Derivation, Simulation and Validation of Poroelastic Models in Dental Biomechanics

Doctoral Dissertation submitted to the
Faculty of Informatics of the Università della Svizzera Italiana
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

presented by
Marco Favino

under the supervision of
Rolf Krause

November 2014

Dissertation Committee

Martin Gander	Université de Genève, Switzerland
Alfio Grillo	Politecnico di Torino, Italy
Rolf Krause	Università della Svizzera italiana, Switzerland
Igor Pivkin	Università della Svizzera italiana, Switzerland
Olaf Schenk	Università della Svizzera italiana, Switzerland

Dissertation accepted on 26 November 2014

Research Advisor	PhD Program Director
Rolf Krause	**Igor Pivkin, Stefan Wolf**

I certify that, except where due acknowledgement has been given, the work presented in this thesis is that of the author alone; the work has not been submitted previously, in whole or in part, to qualify for any other academic award; and the content of the thesis is the result of work which has been carried out since the official commencement date of the approved research program.

Marco Favino
Lugano, 26 November 2014

Abstract

Poroelasticity and mechanics of growth are playing an increasingly relevant role in biomechanics. This work is a self-contained and holistic presentation of the modeling and simulation of non-linear poroelasticity with and without growth inhomogeneities.

Balance laws of poroelasticity are derived in Cartesian coordinates. These allow to write the governing equations in a form that is general but also readily implementable. Closure relations are formally derived from the study of dissipation.

We propose an approximation scheme for the poroelasticity problem based on an implicit Euler method for the time discretization and a finite element method for the spatial discretization. The non-linear system is solved by means of Newton's method. Time integration of the growth tensor is discussed for the specific case in which the rate of inelastic deformations is prescribed.

We discuss the stability of the mixed finite element discretization of the arising saddle-point problem. We show that a linear finite element approximation of both the unknowns, that is not LBB compliant for the elasticity problem, is nevertheless stable when applied to the linearized poroelasticity problem. This choice enables a fast assembling phase.

The discretization of the poroelastic system may present unphysical oscillations if the spatial and temporal step-sizes are not properly chosen. We study the source of these wiggles by comparing the pressure Schur complement to a reaction-diffusion problem. From our analysis, we define a novel Péclet number for the poroelastic system and we show how it depends on the shear and bulk moduli of the solid phase. This number allows to introduce a stability condition that ensures that the solution is free of unphysical oscillations. If this condition on the Péclet number is not met, we introduce a fluid pressure Laplacian stabilization in order to remove the wiggles. This stabilization technique depends on a numerical parameter, whose optimal value is given by the derived Péclet number.

Finally, we propose a coupled elastic-poroelastic model for the simulation of a tooth-periodontal ligament system. Because of the high resolution required by this system, we develop an efficient multigrid Newton's method for the non-linear poroelasticity system. The stability condition has again a significant influence on the performances of this solver. If the condition on the Péclet number is not satisfied on all levels of the multigrid algorithm, poor convergence rates or even divergence of the solver can be

observed. The stabilization of the coarse grid operators with the optimal fluid pressure Laplacian method is a simple and efficient method to improve the convergence rate of the multigrid solver applied to this saddle-point system. We validate our coupled model against experimental measurements realized by the group of Prof. Bourauel at the University of Bonn.

Acknowledgements

Un sentito ringraziamento al mio relatore Prof. Krause, per la dedizione al lavoro e la passione che mi ha trasmesso. A lui esprimo la mia gratitudine per l'opportunità concessami di svolgere questo dottorato di ricerca e per avermi guidato e sostenuto in questa tappa fondamentale del mio percorso professionale e personale.

I also want to express my gratitude to the other members of my dissertation committee, Prof. Martin Gander, Prof. Igor Pivkin, and Prof. Olaf Schenk, for their time and interest. I am thankful to Prof. Alfio Grillo and Prof. Christoph Bourauel for the many interesting and insightful conversations.

Un ringraziamento ai tutti i miei colleghi di Lugano e Bonn, che hanno condiviso le fatiche e i piaceri del dottorato. Un grazie particolare al Dr. Dorian Krause, che durante le sue innumerevoli cene a base di pizza mi ha concesso il tempo per profonde discussioni sul principio del massimo.

I want to thank my parents and family for encouraging me during these years. In particolare grazie a mia sorella, la più brava e la più magra.

