_c}{dS_1}$$

The introducing of these functions becomes clear when one re-write the system of equations using S_1 the saturation of phase 1 (wetting phase) and p_2 the pressure of phase 2 as the primary unknowns. Note we have

$$S_2 = 1 - S_1 \quad p_1 = p_2 - p_c(S_1)$$

Notably, let's write the total mass flux $\mathbf{M}_t$

$$\mathbf{M}_t = \rho_1 \mathbf{q}_1 + \rho_2 \mathbf{q}_2$$

as function of S_1 and p_2 :

$$\begin{aligned} \mathbf{M}_t &= \rho_1 \mathbf{q}_1 + \rho_2 \mathbf{q}_2 \\ &= -\rho_1 \lambda_1(S_1) \mathbb{K} (\nabla p_1 - \rho_1 \mathbf{g}) - \rho_2 \lambda_2(S_1) \mathbb{K} (\nabla p_2 - \rho_2 \mathbf{g}) \\ &= -\rho_1 \lambda_1(S_1) \mathbb{K} (\nabla p_2 - \nabla p_c - \rho_1 \mathbf{g}) - \rho_2 \lambda_2(S_1) \mathbb{K} (\nabla p_2 - \rho_2 \mathbf{g}) \\ &= -\lambda(S_1, p_2) \mathbb{K} (\nabla p_2 - \rho \mathbf{g}) + \rho_1 \lambda_1(S_1) \mathbb{K} \nabla p_c \\ &= -\lambda(S_1, p_2) \mathbb{K} (\nabla p_2 - f_1(S_1, p_2) \nabla p_c - \rho \mathbf{g}) \end{aligned}$$

finally

$$\mathbf{M}_t = -\lambda(S_1, p_2) \mathbb{K} \left(\nabla p_2 - f_1(S_1, p_2) \frac{dp_c}{dS_1} \nabla S_1 - \rho(S_1, p_2) \mathbf{g} \right)$$

and we can express the phases mass flux as

$$\begin{aligned} \rho_1 \mathbf{q}_1 &= \mathbf{M}_t - \rho_2 \mathbf{q}_2 = -\lambda_1 \rho_1 \mathbb{K} \left(\nabla p_2 - \frac{dp_c}{dS_1} \nabla S_1 - \rho_1 \mathbf{g} \right) \\ &= f_1(S_1, p_2) \mathbf{M}_t + \underbrace{\chi(S_1, p_2) \frac{dp_c}{dS_1}}_a \mathbb{K} \nabla S_1 + b(S_1, p_2) \mathbb{K} \mathbf{g} \end{aligned}$$

(do the derivation as exercice) and

$$\rho_2 \mathbf{q}_2 = f_2(S_1, p_2) \mathbf{M}_t - \chi \frac{dp_c}{dS_1} \mathbb{K} \nabla S_1 - b(S_1, p_2) \mathbb{K} \mathbf{g}$$

12.1.1.3 System in term of S_1 and p_2

Summing up the mass conservation for the two phases, we obtain

$$\frac{\partial \phi (\rho_1 S_1 + \rho_2 (1 - S_1))}{\partial t} + \nabla \cdot \mathbf{M}_t = \rho_1 \gamma_1 + \rho_2 \gamma_2 \quad (12.1)$$

and the total flux $\mathbf{M}_t$ has been expressed previously as function of p_2 and S_1 . Werewrite it here to highlight the terms

$$\mathbf{M}_t = -\lambda(S_1, p_2) \mathbb{K} \left(\nabla p_2 - f_1(S_1, p_2) \frac{dp_c}{dS_1} \nabla S_1 - \rho(S_1, p_2) \mathbf{g} \right)$$

such that we have

$$\frac{\partial \phi (\rho_1 S_1 + \rho_2 (1 - S_1))}{\partial t} + \nabla \cdot \left(-\lambda(S_1, p_2) \mathbb{K} \left(\nabla p_2 - f_1(S_1, p_2) \frac{dp_c}{dS_1} \nabla S_1 - \rho(S_1, p_2) \mathbf{g} \right) \right) = \rho_1 \gamma_1 + \rho_2 \gamma_2$$

We then keep the mass conservation for the first phase and write it via the expression of $\rho_1 \mathbf{q}_1$

$$\frac{\partial \phi \rho_1 S_1}{\partial t} + \nabla \cdot \left(f_1(S_1, p_2) \mathbf{M}_t + \chi(S_1, p_2) \frac{dp_c}{dS_1} \mathbb{K} \nabla S_1 + b(S_1, p_2) \mathbb{K} \mathbf{g} \right) = \rho_1 \gamma_1 \quad (12.2)$$

12.1.1.4 Initial and boundary conditions

Initial values for saturation S_1 and pressure p_2 everywhere in the domain Ω . then

$$S_1(\mathbf{x}, t) = S_1^g(\mathbf{x}, t) \quad \text{on } \Gamma_d^{S_1} \quad p_2(\mathbf{x}, t) = p_2^g(\mathbf{x}, t) \quad \text{on } \Gamma_d^{p_2}$$

and

$$\rho_1 \mathbf{q}_1(\mathbf{x}, t) \mathbf{n} = Q_1(t) \quad \text{on } \Gamma_n^{S_1} \quad \rho_2 \mathbf{q}_2(\mathbf{x}, t) \mathbf{n} = Q_1(t) \quad \text{on } \Gamma_n^{p_2}$$

with $\partial\Omega = \Gamma_d^{S_1} \cup \Gamma_n^{S_1} = \Gamma_d^{p_2} \cup \Gamma_n^{p_2}$ and the usual $\Gamma_d^{S_1} \cap \Gamma_n^{S_1} = \emptyset = \Gamma_d^{p_2} \cap \Gamma_n^{p_2}$.

12.2 FEM

It is typical to use linear shape function for p_2 and S_1 such that we can also obtain S_2 , p_c and p_1 at the nodes. The resulting coupled PDE problem is highly non-linear due to the relative permeability and capillary pressure functions. Some strategies involved: doing a staggered coupling, solving implicit for pressure (implicit in time) then evolving the saturation in an explicit manner). Other due a full implicit coupling. The non-linearity can be either solved via a Newton-Raphson scheme or fixed point (the latter sometimes require relaxation or Anderson acceleration to converge).

Let's do a bit further massaging of the 2 conservation equations.

$$\frac{\partial \phi (\rho_1 S_1 + \rho_2 (1 - S_1))}{\partial t} + \nabla \cdot \left(-\lambda(S_1, p_2) \mathbb{K} \left(\nabla p_2 - f_1(S_1, p_2) \frac{dp_c}{dS_1} \nabla S_1 - \rho(S_1, p_2) \mathbf{g} \right) \right) = \rho_1 \gamma_1 + \rho_2 \gamma_2$$

$$\frac{\partial \phi \rho_1 S_1}{\partial t} + \nabla \cdot \left(f_1(S_1, p_2) \mathbf{M}_t + \chi(S_1, p_2) \frac{dp_c}{dS_1} \mathbb{K} \nabla S_1 + b(S_1, p_2) \mathbb{K} \mathbf{g} \right) = \rho_1 \gamma_1 \quad (12.3)$$

we can rewrite the second one as

$$\frac{\partial \phi \rho_1 S_1}{\partial t} + \nabla \cdot \left(-\lambda_1(S_1, p_2) \rho_1(S_1, p_2) \mathbb{K} \left(\nabla p_2 - \frac{dp_c}{dS_1} \nabla S_1 - \rho_1(S_1, p_2) \mathbf{g} \right) \right) = \rho_1 \gamma_1$$

we thus see a sort of symmetry between the spatial operator of the resulting system p_2, S_1 .

For the time derivative we can rewrite

$$\frac{\partial \phi \rho_1 S_1}{\partial t} = \left(\rho_1(S_1, p_2) S_1 \frac{d\phi}{dp_2} + \phi S_1 \frac{d\rho_1}{dp_2} \right) \frac{\partial p_2}{\partial t} + \rho_1 \phi \frac{\partial S_1}{\partial t}$$

and similarly

$$\begin{aligned} \frac{\partial \phi (\rho_1 S_1 + \rho_2 (1 - S_1))}{\partial t} &= \left\{ (\rho_1(S_1, p_2) S_1 + \rho_2(S_1, p_2) (1 - S_1)) \frac{d\phi}{dp_2} + \phi S_1 \frac{d\rho_1}{dp_2} + \phi (1 - S_1) \frac{d\rho_2}{dp_2} \right\} \frac{\partial p_2}{\partial t} \\ &\quad + \phi (\rho_1 - \rho_2) \frac{\partial S_1}{\partial t} \end{aligned}$$

Note that we have the pore-compressibility β_ϕ

$$\beta_\phi = \frac{d\phi}{dp_2}$$

and

$$\frac{d\rho_1}{dp_2} = \frac{d\rho_1}{dp_1} \times \frac{dp_1}{dp_2} = \frac{d\rho_1}{dp_1} \times \left(1 - \frac{dp_c}{dp_2} \right)$$

for a liquid phases - we can linearize and introduce liquid compressibilities β_α

$$\beta_\alpha = \frac{d\rho_\alpha}{dp_\alpha}$$

After FEM discretization, the final ODE will look like

$$\begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{21} \\ \mathbf{M}_{21} & \mathbf{M}_{22} \end{bmatrix} \cdot \frac{\partial}{\partial t} \begin{bmatrix} \mathbf{S}_1 \\ \mathbf{p}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{K}_{11} & \mathbf{K}_{12} \\ \mathbf{K}_{21} & \mathbf{K}_{22} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{S}_1 \\ \mathbf{p}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \end{bmatrix}$$

where the mass matrices and the conductivities matrices are non-linear function of S_1 and p_2 (and the force terms have buoyant effects non-linear as well). Note that the full system is not symmetric (e.g. $K_{12} \neq K_{21}$, same for mass).