Lastly, I am forever thankful to Diana, for her everlasting moral and practical support. Her continued encouragement and inspiration made this thesis possible.

This work was partially funded by the German Research Foundation (DFG).

Contents

Figures

Tables

Introduction and motivation

Poroelasticity studies the mechanical behaviour of deformable porous media infused with a fluid. Early applications of this discipline were in the field of geomechanics. In the last decades, poroelasticity started to play an increasingly relevant role also in biomechanics. Biphasic models have been used to describe biological tissues like cartilage [FGLR+05, EM02], myocardial walls [HvCAH91], intervertebral discs [MNL11, RNL13], and periodontal ligament (PDL) [FGD+11, NPSS02, Ber08]. The equations of poroelasticity in biomechanics are highly non-linear because of the large displacements and the complex material laws of soft tissues. Moreover, the "living nature" of biomaterials is responsible for a change of internal structure and, hence, models of reorganization and growth have also to be considered.

While in elasticity comprehensive works exist, in the field of biphasic continua studies generally tend to be specialized in only one of the following aspects.

- Derivation of the governing equations. These works aim to study the formal setting of poroelasticity but, because of their abstraction, they are not directly applicable for implementation and simulation purposes.

- Analysis of material laws. These works aim to apply the biphasic approach to different physical phenomena. Simulations are usually realized with commercial softwares. Hence, the solution methods are not tuneable and the choice of material laws is restricted to the available ones.

- Numerical analysis of the linear problem. These works focus on the discretization methods for the linear problem. In this setting, stability and convergence can be studied but the obtained results are not always extendable to the non-linear case.

This thesis deals with modeling, discretization and efficient simulation of poroelastic materials. In this work, we want to present the derivation of balance laws of continuum mechanics in a more analytical setting. We want also to clarify how closure relations of poroelastic models (e.g., Darcy's law) can be derived from the study of dissipation. The arising system of equations is a non-linear saddle-point problem coupled with an evolution equation for the tensor that describes volumetric growth. Hence, we present a discretization method which is directly applicable to the solution of this differential problem.

A widely discussed and not yet completely understood problem in the discretization of the equations of poroelasticity is the phenomenon of unphysical oscillations. They are sometimes attributed to a violation of the LBB condition for saddle-point problems. We show that finite element discretizations with spaces of the same order are stable for the poroelastic system even if this choice is not LBB compliant. Moreover, spaces that are LBB stable can present these oscillations.

By comparison with reaction-diffusion problems, we define a corresponding Péclet number for the poroelasticity system. We show how the Péclet number depends on the mechanical parameters and the finite element spaces. We find a condition on this Péclet number that ensures the stability of the system. Furthermore, when this condition is not met the Péclet number provides an optimal stabilization parameter for the fluid pressure Laplacian stabilization.

Finally, we validate a poroelastic model in a fundamental and challenging problem in the filed of dental biomechanics, i.e. the simulation of the PDL. The mechanics of this soft tissue is highly elaborated because of its internal structure. The fibers that connect the tooth and the jaw bone have an elastic behaviour while the interstitial fluid has a damping effect. Because of this composition, force response of the PDL to mechanical solicitations is highly non-linear and time dependent.

Even if the biphasic nature of the PDL is well-known, this aspect has rarely been investigated. Classical mechanical models for PDL usually focus on one of the following characteristics: incompressibility, viscoelasticity, or anisotropy.

In order to reproduce the loading conditions of PDL, we created a coupled elastic-poroelastic model for the tooth-PDL system. The simulation of this system requires a very fine mesh in order to obtain physically meaningful results. To efficiently solve the large arising linear system, we use a multigrid strategy. The following aspects have to be considered concerning the application of this solution method:

- the saddle-point nature of the poroelastic system;
- the large jumps in the coefficients between tooth and PDL;
- the stability condition that must not be violated on all levels.

In this work, we present a simple, directly applicable, and efficient simulation method for the poroelasticity system. We analyze the problems that arise in the discretization and may affect convergence of efficient solution methods. We employ our simulation strategy to reproduce the loading conditions of the PDL.

This thesis is organized as follows.

Chapter 1 We derive the equations that govern the motion of elastic and poroelastic bodies.

Chapter 2 We show how the previously derived balance equations change when inhomogeneities due to growth are considered.