Bibliography

- Carslaw, H. & Jaeger, J. (1959), *Conduction of heat in solids*, Oxford Univ Press.
- Cornet, F. H. (2015), *Elements of Crustal Geomechanics*, Cambridge University Press.
- Coussy, O. (2004), *Poromechanics*, Wiley, New-York.
- Coussy, O. (2007), ‘Revisiting the constitutive equations of unsaturated porous solids using a Lagrangian saturation concept’, *Int. J. Numer. Anal. Meth. Geomech.* **31**(15), 1675–1694.
- Coussy, O. (2010), *Mechanics and Physics of Porous Solids*, Wiley & Sons.
- Crank, J. (1979), *The mathematics of diffusion*, Oxford university press.
- Davis, R. O. & Selvadurai, A. P. (2005), *Plasticity and geomechanics*, Cambridge university press.
- de Souza Neto, E. A., Peric, D. & Owen, D. R. (2011), *Computational methods for plasticity: theory and applications*, John Wiley & Sons.
- Detournay, E. & Cheng, A. (1993), Fundamentals of poroelasticity, in ‘Comprehensive Rock Engineering’, Pergamon Press.
- Guermond, J.-L. & Pasquetti, R. (2013), ‘A correction technique for the dispersive effects of mass lumping for transport problems’, *Computer Methods in Applied Mechanics and Engineering* **253**, 186–198.
- Harr, M. E. (1962), *Groundwater and seepage*, Dover.
- Hughes, T.J.R. (1987), *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*, Dover.
- Lubliner, J. (1990), *Plasticity Theory*, MacMillan Publishing Company.
- Muir Wood, D. (1990), *Soil Behaviour and Critical State Soil Mechanics*, Cambridge Univ Press.

- Rice (2010), ‘Solid mechanics’, available at http://esag.harvard.edu/rice/e0_Solid_Mechanics_94_10.pdf, accessed Aug. 23, 2019.
- Rice, J. & Cleary, M. (1976), ‘Some basic stress diffusion solutions for fluid-saturated elastic porous media with compressible constituents’, *Rev. Geoph.* **14**(2), 227–241.
- Simo, J. C. & Hughes, T. J. R. (1998), *Computational Inelasticity*, Springer.
- Wang, H. (2000), *Theory of linear poroelasticity with applications to geomechanics and hydrogeology*, Princeton University Press.
- Wriggers, P. (2006), *Computational contact mechanics*, Springer Science & Business Media.
- Zienkiewicz, O. & Taylor, R. (2005), *The Finite Element Method for Solid and Structural Mechanics*, Fifth Edition edn, Elsevier.

Appendix A

Displacement, Strain, Stress - some recaps

Review this chapter at your own pace to refresh your continuum mechanics - it will NOT be covered in class.

Convention of summation on repeated/dummy indices is used otherwise stated.

A.1 Transformation, displacement and strain

Let denote $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ an orthonormal cartesian basis, and consider a solid material. Let's adopt a reference configuration (undeformed / initial) in which the solid occupy the volume V_o . In this configuration, the particles of the solid can be located by their position vector $\mathbf{X}$ (coordinates X_i), and $d\mathbf{X}$ is an infinitesimal vector which links the solid particle located at $\mathbf{X}$ with the nearby particle located at $\mathbf{X} + d\mathbf{X}$. Now, let's assume the solid undergoes a deformation from the initial to a current (deformed) configuration at time t . In this new/current configuration, the particle whose initial position vector was $\mathbf{X}$ can be located by its current position vector $\mathbf{x}$ which is a function of $\mathbf{X}$ and time

$$\mathbf{x} = \mathbf{x}(\mathbf{X}, t) = \mathbf{X} + \mathbf{u}(\mathbf{X}, t) \quad (\text{A.1})$$

where $\mathbf{u}$ is the displacement vector from the initial to the current deformation. After deformation the infinitesimal vector $d\mathbf{X}$ linking 2 nearby particles in the original configuration is now $d\mathbf{x}$, and we simply have by differentiation of the previous equation:

$$\begin{aligned} dx_i &= \frac{\partial x_i}{\partial X_j} dX_j = \left(\delta_{ij} + \frac{\partial u_i}{\partial X_j} \right) dX_j \\ d\mathbf{x} &= dx_i \mathbf{e}_i \end{aligned}$$

$F_{ij} = \frac{\partial x_i}{\partial X_j}$ is typically denoted the deformation gradient.

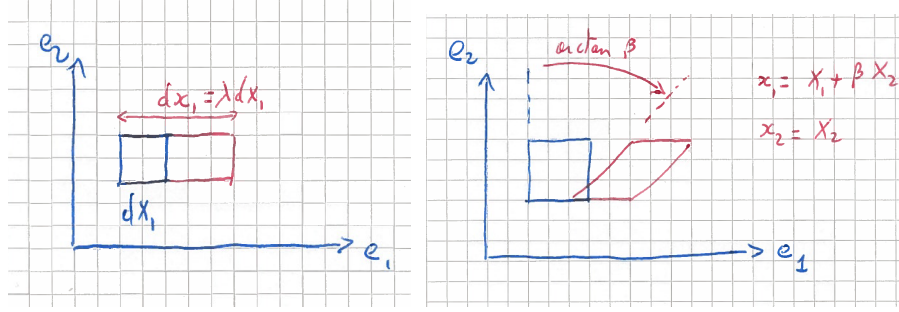


Figure A.1: Uniaxial stretch (left), simple shear (right) transformations

Let's take two simple examples, starting from an infinitesimal element dX_i , originally aligned with the orthonormal basis.

- Uniaxial transformation: $x_1 = a + \lambda X_1$, $x_2 = b + X_2$ (and $x_3 = c + X_3$), then we change in the length of the infinitesimal element in direction 1:

$$\frac{dx_1 - dX_1}{dX_1} = \frac{\partial u_1}{\partial X_1} = \lambda - 1$$

- Simple shear : $x_1 = X_1 + \beta X_2$, $x_2 = X_2$ (and $x_3 = X_3$),

$$\begin{aligned} dx_1 &= dX_1 + \beta dX_2 \\ dx_2 &= dX_2 \end{aligned}$$

Note that there is no change of volume under such a transformation. (area of parallelogram defined by two vectors $\mathbf{v}, \mathbf{w}$ with angle θ between them is $A = \|\mathbf{v}\| \|\mathbf{w}\| \sin \theta \dots$). Also, the angle between dx_1 and dx_2 is simply $\pi/2 - \arctan \beta \approx \pi/2 - \beta$ for $\beta \ll 1$. The change in angle (between direction 1 and 2) is β (to the first order).

The same apply to a small material volume dV_o in the initial configuration to the current one dV . The volume of a parallelepiped dV is given by $dV = d\mathbf{x}_3 \cdot (d\mathbf{x}_1 \times d\mathbf{x}_2) \dots$ show that

$$dV = \left| \frac{\partial x_i}{\partial X_j} \right| dV_o$$

Deformation induces change in the length of the material vectors and the angle they form between each others.

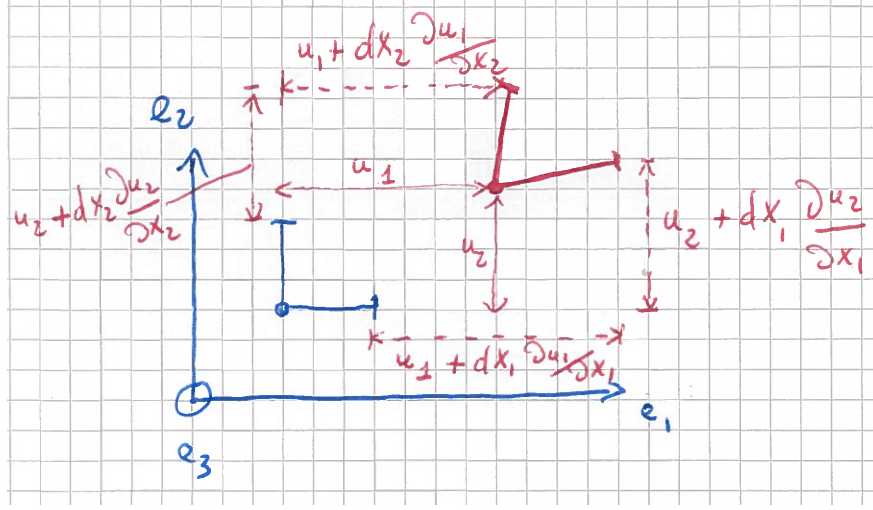


Figure A.2: Transformation and small strain

A.1.1 Small Strain

$$\left\| \frac{\partial u_i}{\partial X_j} \right\| \ll 1 \text{ for all } i \text{ and } j !$$

To grasp its meaning, let's restrict to 2D. Let's look at the deformation of an initially square parallelogram and two infinitesimal line segments in the direction 1 and 2 in the original configuration (blue in Fig. A.1). Displacements are labeled such that u_i denotes the displacement of the point initially at (X_1, X_2) , the displacement of the point originally at $X_1 + dX_1$ is $u_i + dX_1 \partial u_i / \partial X_1$. If the material fibers shown in the 1 and 2 direction did not rotate, then the strain defined as their fractional change in length (e.g. $(dx_1 - dX_1)/dX_1$) would be exactly $\epsilon_{11} = \partial u_1 / \partial X_1$ and $\epsilon_{22} = \partial u_2 / \partial X_2$. These expressions remain correct to the first order in $\partial u / \partial X$ if the fibers do rotate. The angle reduction between two originally orthogonal fibers in the direction of increasing X_1 and X_2 is $\partial u_1 / \partial X_2 + \partial u_2 / \partial X_1$ (to the first order accuracy of $\partial u / \partial X$), and this define a shear strain γ_{12} . For the shear strain ϵ_{12} , we use half of γ_{12} however. Thus considering all shear and extensional strain in all directions at a material point, we obtain the following strain tensor:

$$\epsilon_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial X_j} + \frac{\partial u_j}{\partial X_i} \right)$$

Under the small strain assumptions, we may write $\partial / \partial X_i \sim \partial / \partial x_i$.

Green-Lagrange Alternatively (mathematically), we can define the Green-Lagrange strain tensor as relating the change in the dot product of two infinitesimal material vectors $d\mathbf{X}$ and $d\mathbf{Y}$ which transform into $d\mathbf{x}$ and $d\mathbf{y}$ during the deformation

$$d\mathbf{x} \cdot d\mathbf{y} - d\mathbf{X} \cdot d\mathbf{Y} = d\mathbf{X} \cdot (\mathbf{F}^T \mathbf{F} - \mathbf{I}) \cdot d\mathbf{Y} = 2d\mathbf{X} \cdot \mathbf{D} \cdot d\mathbf{Y} \quad (\text{A.2})$$

where $\mathbf{F}$ denotes the deformation gradient tensor $F_{ij} = \frac{\partial x_i}{\partial X_j}$. From the definition of the displacement vector, we obtain

$$D_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial X_j} + \frac{\partial u_j}{\partial X_i} + \frac{\partial u_k}{\partial X_j} \frac{\partial u_k}{\partial X_i} \right)$$

and we immediatly recognized that the small strain tensor ϵ_{ij} is the first order linearization of D_{ij} accurate as long as $\left\| \frac{\partial u_i}{\partial X_j} \right\| \ll 1$ for all i and j . Now, taking $d\mathbf{y} = d\mathbf{x}$ in eq.(A.2), we obtain after dividing by $\|d\mathbf{X}\|^2$:

$$\frac{d\mathbf{x} - d\mathbf{X}}{\|d\mathbf{X}\|} \frac{d\mathbf{x} + d\mathbf{X}}{\|d\mathbf{X}\|} = 2 \frac{d\mathbf{X} \cdot \mathbf{D} \cdot d\mathbf{X}}{\|d\mathbf{X}\|^2}$$

under the small strain linear assumption, we have $\mathbf{D} = \epsilon$ and $\frac{d\mathbf{x} + d\mathbf{X}}{\|d\mathbf{X}\|} \sim 2$, hence

$$\frac{d\mathbf{x} - d\mathbf{X}}{\|d\mathbf{X}\|} = N_i \epsilon_{ij} N_j$$

where $N_i = dX_i / \|d\mathbf{X}\|$ is the unit vector in the direction $d\mathbf{X}$. For example if $\mathbf{N} = \mathbf{e}_1$, we see that ϵ_{11} is the fractional change in length in the $\mathbf{e}_1$ direction.

A.1.2 Material derivatives - Lagrangian vs Eulerian

The velocity $\mathbf{v}$ is the time-derivative of the material position vector in the current configuration $\mathbf{x}$:

$$\mathbf{v} = \frac{\partial \mathbf{x}}{\partial t} = \frac{\partial \mathbf{u}}{\partial t}$$

So-far we have only taken the Lagrangian point of view, which always refer to the change with respect to an initial configuration (i.e. like following an individual particle). On the contrary, Eulerian description focus on a specific geometrical position and specify e.g. the flow field with respect to that position.

Let G be a quantity function of the current position and time, i.e. $G = G(\mathbf{x}(\mathbf{X}, t), t)$, the chain rules give the following material derivative:

$$\frac{dG}{dt} = \frac{\partial G}{\partial t} + \frac{\partial x_i}{\partial t} \frac{\partial G}{\partial x_i} = \frac{\partial G}{\partial t} + \mathbf{v} \cdot \nabla G$$

In the case of solid(and a porous solid), a Lagrangian description is usually preferred, while Eulerian description is used in fluid mechanics. Note also that in the flow in porous media litterature, an Eulerian description is sometimes used (implicitly). Note that for small strain, Eulerian and Lagrangian description can eventually be merged provided only spatial derivations are involved - which is the case for creeping flow (Stokes-flow) typically occuring in porous media. Note, however, that the concept of displacement is only relevant in a Lagrangian description.

A.1.3 Small strain & compatibility conditions

The small strain tensor is symmetric and has 6 independent components. We need additional relations to integrate it to obtain the displacement vector, i.e. the strain compatibility conditions.

The spin tensor (anti-symmetric part of the displacement gradient) is defined as

$$\omega_{ij} = \frac{1}{2} (u_{i,j} - u_{j,i})$$

and we see that

$$w_{ij,k} = \frac{1}{2} (u_{i,jk} - u_{j,ik}) = \frac{1}{2} (u_{i,kj} + u_{k,ij} - u_{j,ki} - u_{k,ji}) = \epsilon_{ki,j} - \epsilon_{jk,i}$$

We can deduce the compatibility conditions on ϵ_{ij} , from the fact that:

$$w_{ij,kl} = w_{ij,lk}$$

which gives

$$\epsilon_{ij,kl} + \epsilon_{kl,ij} - \epsilon_{ik,jl} - \epsilon_{jl,ik} = 0$$

A.2 Stress and equations of motion

Let denote $\mathbf{x}$ the position vector of a point in space. Let $\mathbf{v}(\mathbf{x}, t)$ be the current velocity and $\rho(\mathbf{x})$ the mass density of the material. The linear momentum of a material element dV is $\rho \mathbf{v} dV$ and its angular momentum relative to the coordinate of origin is $\mathbf{x} \times \rho \mathbf{v} dV$. Linear and angular momentum equilibrium express that the rate of change of linear momentum (resp. angular) are equal to the sum of forces (resp. torques) applied to the body. The latter being a combination of body forces density $\mathbf{f}$ and surface traction $\mathbf{T}$ (function of the surface normal $\mathbf{n}$):

$$\begin{aligned} \frac{d}{dt} \int_V \rho \mathbf{v} dV &= \int_V \mathbf{f} dV + \int_S \mathbf{T}(\mathbf{n}) dS \\ \frac{d}{dt} \int_V \mathbf{x} \times \rho \mathbf{v} dV &= \int_V \mathbf{x} \times \mathbf{f} dV + \int_S \mathbf{x} \times \mathbf{T}(\mathbf{n}) dS \end{aligned}$$

Restricting to quasi-static conditions, the balance of linear and angular momentum reduces to:

$$\begin{aligned} \int_V \mathbf{f} dV + \int_S \mathbf{T}(\mathbf{n}) dS &= 0 \\ \int_V \mathbf{x} \times \mathbf{f} dV + \int_S \mathbf{x} \times \mathbf{T}(\mathbf{n}) dS &= 0 \end{aligned}$$

Action-Reaction law Let consider an infinitely thin cylinder (of volume $\pi R^2 h$ where R is the cylinder radius and h its thickness) with S_{out} the outer cylindrical surface (size $2\pi R h$), S^+ and S^- the two -end surfaces oriented by the unit outward vector $\mathbf{n}$ and $-\mathbf{n}$. Applying the linear momentum, in the limit where the cylinder thickness tends to zero such that the outer cylindrical surface S_{out} and volume V vanishes, we obtain:

$$\int_{S^+} \mathbf{T}(\mathbf{n}) dS + \int_{S^-} \mathbf{T}(-\mathbf{n}) dS = 0$$

Since it holds whatever dS , we obtain the action-reaction law:

$$\mathbf{T}(-\mathbf{n}) = -\mathbf{T}(\mathbf{n})$$

Small tetrahedron Let consider an infinitesimal tetrahedron, whose faces S_j are parallel to the coordinate planes and oriented by $-\mathbf{e}_j$. These surfaces A_j are linked to the base surface S of unit normal $\mathbf{n}$,

$$S_j = S n_j$$

[Note $n_j = 1/\sqrt{3}$, $S = h^2\sqrt{3}/2$, $S_j = h^2/2$]. The balance of momentum, for an infinitesimally small tetrahedron reduces to

$$\frac{hS}{3} O(f) + S\mathbf{T}(\mathbf{n}) + S_j\mathbf{T}(-\mathbf{e}_j) = 0$$

using the action-reaction law, and letting $h \rightarrow 0$, we obtain

$$\mathbf{T}(\mathbf{n}) = \mathbf{T}(\mathbf{e}_j)n_j$$

Now, suppose that we consider an element of surface dS through a point $\mathbf{x}$ so that its outer normal points in the positive j direction ($\mathbf{n} = \mathbf{e}_j$). Let $\sigma_{1j}, \sigma_{2j}, \sigma_{3j}$ define the cartesian components of the stress vector $\mathbf{T}$ acting on this element of surface dS is, given by:

$$\mathbf{T}(\mathbf{e}_j) = \sigma_{1j}\mathbf{e}_1 + \sigma_{2j}\mathbf{e}_2 + \sigma_{3j}\mathbf{e}_3 = \sigma_{ij}\mathbf{e}_i$$

For a general element of surface dS of normal $\mathbf{n} = n_i\mathbf{e}_i$, we have $\mathbf{T}(\mathbf{n}) = n_j\sigma_{ij}\mathbf{e}_i$, where σ_{ij} is the stress tensor (dimension of Force / area). It is a linear operator relating the stress vector $\mathbf{T}(\mathbf{n})$ to the normal $\mathbf{n}$ of the element of surface.