Chapter 3 We present the necessary instruments for the discretization and the iterative solution of the non-linear poroelastic model. We discuss in particular: time integration of differential algebraic equations, principle of virtual work, Newton's method, stability analysis of saddle-point problems, and the mixed finite element method.

Chapter 4 We discuss the presence of unphysical oscillations in the solution of the discretized poroelastic system in the linear case. We present a stability condition on the spatial and temporal step-size in the one-dimensional case. Then, by means of numerical experiments, we derive a new stability condition for the two- and three-dimensional cases and for different finite element discretizations.

Chapter 5 We present the discretization of the non-linear poroelastic system. We validate our solution algorithm and our implementation in a confined compression experiment.

Chapter 6 We apply the solution algorithm to the simulation of a coupled elastic-poroleastic model that describes a tooth-PDL system. We present a suitable multigrid solution strategy for the coupled system. Finally, we validate our model against measurements performed on a porcine and a human tooth.

The complete solution strategy has been implemented in the toolbox UG/obslib++. In particular, we created a new library for the assembling routine of non-linear mechanical problems and an interface to the PETSc solver. The tests on the stability have been implemented in Matlab and in FreeFem++. The meshes of the teeth have been provided by our co-workers from the group of Prof. Bourauel at the University of Bonn.

Chapter 1

Mathematical modelling of elasticity and poroelasticity

In this first chapter, we derive the equations governing the behavior of elastic and poroelastic materials. The former are presented in the first section and the latter in the second. The main goal of this introductory chapter is twofold: to introduce the notation and basic definitions, and to present the derivation of balance equations for these two models. We will refer to elastic bodies as monophasic continua in contrast to the biphasic continua studied in poroelasticity.

1.1 Elasticity

Continuum mechanics is a branch of mechanics that aims at the mathematical description of media modeled as a continuous body. No presumption is done on the structure of matter. In continuum mechanics, standard balances of mass, momentum, angular momentum, and energy are written as integral equations on an arbitrary subpart of the body. This approach can be used when the level of investigation (the size of the considered subpart) is coarser than the molecular length. From this point of view, the above mentioned conservation laws can be considered as averaged balances of the molecular actions. Finally, under certain mathematical hypotheses the integral balances may be written as differential pointwise equations. Continuum mechanics is divided in two main branches, that are solid mechanics and fluid mechanics. In particular, in solid mechanics the term *elasticity* refers to materials that return to their zero stress configuration when external loadings are removed.

There are several treatises on the mathematical formulation of continuum mechanics and elasticity. In particular for the presentation of kinematics we will follow [TN04], while our notation is based on [VF06].

1.1.1 Kinematics

A *body* $\mathscr{B}$ is a n-dimensional differentiable manifold (or smooth n-manifold) with $n = 1, 2, 3$. The elements thereof are called *particles* and are denoted by P. A manifold may be referred to a system of coordinates. This means that exists a one-to-one function $\mathsf{P}(\cdot)$ between particles and n-tuplet of real numbers denoted by $\mathbf{p} = (p^1, \ldots, p^n)$:

$$\mathsf{P} = \mathsf{P}(\mathbf{p}).$$

These coordinates $\mathbf{p}$ are known as *material coordinates*.

A *deformation*, or placement, of $\mathscr{B}$ is a mapping

$$\mathscr{B} \ni \mathsf{P} \longmapsto \mathsf{x} = \boldsymbol{\chi}(\mathsf{P}) \in \mathbb{R}^3.$$

onto the three dimensional Euclidean space. The point x is the *place* occupied by the particle P. If a system of Cartesian coordinate is given, $\mathbf{x} = (x^1, x^2, x^3)$ represents the *spatial* coordinates with respect to the canonical basis $\mathbf{e}^k$ of the Euclidean space. Since particles and places can be identified with their coordinates, we can write both

$$\mathbf{x} = \boldsymbol{\chi}(\mathsf{P}) \quad \text{and} \quad \mathbf{x} = \boldsymbol{\chi}(\mathbf{p}).$$

A *motion* is a time-dependent family $\boldsymbol{\chi}_t$ of deformations. The place

$$\mathbf{x} = \boldsymbol{\chi}_t(\mathsf{P}) = \boldsymbol{\chi}(\mathsf{P}, t)$$

denotes the coordinates of the place occupied by particle P at time t. The image $\mathscr{B}_t = \boldsymbol{\chi}_t(\mathscr{B})$ is the *current* configuration. The terms deformation and configuration are also used as synonyms.