Introducing in the linear momentum, we obtain:

$$\int_V f_i dV + \int_{\partial V} n_j \sigma_{ij} dS = 0$$

Applying's the divergence theorem, we obtain the local form of the balance of momentum (under quasi-static conditions)

$$\frac{\partial \sigma_{ij}}{\partial x_j} + f_i = 0$$

[Exercise: Show that the stress tensor is symmetric using the balance of angular of momentum].

Angular momentum quasi-static case

$$\int_V (x_i f_j - x_j f_i) dV + \int_S (x_i T_j - x_j T_i) dS = 0$$

but $T_i = \sigma_{ik} n_k$, the second term becomes:

$$\begin{aligned} \int_S (x_i T_j - x_j T_i) dS &= \int_S x_i \sigma_{jk} n_k - x_j \sigma_{ik} n_k dS = \int_V \frac{\partial}{\partial x_k} (x_i \sigma_{jk} - x_j \sigma_{ik}) dV = \\ &= \int_V (\sigma_{ji} - \sigma_{ij} + x_i \frac{\partial \sigma_{jk}}{\partial x_k} - x_j \frac{\partial \sigma_{ik}}{\partial x_k}) dV \end{aligned}$$

using the balance of linear momentum, we obtain the symmetry of the stress tensor:

$$\sigma_{ij} = \sigma_{ji}$$

A.2.1 Principal stresses - Mohr Circles etc.

A.2.1.1 Coordinates transform

Let consider another orthonormal system of coordinates $(\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3)$, let R_{ij} be the matrix containing the coordinates of $\mathbf{e}'_i$ in the original system $\mathbf{e}_i$, such that

$$\mathbf{e}'_i = R_{ij} \mathbf{e}_j$$

it defines the transformation between the two orthonormal frames, and $R_{ik} R_{kj}^T = \delta_{ij}$ ($R_{ik} R_{jk} = \delta_{ij}$); i.e. we also have. $\mathbf{e}_i = R_{ij}^T \mathbf{e}'_j$. Note that for a vector $\mathbf{v}$, its component in the new frame is $v'_i = R_{ij} v_j$.

For the stress vector, its definition in either of the 2 frames are

$$\mathbf{T} = n_j \sigma_{ij} \mathbf{e}_i = n'_k \sigma'_{kl} \mathbf{e}'_l$$

with the stress tensor in the frame $(\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3)$ therefore given by

$$\sigma'_{kl} = R_{kj} \sigma_{ij} R_{il}^T$$

.

A.2.1.2 Principal stresses & invariants

Due to the symmetry of the stress tensor, and the coordinates transform previously stated, at a given point, it exists three mutually independent direction forming an orthonormal frame such that the stress tensor expressed in that frame is diagonal. These are the principal stresses

direction, to which are associated the principal stress values which we can algebraically denote σ_1, σ_2 and σ_3 (in increasing order). They are solution of an eigenvalue problem i.e.

$$\sigma_{ij}p_j = \sigma p_i$$

the eigenvalues (three) are solution of the characteristic equation

$$\det(\sigma_{ij} - \sigma \delta_{ij}) = 0 = -\sigma^3 + I_1\sigma^2 + I_2\sigma + I_3 = 0$$

Stress invariants independent of the chosen coordinates system, to the number of 3.

$$\begin{aligned} I_1 &= \sigma_{ii} \\ I_2 &= \frac{1}{2}\sigma_{ij}\sigma_{ij} \\ I_3 &= \frac{1}{3}\sigma_{ij}\sigma_{jk}\sigma_{ki} \end{aligned}$$

A.2.1.3 Normal and shear stress to a plane

Let denote a plan of normal $\mathbf{n}$, with $\mathbf{s}$ a vector in the plane (such that $s_i n_i = 0$), the traction vector T can be decomposed in a normal and shear component in this plane. The normal component of the traction vector is denoted the normal stress σ_n :

$$T_i n_i = n_i \sigma_{ij} n_j \equiv \sigma_n$$

and the tangential stress in the direction $\mathbf{s}$ is

$$T_i s_i = s_i \sigma_{ij} n_j$$

Note that in a given plane, where 2 orthonormal vector define a local basis, they are of course two associated tangential stresses. Taking the norm of these tangential stresses, we obtain the so-called shear stress. It can be directly expressed as

$$\tau^2 = n_i \sigma_{ij} \sigma_{jk} n_k - (\sigma_{ij} n_i n_j)^2$$

A.2.1.4 Mohr Circles

Let's take an orthonormal basis made of the principal direction of the stress tensor (I, II, III) with principal values $\sigma_I > \sigma_{II} > \sigma_{III}$. For a given normal $\mathbf{n}$ to an infinitesimal surface, the traction vector reduce to

$$T_I = \sigma_I n_I \quad T_{II} = \sigma_{II} n_{II} \quad T_{III} = \sigma_{III} n_{III}$$

The normal stress

$$\sigma_n = \sigma_I n_I^2 + \sigma_{II} n_{II}^2 + \sigma_{III} n_{III}^2$$

and the shear stress:

$$\tau^2 + \sigma_n^2 = \sigma_I^2 n_I^2 + \sigma_{II}^2 n_{II}^2 + \sigma_{III}^2 n_{III}^2$$

Let's take as a particular example, the case where the normal to the segment lie in the plane (I, II) i.e. perpendicular to direction III, with an angle θ to the direction I: $n_I = \cos \theta$, $n_{II} = \sin \theta$, and $n_{III} = 0$. The normal stress is:

$$\begin{aligned}\sigma_n &= \frac{\sigma_I + \sigma_{II}}{2} + \frac{\sigma_I - \sigma_{II}}{2} \cos 2\theta \\ \tau &= -\frac{\sigma_I - \sigma_{II}}{2} \sin 2\theta\end{aligned}$$

[Note that θ lies in the range $[0, \pi/2]$, the shear stress is obviously null for $\theta = 0$ and $\pi/2$ and maximum for $\theta = \pi/4$.]. We can see that the traction vector in that case, is located in a circle (parametrized by the angle θ defining here the normal to the infinitesimal plane) in the σ_n, τ plane, with center $(\frac{\sigma_I + \sigma_{II}}{2}, 0)$ and radius $\frac{\sigma_{II} - \sigma_I}{2}$.

We can do similarly for the cases where the normal to the segment lies in the plane (I, III) and (I, II). We therefore obtain three circles, and we can show that the traction vector to any surface with normal $\mathbf{n}$ lies in between the two small Mohr circles defined by $\sigma_{II} - \sigma_I$, and $\sigma_{III} - \sigma_{II}$ and the largest Mohr circle defined by $\sigma_{III} - \sigma_I$.

A.2.1.5 Mean and deviatoric stress

It is sometimes worthwhile to decompose the stress tensor in a volumetric (mean) and a deviatoric part:

$$\sigma_{ij} = \frac{\sigma_{kk}}{3} \delta_{ij} + s_{ij}$$

By definition $s_{kk} = 0$. Note the invariants J_s of the deviatoric stress tensor s_{ij} are

$$J_1 = 0, \quad J_2 = I_2 - I_1^2/6 \quad J_3 = I_3 - 2I_1 J_2/3 - I_1^3/27$$

It is interesting to note that for an element of surface with normal $(\sqrt{3}/3, \sqrt{3}/3, \sqrt{3}/3)$ in the principal stress direction basis, the normal and shear stress are

$$\sigma_n = I_1/3 = \sigma_{kk}/3 \quad \tau = \sqrt{\frac{3}{2} J_2}$$

A.3 Exercises

1. Let's assume a Mohr-Coulomb failure envelope, i.e. $\tau \leq C - \sigma_n \tan \varphi$ representative of rocks. Rocks are weak in tension, and fail at higher shear stress when the compressive normal stress is higher (tensile stresses are positive in the convention used in this chapter). In a bi-axial test (i.e. 2 principal stresses σ_I and σ_{II}), what is the angle (with respect to the largest principal stresses σ_I) of the element of surface which fails first (i.e. touch the failure surface first)?

2. An irrotational displacement field is such that $\nabla \times \mathbf{u} = \mathbf{0}$. Provide two simple examples (in 2D).
3. Derive the strain compatibility conditions (compatibility conditions for integration of the strain tensor).
4. Prove that the stress tensor is symmetric (hints: start from the conservation of angular of momentum)

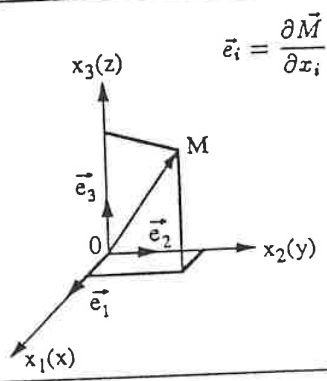
Appendix B

Coordinates systems

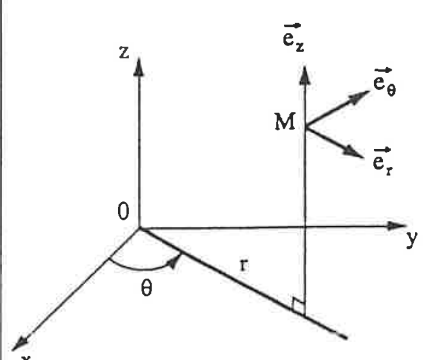
[in french ;), my 20 years old indispensable list of differential operators in the different coordinates systems]

Tableau de passage des coordonnées

- Coordonnées cartésiennes orthonormées,
- Coordonnées cylindriques,
- Coordonnées sphériques,

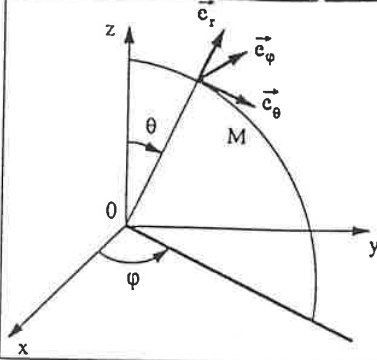
	Coordonnées cartésiennes orthonormées
Définitions	 $\vec{e}_i = \frac{\partial \vec{M}}{\partial x_i}$
coordonnées	$\vec{OM} = x_i \vec{e}_i$
champ de vecteurs $\vec{u}(M)$	$\vec{u}(M) = u_i(M) \vec{e}_i$
vecteurs unitaires de base	$d\vec{e}_i = 0$ $d\vec{M} = dx_i \vec{e}_i$
gradient d'un scalaire f $df = \overrightarrow{\text{grad}} f \cdot d\vec{M}$	$\overrightarrow{\text{grad}} f = \frac{\partial f}{\partial x_i} \vec{e}_i = f_{,i} \vec{e}_i$
gradient d'un champ de vecteurs $\vec{u}$ $d\vec{u} = (\text{grad } \vec{u}) \cdot d\vec{M}$	$\text{grad } \vec{u} = \frac{\partial u_i}{\partial x_j} \vec{e}_i \otimes \vec{e}_j$ $= u_{i,j} \vec{e}_i \otimes \vec{e}_j$
champ de déformations $\epsilon = \frac{1}{2}(\text{grad } \vec{u} + t \text{grad } \vec{u})$	$\epsilon_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i})$

	Coordonnées cartésiennes orthonormées
divergence d'un champ de vecteurs $\text{div } \vec{u} = \text{Tr}(\text{grad } \vec{u})$	$\text{div } \vec{u} = u_{k,k}$
laplacien d'un scalaire $\Delta f = \text{div}(\text{grad } f)$	$\Delta f = \sum_i \frac{\partial^2 f}{\partial x_i^2} = f_{,ii}$
laplacien d'un champ de vecteurs $\vec{\Delta} \vec{u} = \text{div}(\text{grad } \vec{u})$	$\vec{\Delta} \vec{u} = u_{i,kk} \vec{e}_i$
champ de tenseurs du second ordre T	$T = T_{ij} \vec{e}_i \otimes \vec{e}_j$
divergence d'un champ de tenseurs du second ordre symétriques $\text{div } \sigma$	$\text{div } \sigma = \sigma_{ij,j} \vec{e}_i$
changement de coordonnées pour les vecteurs de base	
changement de coordonnées pour un vecteur $\vec{u}$	$\vec{u} = u_i \vec{e}_i$
changement de coordonnées pour un tenseur du second ordre symétrique ϵ	$\epsilon = \epsilon_{ij} \vec{e}_i \otimes \vec{e}_j$

Coordonnées cylindriques	
Définitions	 $\left\{ \begin{array}{l} \vec{e}_r = \frac{\partial \vec{M}}{\partial r} \\ \vec{e}_\theta = \frac{1}{r} \frac{\partial \vec{M}}{\partial \theta} \\ \vec{e}_z = \frac{\partial \vec{M}}{\partial z} \end{array} \right.$
coordonnées	$\overrightarrow{OM} = r\vec{e}_r + z\vec{e}_z$
champ de vecteurs $\vec{u}(M)$	$\vec{u}(M) = u_r\vec{e}_r + u_\theta\vec{e}_\theta + u_z\vec{e}_z$
vecteurs unitaires de base	$d\vec{e}_r = d\theta\vec{e}_\theta \quad d\vec{e}_\theta = -d\theta\vec{e}_r \quad d\vec{e}_z = 0$ $d\vec{M} = dr\vec{e}_r + r d\theta\vec{e}_\theta + dz\vec{e}_z$
gradient d'un scalaire f $df = \overrightarrow{\text{grad}} f \cdot d\vec{M}$	$\overrightarrow{\text{grad}} f = \frac{\partial f}{\partial r}\vec{e}_r + \frac{1}{r} \frac{\partial f}{\partial \theta}\vec{e}_\theta + \frac{\partial f}{\partial z}\vec{e}_z$
gradient d'un champ de vecteurs $\vec{u}$ $d\vec{u} = (\text{grad } \vec{u}) \cdot d\vec{M}$	$\text{grad } \vec{u} = \begin{bmatrix} \frac{\partial u_r}{\partial r} & \frac{1}{r} \left(\frac{\partial u_r}{\partial \theta} - u_\theta \right) & \frac{\partial u_r}{\partial z} \\ \frac{\partial u_\theta}{\partial r} & \frac{1}{r} \left(\frac{\partial u_\theta}{\partial \theta} + u_r \right) & \frac{\partial u_\theta}{\partial z} \\ \frac{\partial u_z}{\partial r} & \frac{1}{r} \frac{\partial u_z}{\partial \theta} & \frac{\partial u_z}{\partial z} \end{bmatrix}$
champ de déformations $\epsilon = \frac{1}{2}(\text{grad } \vec{u} + {}^t\text{grad } \vec{u})$	$\epsilon = \begin{bmatrix} \frac{\partial u_r}{\partial r} & \frac{1}{2} \frac{\partial u_\theta}{\partial r} + \frac{1}{2r} \left(\frac{\partial u_r}{\partial \theta} - u_\theta \right) & \frac{1}{2} \left(\frac{\partial u_r}{\partial z} + \frac{\partial u_z}{\partial r} \right) \\ - & \frac{1}{r} \left(\frac{\partial u_\theta}{\partial \theta} + u_r \right) & \frac{1}{2} \left(\frac{\partial u_\theta}{\partial z} + \frac{1}{r} \frac{\partial u_z}{\partial \theta} \right) \\ \text{sym} & - & \frac{\partial u_z}{\partial z} \end{bmatrix}$

	Coordonnées cylindriques
divergence d'un champ de vecteurs $\text{div } \vec{u} = \text{Tr}(\text{grad } \vec{u})$	$\text{div } \vec{u} = \frac{\partial u_r}{\partial r} + \frac{1}{r} \frac{\partial u_\theta}{\partial \theta} + \frac{u_r}{r} + \frac{\partial u_z}{\partial z}$
laplacien d'un scalaire $\Delta f = \text{div}(\text{grad } f)$	$\Delta f = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r} \frac{\partial f}{\partial r} + \frac{1}{r^2} \frac{\partial^2 f}{\partial \theta^2} + \frac{\partial^2 f}{\partial z^2}$
laplacien d'un champ de vecteurs $\Delta \vec{u} = \text{div}(\text{grad } \vec{u})$	$\begin{aligned} \Delta \vec{u} = & \left(\Delta u_r - \frac{2}{r^2} \frac{\partial u_\theta}{\partial \theta} - \frac{u_r}{r^2} \right) \vec{e}_r \\ & + \left(\Delta u_\theta + \frac{2}{r^2} \frac{\partial u_r}{\partial \theta} - \frac{u_\theta}{r^2} \right) \vec{e}_\theta + \Delta u_z \vec{e}_z \end{aligned}$
champ de tenseurs du second ordre $\mathbf{T}$	$\begin{aligned} \mathbf{T} = & T_{rr} \vec{e}_r \otimes \vec{e}_r + T_{r\theta} \vec{e}_r \otimes \vec{e}_\theta + T_{rz} \vec{e}_r \otimes \vec{e}_z \\ & + T_{\theta r} \vec{e}_\theta \otimes \vec{e}_r + T_{\theta\theta} \vec{e}_\theta \otimes \vec{e}_\theta + T_{\theta z} \vec{e}_\theta \otimes \vec{e}_z \\ & + T_{zr} \vec{e}_z \otimes \vec{e}_r + T_{z\theta} \vec{e}_z \otimes \vec{e}_\theta + T_{zz} \vec{e}_z \otimes \vec{e}_z \end{aligned}$
divergence d'un champ de tenseurs du second ordre symétriques $\text{div } \vec{\sigma}$	$\begin{aligned} \text{div } \vec{\sigma} = & \left(\frac{\partial \sigma_{rr}}{\partial r} + \frac{1}{r} \frac{\partial \sigma_{r\theta}}{\partial \theta} + \frac{\partial \sigma_{rz}}{\partial z} + \frac{\sigma_{rr} - \sigma_{\theta\theta}}{r} \right) \vec{e}_r \\ & + \left(\frac{\partial \sigma_{\theta r}}{\partial r} + \frac{1}{r} \frac{\partial \sigma_{\theta\theta}}{\partial \theta} + \frac{\partial \sigma_{\theta z}}{\partial z} - \frac{2\sigma_{r\theta}}{r} \right) \vec{e}_\theta \\ & + \left(\frac{\partial \sigma_{zr}}{\partial r} + \frac{1}{r} \frac{\partial \sigma_{z\theta}}{\partial \theta} + \frac{\partial \sigma_{zz}}{\partial z} + \frac{\sigma_{zr}}{r} \right) \vec{e}_z \end{aligned}$
changement de coordonnées pour les vecteurs de base	$\begin{aligned} \vec{e}_r &= \cos \theta \vec{e}_1 + \sin \theta \vec{e}_2 \\ \vec{e}_\theta &= -\sin \theta \vec{e}_1 + \cos \theta \vec{e}_2 \\ \vec{e}_z &= \vec{e}_3 \\ \\ \vec{e}_1 &= \cos \theta \vec{e}_r - \sin \theta \vec{e}_\theta \\ \vec{e}_2 &= \sin \theta \vec{e}_r + \cos \theta \vec{e}_\theta \\ \vec{e}_3 &= \vec{e}_z \end{aligned}$