In solid mechanics, physical quantities are functions defined on the current configuration of the body. This is a spatial domain that changes with time. Hence, it is customary to introduce a *reference* configuration that is a fixed domain in space employed to define physical quantities in the course of motion. This reference configuration is the image of a deformation $\boldsymbol{\varkappa}$ that maps material points P into $\mathbb{R}^3$:

$$\mathscr{B} \ni \mathsf{P} \longmapsto \mathsf{X} = \boldsymbol{\varkappa}(\mathsf{P}) \in \mathbb{R}^3.$$

If a system of general coordinates is given, $\mathbf{X} = (X^1, X^2, X^3)$ represents the coordinates of X with respect to a basis $\mathbf{E}^k$. Since we can identify a position in the reference configuration with its coordinates, we will also write

$$\mathbf{X} = \boldsymbol{\varkappa}(\mathsf{P}) \quad \text{and} \quad \mathbf{X} = \boldsymbol{\varkappa}(\mathbf{p}).$$

The reference configuration will be denoted by $\mathscr{B}_r = \boldsymbol{\varkappa}(\mathscr{B})$. The motion $\boldsymbol{\chi}(\mathsf{P}, t)$ can also be written with respect to the reference configuration $\boldsymbol{\varkappa}$ by means of the following definition

$$\mathbf{x} = \boldsymbol{\chi}(\mathbf{X}, t) := \boldsymbol{\chi}(\boldsymbol{\varkappa}^{-1}(\mathbf{X}), t). \tag{1.1}$$

We used the same symbol for the deformation defined on $\mathscr{B}$ and $\mathscr{B}_r$. In Figs. 1.1 and 1.2, we reported the plots of configurations and deformations.

The general definitions above can be used for solid continua but allow also to deal with rods, shells, and "exotic" materials such as liquid crystals. Since we will consider only three-dimensional bodies, the following simplifications can and will be introduced:

A1: we consider *simple* bodies [MH94], i.e. $\mathscr{B}$ is an open set of $\mathbb{R}^3$;

A2: we regard the inverse of the reference configuration as a definition of coordinate system in space [TN04], i.e. $\boldsymbol{\varkappa}^{-1}(\,\cdot\,) = \mathsf{P}(\,\cdot\,)$. Hence, a particle P and its material positions X have the same coordinate

$$\mathbf{p} = \mathbf{X} = \boldsymbol{\varkappa}(\mathsf{P}),$$

and the deformation $\boldsymbol{\varkappa}$ is an embedding of $\mathscr{B}$ into $\mathbb{R}^3$. We use the reference coordinates $\mathbf{X}$ as material coordinates;

A3: we will also consider the basis $\mathbf{E}^k$ as a Cartesian basis and we will impose $\mathbf{E}^k = \mathbf{e}^k$ [Bon08]. This allows to define the displacement vector as

$$\mathbf{u}(\mathbf{X}, t) = \boldsymbol{\chi}(\mathbf{X}, t) - \mathbf{X}. \tag{1.2}$$

For the assumptions above, if $\mathscr{I}$ is an interval in $\mathbb{R}$, representing a time interval, we can consider a motion as a function

$$\mathscr{B}_r \times \mathscr{I} \ni (\mathbf{X}, t) \longmapsto \mathbf{x} = \boldsymbol{\chi}(\mathbf{X}, t) \in \mathbb{R}^3. \tag{1.3}$$

The simplifications introduced are sketched in Fig. 1.3.

We will consider only regular motions for which:

B1: $\boldsymbol{\chi}$ is C^2 with respect to time;

B2: the correspondence between the material point and spatial position at each time is bijective.