	Coordonnées cylindriques
changement de coordonnées pour un vecteur $\vec{u}$	$u_1 = u_r \cos \theta - u_\theta \sin \theta$ $u_2 = u_r \sin \theta + u_\theta \cos \theta$ $u_3 = u_z$ $u_r = u_1 \cos \theta + u_2 \sin \theta$ $u_\theta = -u_1 \sin \theta + u_2 \cos \theta$ $u_z = u_3$
changement de coordonnées pour un tenseur du second ordre symétrique ϵ	$\epsilon_{11} = \epsilon_{rr} \cos^2 \theta + \epsilon_{\theta\theta} \sin^2 \theta - 2\epsilon_{r\theta} \sin \theta \cos \theta$ $\epsilon_{22} = \epsilon_{rr} \sin^2 \theta + \epsilon_{\theta\theta} \cos^2 \theta + 2\epsilon_{r\theta} \sin \theta \cos \theta$ $\epsilon_{33} = \epsilon_{zz}$ $\epsilon_{12} = (\epsilon_{rr} - \epsilon_{\theta\theta}) \sin \theta \cos \theta + \epsilon_{r\theta}(\cos^2 \theta - \sin^2 \theta)$ $\epsilon_{13} = \epsilon_{rz} \cos \theta - \epsilon_{\theta z} \sin \theta$ $\epsilon_{23} = \epsilon_{rz} \sin \theta + \epsilon_{\theta z} \cos \theta$ $\epsilon_{rr} = \epsilon_{11} \cos^2 \theta + \epsilon_{22} \sin^2 \theta + 2\epsilon_{12} \sin \theta \cos \theta$ $\epsilon_{\theta\theta} = \epsilon_{11} \sin^2 \theta + \epsilon_{22} \cos^2 \theta - 2\epsilon_{12} \sin \theta \cos \theta$ $\epsilon_{zz} = \epsilon_{33}$ $\epsilon_{r\theta} = (\epsilon_{22} - \epsilon_{11}) \sin \theta \cos \theta + \epsilon_{12}(\cos^2 \theta - \sin^2 \theta)$ $\epsilon_{rz} = \epsilon_{13} \cos \theta + \epsilon_{23} \sin \theta$ $\epsilon_{\theta z} = -\epsilon_{13} \sin \theta + \epsilon_{23} \cos \theta$

Coordonnées sphériques	
Définitions	 $\begin{cases} \vec{e}_r = \frac{\partial \vec{M}}{\partial r} \\ \vec{e}_\theta = \frac{1}{r} \frac{\partial \vec{M}}{\partial \theta} \\ \vec{e}_\varphi = \frac{1}{r \sin \theta} \frac{\partial \vec{M}}{\partial \varphi} \end{cases}$
coordonnées	$\overrightarrow{OM} = r \vec{e}_r$
champ de vecteurs $\vec{u}(M)$	$\vec{u}(M) = u_r \vec{e}_r + u_\theta \vec{e}_\theta + u_\varphi \vec{e}_\varphi$
vecteurs unitaires de base	$\begin{aligned} d\vec{e}_r &= d\theta \vec{e}_\theta + \sin \theta d\varphi \vec{e}_\varphi \\ d\vec{e}_\theta &= -d\theta \vec{e}_r + \cos \theta d\varphi \vec{e}_\varphi \\ d\vec{e}_\varphi &= -\sin \theta d\varphi \vec{e}_r - \cos \theta d\varphi \vec{e}_\theta \\ d\vec{M} &= dr \vec{e}_r + r d\theta \vec{e}_\theta + r \sin \theta d\varphi \vec{e}_\varphi \end{aligned}$
gradient d'un scalaire f $df = \overrightarrow{\text{grad}} f \cdot d\vec{M}$	$\overrightarrow{\text{grad}} f = \frac{\partial f}{\partial r} \vec{e}_r + \frac{1}{r} \frac{\partial f}{\partial \theta} \vec{e}_\theta + \frac{1}{r \sin \theta} \frac{\partial f}{\partial \varphi} \vec{e}_\varphi$
gradient d'un champ de vecteurs $\vec{u}$ $d\vec{u} = (\text{grad } \vec{u}) \cdot d\vec{M}$	$\text{grad } \vec{u} = \begin{bmatrix} \frac{\partial u_r}{\partial r} & \frac{1}{r} \left(\frac{\partial u_r}{\partial \theta} - u_\theta \right) & \frac{1}{r} \left(\frac{1}{\sin \theta} \frac{\partial u_r}{\partial \varphi} - u_\varphi \right) \\ \frac{\partial u_\theta}{\partial r} & \frac{1}{r} \left(\frac{\partial u_\theta}{\partial \theta} + u_r \right) & \frac{1}{r} \left(\frac{1}{\sin \theta} \frac{\partial u_\theta}{\partial \varphi} - \frac{u_\varphi}{\tan \theta} \right) \\ \frac{\partial u_\varphi}{\partial r} & \frac{1}{r} \frac{\partial u_\varphi}{\partial \theta} & \frac{1}{r} \left(\frac{1}{\sin \theta} \frac{\partial u_\varphi}{\partial \varphi} + \frac{u_\theta}{\tan \theta} + u_r \right) \end{bmatrix}$
champ de déformations $\epsilon = \frac{1}{2} (\text{grad } \vec{u} + {}^t \text{grad } \vec{u})$	$\epsilon = \begin{bmatrix} \frac{\partial u_r}{\partial r} & \frac{1}{2} \frac{\partial u_\theta}{\partial r} + \frac{1}{2r} \left(\frac{\partial u_r}{\partial \theta} - u_\theta \right) & \frac{1}{2} \frac{\partial u_\varphi}{\partial r} + \frac{1}{2r} \left(\frac{1}{\sin \theta} \frac{\partial u_r}{\partial \varphi} - u_\varphi \right) \\ - & \frac{1}{r} \left(\frac{\partial u_\theta}{\partial \theta} + u_r \right) & \frac{1}{2r} \frac{\partial u_\varphi}{\partial \theta} + \frac{1}{2r} \left(\frac{1}{\sin \theta} \frac{\partial u_\theta}{\partial \varphi} - \frac{u_\varphi}{\tan \theta} \right) \\ \text{sym} & & \frac{1}{r} \left(\frac{1}{\sin \theta} \frac{\partial u_\varphi}{\partial \varphi} + \frac{u_\theta}{\tan \theta} + u_r \right) \end{bmatrix}$