The regularity assumption **B1** ensures that the material velocity

$$\dot{\mathbf{x}}(\mathbf{X}, t) = \frac{\partial \boldsymbol{\chi}}{\partial t}(\mathbf{X}, t) \tag{1.4}$$

and the material acceleration

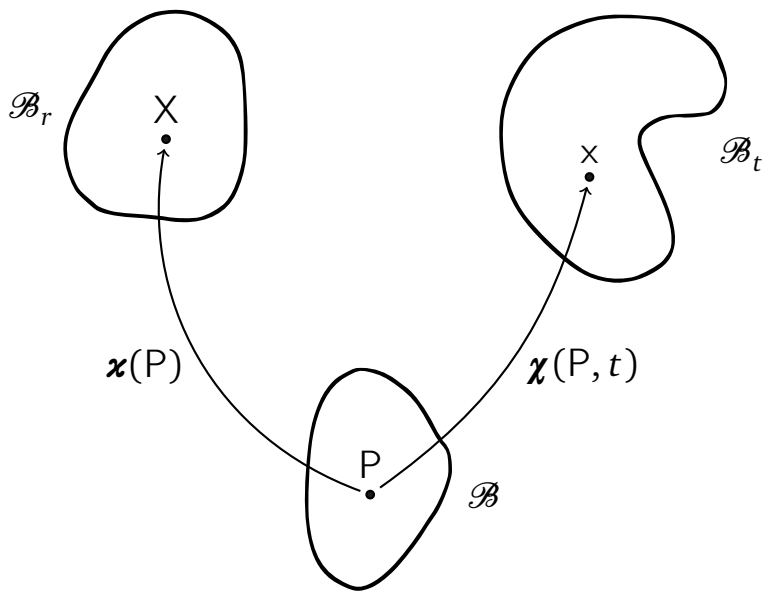


Figure 1.1. Body manifold, reference configuration, and current configuration

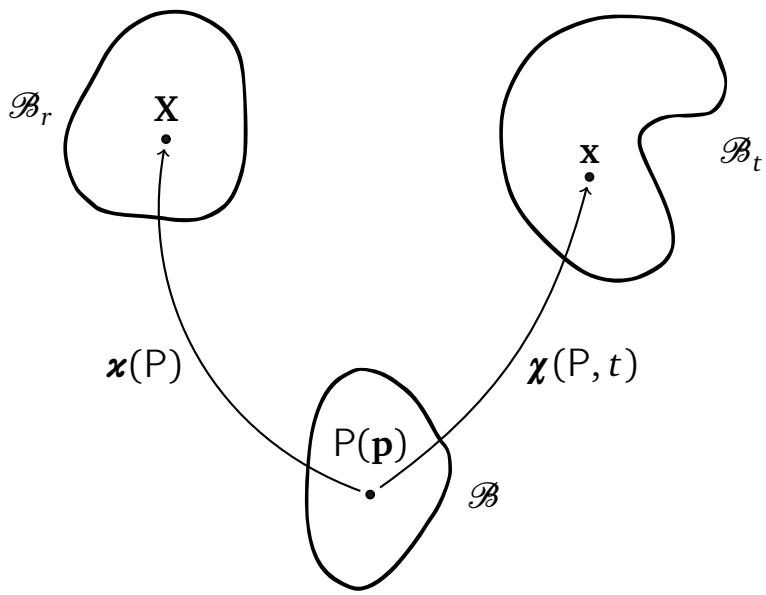


Figure 1.2. Body manifold, reference configuration, and current configuration using coordinates.

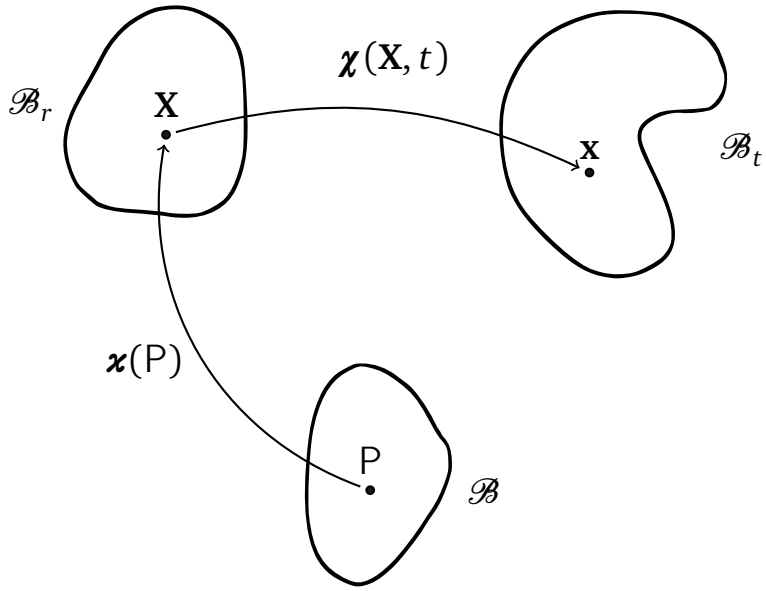


Figure 1.3. When considering simple three-dimensional bodies, the coordinates of the reference configuration can be seen as coordinates of material points.

$$\ddot{\mathbf{x}}(\mathbf{X}, t) = \frac{\partial^2 \boldsymbol{\chi}}{\partial t^2}(\mathbf{X}, t) \tag{1.5}$$

are well defined. Moreover, because of assumption **B2**, at each time t the function $\boldsymbol{\chi}(\,\cdot\,, t)$ is invertible and allows to define the inverse motion

$$\mathbf{X} = \pi(\mathbf{x}, t) \tag{1.6}$$

for which the following properties hold:

$$\boldsymbol{x} = \boldsymbol{\chi}(\pi(\mathbf{x}, t), t) \quad \text{and} \quad \mathbf{X} = \pi(\boldsymbol{\chi}(\boldsymbol{X}, t), t).$$

The image of $\boldsymbol{\chi}$ is the *trajectory* and it is defined as

$$\mathscr{T}_{\boldsymbol{\chi}} := \{(\mathbf{x}, t) \in \mathbb{R}^3 \times \mathscr{I} \mid \mathbf{x} = \boldsymbol{\chi}(\mathbf{X}, t), \mathbf{X} \in \mathscr{B}, t \in \mathscr{I}\}.$$

The trajectory is a subset of $\mathbb{R}^3 \times \mathscr{I}$ that can not be written as a Cartesian product of a subset of $\mathbb{R}^3$ and a time interval (cf. Def. (1.3)). The trajectory is the domain of the inverse motion. Hence π is a function whose domain changes with time but whose image remains constant.

The gradient of the motion with respect to the material coordinates is called *deformation gradient* and it is denoted by

$$\mathbf{F} := \nabla f = D_X f.$$

Formally $\mathbf{F}$ is a $(1,1)$ second order tensor [MH94], and its components are

$$F^k_\alpha = \frac{\partial \chi^k}{\partial X^\alpha}.$$

Here we consider only tensors defined on Euclidean space and we identify the deformation determinant with a linear transformation whose components are

$$F_{ij} = \frac{\partial \chi_i}{\partial X_j}.$$

The product $\mathbf{C} = \mathbf{F}^T F$ defines the right Green deformation tensor that allows to measure distances in the deformed configuration.

The determinant $J = J(X,t)$ of the deformation gradient is called *deformation determinant* and accounts for volume changes between the reference and the current configuration. The deformation determinant has to be strictly positive in order have local admissibility of the deformation. Since the deformation gradient is time dependent, the deformation determinant is time dependent too. The time derivative of the deformation determinant can be computed using the chain rule by the formula

$$\dot{J} = J\mathbf{F}^{-T} : \dot{\mathbf{F}} = J\,\mathrm{tr}(\dot{F}F^{-1}). \tag{1.7}$$

The inverse motion can be composed with the material velocity and acceleration in order to obtain the spatial velocity and acceleration, that are defined as

$$\mathbf{v}(\mathbf{x},t) = \dot{\mathbf{x}}(\pi(\mathbf{x},t),t) \tag{1.8}$$

and

$$\mathbf{a}(\mathbf{x},t) = \ddot{\mathbf{x}}(\pi(\mathbf{x},t),t). \tag{1.9}$$

The application of the spatial gradient to Eq. (1.8) gives us

$$\mathbf{L} = \mathrm{grad}\,\mathbf{v} = (\dot{\mathbf{F}}\mathbf{F}^{-1}) \circ \pi. \tag{1.10}$$

1.1.2 Material and spatial fields

A material field is a tensorial function

$$\mathscr{B}_r \times \mathscr{I} \ni (\mathbf{X},t) \longmapsto \mathbf{\Psi}(\mathbf{X},t),$$

where $\mathbf{\Psi}$ is a tensor in Lin_p, i.e. tensor of order $p \in \mathbb{N}$. We also include the cases of scalar functions ($p = 0$) and covectors ($p = 1$). We will identify covectors and vectors since we are in an Euclidean space. Material fields are also called Lagrangian fields.

A spatial field is a tensorial function

$$\mathscr{T}_{\chi} \ni (\mathbf{x}, t) \longmapsto \boldsymbol{\psi}(\mathbf{x}, t),$$

where $\boldsymbol{\psi}$ is a tensor. Differently from a material field, the domain of a spatial field is the image of the motion of the body.