	Coordonnées sphériques
divergence d'un champ de vecteurs $\text{div } \vec{u} = \text{Tr}(\text{grad } \vec{u})$	$\text{div } \vec{u} = \frac{\partial u_r}{\partial r} + \frac{1}{r} \frac{\partial u_\theta}{\partial \theta} + 2 \frac{u_r}{r} + \frac{1}{r \sin \theta} \frac{\partial u_\varphi}{\partial \varphi} + \frac{u_\theta}{r \tan \theta}$
laplacien d'un scalaire $\Delta f = \text{div}(\text{grad } f)$	$\Delta f = \frac{\partial^2 f}{\partial r^2} + \frac{2}{r} \frac{\partial f}{\partial r} + \frac{1}{r^2} \frac{\partial^2 f}{\partial \theta^2} + \frac{1}{r^2 \tan \theta} \frac{\partial f}{\partial \theta} + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 f}{\partial \varphi^2}$
laplacien d'un champ de vecteurs $\Delta \vec{u} = \text{div}(\text{grad } \vec{u})$	$\begin{aligned} \Delta \vec{u} = & \left(\Delta u_r - \frac{2u_r}{r^2} - \frac{2}{r^2 \sin \theta} \frac{\partial(u_\theta \sin \theta)}{\partial \theta} - \frac{2}{r^2 \sin \theta} \frac{\partial u_\varphi}{\partial \varphi} \right) \vec{e}_r \\ & + \left(\Delta u_\theta + \frac{2}{r^2} \frac{\partial u_r}{\partial \theta} - \frac{u_\theta}{r^2 \sin^2 \theta} - \frac{2 \cos \theta}{r^2 \sin^2 \theta} \frac{\partial u_\varphi}{\partial \varphi} \right) \vec{e}_\theta \\ & + \left(\Delta u_\varphi + \frac{2}{r^2 \sin^2 \theta} \frac{\partial u_r}{\partial \varphi} + \frac{2 \cos \theta}{r^2 \sin^2 \theta} \frac{\partial u_\theta}{\partial \varphi} - \frac{u_\varphi}{r^2 \sin^2 \theta} \right) \vec{e}_\varphi \end{aligned}$
champ de tenseurs du second ordre T	$\begin{aligned} T = & T_{rr} \vec{e}_r \otimes \vec{e}_r + T_{r\theta} \vec{e}_r \otimes \vec{e}_\theta + T_{r\varphi} \vec{e}_r \otimes \vec{e}_\varphi \\ & + T_{\theta r} \vec{e}_\theta \otimes \vec{e}_r + T_{\theta\theta} \vec{e}_\theta \otimes \vec{e}_\theta + T_{\theta\varphi} \vec{e}_\theta \otimes \vec{e}_\varphi \\ & + T_{\varphi r} \vec{e}_\varphi \otimes \vec{e}_r + T_{\varphi\theta} \vec{e}_\varphi \otimes \vec{e}_\theta + T_{\varphi\varphi} \vec{e}_\varphi \otimes \vec{e}_\varphi \end{aligned}$
divergence d'un champ de tenseurs du second ordre symétriques $\text{div } \vec{\sigma}$	$\begin{aligned} \text{div } \vec{\sigma} = & \left[\frac{\partial \sigma_{rr}}{\partial r} + \frac{1}{r} \frac{\partial \sigma_{r\theta}}{\partial \theta} + \frac{1}{r \sin \theta} \frac{\partial \sigma_{r\varphi}}{\partial \varphi} + \frac{1}{r} (2\sigma_{rr} - \sigma_{\theta\theta} - \sigma_{\varphi\varphi} + \sigma_{r\theta} \cot \theta) \right] \vec{e}_r \\ & + \left[\frac{\partial \sigma_{\theta r}}{\partial r} + \frac{1}{r} \frac{\partial \sigma_{\theta\theta}}{\partial \theta} + \frac{1}{r \sin \theta} \frac{\partial \sigma_{\theta\varphi}}{\partial \varphi} + \frac{1}{r} (\sigma_{\theta\theta} \cot \theta - \sigma_{\varphi\varphi} \cot \theta + 3\sigma_{r\theta}) \right] \vec{e}_\theta \\ & + \left[\frac{\partial \sigma_{\varphi r}}{\partial r} + \frac{1}{r} \frac{\partial \sigma_{\varphi\theta}}{\partial \theta} + \frac{1}{r \sin \theta} \frac{\partial \sigma_{\varphi\varphi}}{\partial \varphi} + \frac{1}{r} (3\sigma_{r\varphi} + \frac{2\sigma_{\theta\varphi}}{\tan \theta}) \right] \vec{e}_\varphi \end{aligned}$
changement de coordonnées pour les vecteurs de base	$\begin{aligned} \vec{e}_r &= \sin \theta \cos \varphi \vec{e}_1 + \sin \theta \sin \varphi \vec{e}_2 + \cos \theta \vec{e}_3 \\ \vec{e}_\theta &= \cos \theta \cos \varphi \vec{e}_1 + \cos \theta \sin \varphi \vec{e}_2 - \sin \theta \vec{e}_3 \\ \vec{e}_\varphi &= -\sin \varphi \vec{e}_1 + \cos \varphi \vec{e}_2 \\ \vec{e}_1 &= \sin \theta \cos \varphi \vec{e}_r + \cos \theta \cos \varphi \vec{e}_\theta - \sin \varphi \vec{e}_\varphi \\ \vec{e}_2 &= \sin \theta \sin \varphi \vec{e}_r + \cos \theta \sin \varphi \vec{e}_\theta + \cos \varphi \vec{e}_\varphi \\ \vec{e}_3 &= \cos \theta \vec{e}_r - \sin \theta \vec{e}_\theta \end{aligned}$

	Coordonnées sphériques
changement de coordonnées	$u_1 = u_r \sin \theta \cos \varphi + u_\theta \cos \theta \cos \varphi - u_\varphi \sin \varphi$ $u_2 = u_r \sin \theta \sin \varphi + u_\theta \cos \theta \sin \varphi + u_\varphi \cos \varphi$ $u_3 = u_r \cos \theta - u_\theta \sin \theta$
pour un vecteur $\vec{u}$	$u_r = u_1 \sin \theta \cos \varphi + u_2 \sin \theta \sin \varphi + u_3 \cos \theta$ $u_\theta = u_1 \cos \theta \cos \varphi + u_2 \cos \theta \sin \varphi - u_3 \sin \theta$ $u_\varphi = -u_1 \sin \varphi + u_2 \cos \varphi$
changement de coordonnées pour un tenseur du second ordre symétrique ϵ	$\epsilon_{11} = \epsilon_{rr} \sin^2 \theta \cos^2 \varphi + \epsilon_{\theta\theta} \cos^2 \theta \cos^2 \varphi + \epsilon_{\varphi\varphi} \sin^2 \varphi$ $+ 2\epsilon_{r\theta} \sin \theta \cos \theta \cos^2 \varphi - 2\epsilon_{r\varphi} \sin \theta \sin \varphi \cos \varphi - 2\epsilon_{\theta\varphi} \cos \theta \sin \varphi \cos \varphi$ $\epsilon_{22} = \epsilon_{rr} \sin^2 \theta \sin^2 \varphi + \epsilon_{\theta\theta} \cos^2 \theta \sin^2 \varphi + \epsilon_{\varphi\varphi} \cos^2 \varphi$ $+ 2\epsilon_{r\theta} \sin \theta \cos \theta \sin^2 \varphi + 2\epsilon_{r\varphi} \sin \theta \sin \varphi \cos \varphi + 2\epsilon_{\theta\varphi} \cos \theta \sin \varphi \cos \varphi$ $\epsilon_{33} = \epsilon_{rr} \cos^2 \theta + \epsilon_{\theta\theta} \sin^2 \theta - 2\epsilon_{r\theta} \sin \theta \cos \theta$ $\epsilon_{12} = \epsilon_{rr} \sin^2 \theta \sin \varphi \cos \varphi + \epsilon_{\theta\theta} \cos^2 \theta \sin \varphi \cos \varphi - \epsilon_{\varphi\varphi} \sin \varphi \cos \varphi$ $+ 2\epsilon_{r\theta} \sin \theta \cos \theta \sin \varphi \cos \varphi + (\epsilon_{r\varphi} \sin \theta + \epsilon_{\theta\varphi} \cos \theta)(\cos^2 \varphi - \sin^2 \varphi)$ $\epsilon_{13} = (\epsilon_{rr} - \epsilon_{\theta\theta}) \sin \theta \cos \theta \cos \varphi + \epsilon_{r\theta}(\cos^2 \theta - \sin^2 \theta) \cos \varphi$ $- \epsilon_{r\varphi} \cos \theta \sin \varphi + \epsilon_{\theta\varphi} \sin \theta \sin \varphi$ $\epsilon_{23} = (\epsilon_{rr} - \epsilon_{\theta\theta}) \sin \theta \cos \theta \sin \varphi + \epsilon_{r\theta}(\cos^2 \theta - \sin^2 \theta) \sin \varphi$ $+ \epsilon_{r\varphi} \cos \theta \cos \varphi - \epsilon_{\theta\varphi} \sin \theta \cos \varphi$ $\epsilon_{rr} = \epsilon_{11} \sin^2 \theta \cos^2 \varphi + \epsilon_{22} \sin^2 \theta \sin^2 \varphi + \epsilon_{33} \cos^2 \theta$ $+ 2\epsilon_{12} \sin^2 \theta \sin \varphi \cos \varphi + 2\epsilon_{13} \sin \theta \cos \theta \cos \varphi + 2\epsilon_{23} \sin \theta \cos \theta \sin \varphi$ $\epsilon_{\theta\theta} = \epsilon_{11} \cos^2 \theta \cos^2 \varphi + \epsilon_{22} \cos^2 \theta \sin^2 \varphi + \epsilon_{33} \sin^2 \theta$ $+ 2\epsilon_{12} \cos^2 \theta \sin \varphi \cos \varphi - 2\epsilon_{13} \sin \theta \cos \theta \cos \varphi - 2\epsilon_{23} \sin \theta \cos \theta \sin \varphi$ $\epsilon_{\varphi\varphi} = \epsilon_{11} \sin^2 \varphi + \epsilon_{22} \cos^2 \varphi - 2\epsilon_{12} \sin \varphi \cos \varphi$ $\epsilon_{r\theta} = \epsilon_{11} \sin \theta \cos \theta \cos^2 \varphi + \epsilon_{22} \sin \theta \cos \theta \sin^2 \varphi - \epsilon_{33} \sin \theta \cos \theta$ $+ 2\epsilon_{12} \sin \theta \cos \theta \sin \varphi \cos \varphi + (\epsilon_{13} \cos \varphi + \epsilon_{23} \sin \varphi)(\cos^2 \theta - \sin^2 \theta)$ $\epsilon_{r\varphi} = (\epsilon_{22} - \epsilon_{11}) \sin \theta \sin \varphi \cos \varphi + \epsilon_{12} \sin \theta (\cos^2 \varphi - \sin^2 \varphi)$ $- \epsilon_{13} \cos \theta \sin \varphi + \epsilon_{23} \cos \theta \cos \varphi$ $\epsilon_{\theta\varphi} = (\epsilon_{22} - \epsilon_{11}) \cos \theta \sin \varphi \cos \varphi + \epsilon_{12} \cos \theta (\cos^2 \varphi - \sin^2 \varphi)$ $+ \epsilon_{13} \sin \theta \sin \varphi - \epsilon_{23} \sin \theta \cos \varphi$

$\gamma = \mu_{12}$

$$f(\gamma) = \frac{1}{6} \left((\gamma_{xx} - \gamma_{yy})^2 + (\gamma_{yy} - \gamma_{zz})^2 + (\gamma_{xx} - \gamma_{zz})^2 \right) + \gamma_{xy}^2 + \gamma_{yz}^2 + \gamma_{xz}^2$$