A spatial field $\boldsymbol{\psi}$ can be composed with the motion (1.1) in order to obtain its material description

$$\boldsymbol{\psi}_m(\mathbf{X}, t) := \boldsymbol{\psi}(\boldsymbol{\chi}(\mathbf{X}, t), t)$$

and a material field $\boldsymbol{\Psi}$ can be composed with the inverse motion (1.6) to obtain its spatial description

$$\boldsymbol{\Psi}_s(\mathbf{x}, t) := \boldsymbol{\Psi}(\boldsymbol{\pi}(\mathbf{x}, t), t).$$

To denote the change of description, we employ the operators $(\,\cdot\,)_m$ and $(\,\cdot\,)_s$ that transform a spatial field into a material one and vice versa. Definitions (1.8) and (1.9) are examples of a change of description.

We will use the operators $\nabla\boldsymbol{\Psi} = D_{\mathbf{X}}\Psi$ and grad $\boldsymbol{\psi} = D_{\mathbf{x}}\boldsymbol{\psi}$ to denote the gradient with respect to the material and the spatial coordinates respectively. Moreover for tensor of order $p \geq 1$, the operators Div $\boldsymbol{\Psi} = \mathrm{tr}(\nabla\boldsymbol{\Psi})$ and div $\boldsymbol{\psi} = \mathrm{tr}\,(\mathrm{grad}\,\boldsymbol{\psi})$ to denote the divergence with respect to the material and the spatial coordinates respectively.

We will also use different symbols for time derivatives of material and spatial fields. As in (1.4), the symbol

$$\dot{\boldsymbol{\Psi}} = \frac{\partial \boldsymbol{\Psi}}{\partial t}$$

denotes the partial time derivative of a spatial field, while the partial time derivative of a material function will be denoted by

$$\boldsymbol{\psi}' = \frac{\partial \boldsymbol{\psi}}{\partial t}.$$

The convective derivative of the spatial field is the total time derivative of a spatial field. It is denoted by D_t and it it defined as

$$D_t\boldsymbol{\psi} = \left(\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\psi}_m\right)_s. \tag{1.11}$$

The argument of operator $(\,\cdot\,)_s$ in (1.11) can be computed as

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\psi}(\boldsymbol{\chi}(\mathbf{X}, t), t) = (\boldsymbol{\psi}')_m + (\mathrm{grad}\boldsymbol{\psi})_m \cdot \dot{\mathbf{x}}, \tag{1.12}$$

and finally written in spatial coordinates as

$$D_t \boldsymbol{\psi}(\mathbf{x}, t) = \boldsymbol{\psi}' + \text{grad}\boldsymbol{\psi} \cdot \mathbf{v}. \tag{1.13}$$

Since the spatial acceleration can be written

$$\mathbf{a} = \left(\frac{\mathrm{d}}{\mathrm{d}t} \mathbf{v}_m \right)_s,$$

the convective derivative has to be used also for the computation of the spatial acceleration. This leads to the following identity

$$\mathbf{a} = \mathbf{v}' + (\text{grad } \mathbf{v})\mathbf{v}. \tag{1.14}$$

Algebraic view of tensors and differential operators

For implementational purposes, it is convenient to consider first order tensor as column vectors and second order tensor as 3×3 matrices. The gradient of a scalar function f is a column vector ∇f with components

$$(\nabla f)_i = \frac{\partial f}{\partial p_i} \quad i = 1, 2, 3.$$

The gradient of a vector function $\boldsymbol{\psi}$ is a second order tensor $\nabla \boldsymbol{\psi}$ with components

$$(\nabla \boldsymbol{\psi})_{ij} = \frac{\partial \psi_i}{\partial p_j} \quad i, j = 1, 2, 3.$$

This means that the gradient of a vector field $\nabla \boldsymbol{\psi}$ can be seen as a matrix whose i-th row equals the gradient of the i-th component of $\boldsymbol{\psi}$.

The divergence of a first order tensor $\boldsymbol{\psi}$ is defined as

$$\text{Div } \boldsymbol{\psi} = \sum_{i=1}^{3} \frac{\partial \psi_i}{\partial p_i}$$

or, using Einstein notation,

$$\text{Div } \boldsymbol{\psi} = \psi_{i,i}$$

where the comma denotes the derivative and we use the Einstein convention of summing on repeated indices.