	E, ν	E, μ	k, ν	k, μ	λ, μ
E	E	E	$3(1-2\nu)k$	$\frac{9k}{1+3k/\mu}$	$\frac{\mu(3+2\mu/\lambda)}{1+\mu/\lambda}$
ν	ν	$-1+E/2\mu$	ν	$\frac{1-2\mu/3k}{2+2\mu/3k}$	$\frac{1}{2(1+\mu/\lambda)}$
μ	$\frac{E}{2(1+\nu)}$	μ	$\frac{3(1-2\nu)k}{2(1+\nu)}$	μ	μ
k	$\frac{E}{3(1-2\nu)}$	$\frac{E}{9-3E/\mu}$	k	k	$\lambda+2\mu/3$
λ	$\frac{E\nu}{(1+\nu)(1-2\nu)}$	$\frac{E(1-2\mu/E)}{3-E/\mu}$	$\frac{3k\nu}{1+\nu}$	$k-2\mu/3$	λ

g, 0

$$\frac{2\mu(1+\nu)}{3(1-2\nu)}$$

module de Young E

coefficient de Poisson ν

module d'élasticité en cisaillement (ou de Coulomb) μ

module d'incompressibilité k

modules de Lamé λ, μ

$$3\lambda + 2\mu = 3k = \frac{E}{1-2\nu}$$

$$3k = \lambda + \frac{2}{3}\mu + \frac{\mu}{3}$$

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

$$\Rightarrow 2\nu = (1-2\nu) \Rightarrow$$

$$2\nu + 2\nu = 1$$

$$4\nu = 1 \quad \nu = \frac{1}{4}$$

Appendix C

Gauss integration rules

Extract taken from the book *Computational contact mechanics* by Wriggers (2006).

A

Gauss integration rules

In finite element computations we always have to evaluate integrals, e.g. the weak form or the tangent matrices. These integrations can be performed on the element level, see Section 8.3, or as for example in the mortar or NITSCHÉ discretization schemes for contact on segment level, see Sections 8.4 and 8.4.3. Since isoparametric elements are usually employed for the discretization, an exact integration is no longer possible. Thus we need numerical integration. These are usually performed on the reference element $\Omega_\square$, see Chapter 7. Here we focus especially on contact problems, hence the integration rules are only stated up to two dimensions. These are needed in Chapter 8. For three-dimensional rules applied for three-dimensional solids we refer to Zienkiewicz and Taylor (2000b) or Dhatt and Touzot (1985).

A.1 One-dimensional Integration

Since integration is carried out in finite element analysis in the reference configuration, $\xi \in [-1, +1]$, all values have to be transformed to this configuration:

$$\int_{(X)} g(X) dX = \int_{-1}^{+1} g(\xi) \frac{dX}{d\xi} d\xi = \int_{-1}^{+1} g(\xi) J_e(\xi) d\xi. \quad (\text{A.1})$$

$g(\xi)$ is the function which has to be integrated and J_e is the JACOBIAN of the transformation to the reference configuration which can be computed using the isoparametric map defined in Section 7.1 by

$$J_e = \frac{dX}{d\xi} = \sum_{I=1}^n \frac{\partial N_I(\xi)}{\partial \xi} X_I, \quad (\text{A.2})$$

where N_I are the shape functions and X_I are the nodal coordinates. The integration will be done numerically, since the product $g(\xi) J_e(\xi)$ is in general

Table A.1. One-dimensional GAUSS integration

n_p	p	ξ_p	W_p	
1	1	0	2	
2	1	$1/\sqrt{3}$	1	
	2	$-1/\sqrt{3}$	1	
3	1	$-\sqrt{3/5}$	$5/9$	
	2	0	$8/9$	
	3	$+\sqrt{3/5}$	$5/9$	

no longer a polynomial. Hence, the integral (A.1) will be approximated by the sum

$$\int_{-1}^{+1} g(\xi) J_e(\xi) d\xi \approx \sum_{p=1}^{n_p} g(\xi_p) J_e(\xi_p) W_p. \quad (\text{A.3})$$

W_p are weighting factors and ξ_p denote the coordinates of the evaluation points. The locations ξ_p and the weighting factors W_p are stated in Table A.1 up to the order of $n_p = 3$ for a GAUSS integration.

Polynomials of order $p = 2n_p - 1$ are integrated exactly by n_p evaluation points. These rules can be used for two-dimensional contact elements, e.g. see Section 8.3.

A.2 Two-dimensional Integration

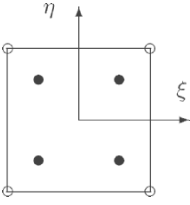
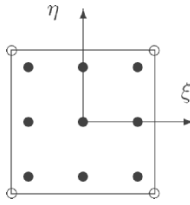
For evaluation of the weak form in (3.59) which is valid for two-dimensional problems, or for evaluation of the contact element (8.60) or (8.65), we need an integration of the interpolation functions and its derivatives over the element domain Ω_e . For this purpose, it is advantageous to transform the integral to the ξ - η coordinate system in the reference element $\Omega_\square$:

$$\int_{(\Omega_e)} g(\mathbf{X}) dA = \int_{(\Omega_\square)} g(\boldsymbol{\xi}) \det \mathbf{J}_e(\boldsymbol{\xi}) d\boldsymbol{\xi} = \int_{-1}^{+1} \int_{-1}^{+1} g(\xi, \eta) \det \mathbf{J}_e d\xi d\eta. \quad (\text{A.4})$$

Integration over $\Omega_\square$ is performed by a numerical quadrature formula, since the product $g(\boldsymbol{\xi}) \det \mathbf{J}_e(\boldsymbol{\xi})$ does in general not yield a polynomial. Thus, we obtain

$$\int_{-1}^{+1} \int_{-1}^{+1} g(\xi, \eta) \det \mathbf{J}_e d\xi d\eta \approx \sum_{p=1}^{n_p} g(\xi_p, \eta_p) \det \mathbf{J}_e(\xi_p, \eta_p) W_p. \quad (\text{A.5})$$

Table A.2. Two-dimensional GAUSS quadrature for rectangular elements

m	n_p	p	ξ_p	η_p	W_p	Position of points
1	1	1	0	0	4	
3	4	1	$-1/\sqrt{3}$	$-1/\sqrt{3}$	1	
		2	$+1/\sqrt{3}$	$-1/\sqrt{3}$	1	
		3	$-1/\sqrt{3}$	$+1/\sqrt{3}$	1	
		4	$+1/\sqrt{3}$	$+1/\sqrt{3}$	1	
5	9	1	$-\sqrt{3/5}$	$-\sqrt{3/5}$	25/81	
		2	0	$-\sqrt{3/5}$	40/81	
		3	$+\sqrt{3/5}$	$-\sqrt{3/5}$	25/81	
		4	$-\sqrt{3/5}$	0	40/81	
		5	0	0	64/81	
		6	$+\sqrt{3/5}$	0	40/81	
		7	$-\sqrt{3/5}$	$+\sqrt{3/5}$	25/81	
		8	0	$+\sqrt{3/5}$	40/81	
		9	$+\sqrt{3/5}$	$+\sqrt{3/5}$	25/81	

The weighting factors W_p and the coordinates of the quadrature points ξ_p and η_p are contained in Table A.2 for a GAUSS quadrature up to a number of $n_p = 3 \times 3$ points. These integration rules are exact for polynomials up to the order $i + k \leq m$. We note that the integration rules follow from the one-dimensional integration rules via a product formula. Usually, GAUSS rules are applied in finite element computations due to their accuracy. Thus we do not discuss other rules here. More quadrature rules can be found in Dhatt and Touzot (1985), for example.

The transformation to the reference element is different for triangular elements. In general, we obtain the following relation:

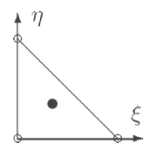
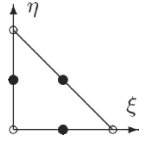
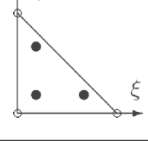
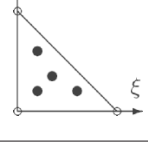
$$\int_{(\Omega_e)} g(\mathbf{X}) dA = \int_0^1 \int_0^{1-\xi} g(\xi, \eta) \det \mathbf{J}_e d\eta d\xi, \quad (\text{A.6})$$

which again can be evaluated using the quadrature rule

$$\int_0^1 \int_0^{1-\xi} g(\xi, \eta) \det \mathbf{J}_e d\eta d\xi \approx \sum_{p=1}^{n_p} g(\xi_p, \eta_p) \det \mathbf{J}_e(\xi_p, \eta_p) W_p. \quad (\text{A.7})$$

Table A.3 contains the associated quadrature points and weighting factors for an element with side length 1. The formulas are exact for polynomials $\xi^k \eta^l$

Table A.3. Two-dimensional GAUSS quadrature for triangular elements

m	n_p	p	ξ_p	η_p	W_p	Position of points
1	1	1	$1/3$	$1/3$	$1/2$	
2	3	1 2 3	$1/2$ 0 $1/2$	$1/2$ $1/2$ 0	$1/6$ $1/6$ $1/6$	
2	3	1 2 3	$1/6$ $2/3$ $1/6$	$1/6$ $1/6$ $2/3$	$1/6$ $1/6$ $1/6$	
3	4	1 2 3 4	$1/3$ $1/5$ $3/5$ $1/5$	$1/3$ $1/5$ $1/5$ $3/5$	$-27/96$ $25/96$ $25/96$ $25/96$	

up to the order m (with $m \geq k + l$). Again, different quadrature rules with different quadrature points or higher accuracy can be found in Zienkiewicz and Taylor (1989) or Dhatt and Touzot (1985), for example.