The divergence of a second order tensor $\boldsymbol{\Psi}$ is a first order tensor whose components are

$$(\text{Div}\boldsymbol{\Psi})_i = \sum_{j=1}^{3} \frac{\partial \psi_{ij}}{\partial p_j}$$

or in Einstein notation

$$(\text{Div } \mathbf{\Psi})_i = \psi_{ij,j}.$$

Thus, the divergence operator applied to a second order tensor $\mathbf{\Psi}$ is the application of the divergence to each row $\mathbf{\Psi}_i$.

1.1.3 Piola tranformation

The balance laws of mechanics are differential conservation equations defined on the current configuration, that is, a spatial domain which changes with time. It is convenient, in order to perform numerical simulations, to write them in the reference configuration. Piola transformations, named after Gabrio Piola, allow to transform conservation laws on the current configuration into conservation laws written on the current configuration using material coordinates. The derivation of such laws is based on two ingredients: a change of integration variable and the chain rule. In the following, $\mathscr{P}$ denotes a subset of $\mathscr{B}_r$ and $\mathscr{P}_t = \boldsymbol{\chi}(\mathscr{P}, t)$ denotes its image through the motion at time t.

Transformation of the gradient of a scalar function

Let ϕ be a spatial scalar field. The integral of its gradient can be written in material coordinates as

$$\int_{\mathscr{P}_t} \text{grad } \phi \, \mathrm{d}\mathbf{x} = \int_{\mathscr{P}} (\text{grad } \phi)_m J \, \mathrm{d}\mathbf{X}. \tag{1.15}$$

The argument of the right hand side of (1.15) is a spatial gradient of a "material" function. It can be computed employing the chain rule, so that we get

$$(\text{grad}\phi)_m = \mathbf{F}^{-T} \nabla \phi_m. \tag{1.16}$$

Finally, substituting (1.16) in (1.15), we obtain

$$\int_{\mathscr{P}_t} \text{grad}\phi \, \mathrm{d}\mathbf{x} = \int_{\mathscr{P}} \mathbf{F}^{-T} \nabla \phi_m J \, \mathrm{d}\mathbf{X}. \tag{1.17}$$

Transformation of the gradient of a vector function

Let $\boldsymbol{\phi} \in \text{Lin}_1$ be a spatial vector field. Since its gradient can be seen as the three row vectors of the gradient of its component, we can simply consider the transposition of (1.17), so that we obtain

$$\int_{\mathscr{P}_t} \text{grad}\boldsymbol{\phi} \, \mathrm{d}\mathbf{x} = \int_{\mathscr{P}} \nabla \boldsymbol{\phi}_m \mathbf{F}^{-1} J \, \mathrm{d}\mathbf{X}. \tag{1.18}$$

Piola transformation of the divergence of a vector function

The transformation of the divergence of a spatial vector field $\boldsymbol{\phi} \in \mathrm{Lin}_1$ is more involved. It can be derived using the Piola identity $\mathrm{Div}(J\mathbf{F}^{-T}) = \mathbf{0}$ [MH94] and the fact that

$$\mathrm{div}\ \boldsymbol{\phi} = \mathrm{tr}\ (\mathrm{grad}\ \boldsymbol{\phi}). \tag{1.19}$$

In a way similar to (1.15), the spatial integral of a divergence can be written in material coordinates as

$$\int_{\mathscr{P}_t} \mathrm{div}\ \boldsymbol{\phi}\,\mathrm{d}\mathbf{x} = \int_{\mathscr{P}} (\mathrm{div}\ \boldsymbol{\phi})_m J\,\mathrm{d}\mathbf{X} = \int_{\mathscr{P}} \mathrm{tr}(\mathrm{grad}\ \boldsymbol{\phi})_m J\,\mathrm{d}\mathbf{X}. \tag{1.20}$$

Inserting (1.17) in the previous equation and using the property (1.19), we get

$$\int_{\mathscr{P}} \mathrm{tr}(\mathrm{grad}\ \boldsymbol{\phi})_m J\,\mathrm{d}\mathbf{X} = \int_{\mathscr{P}} \mathrm{tr}\ (\nabla\boldsymbol{\phi}_m\mathbf{F}^{-1})J\,\mathrm{d}\mathbf{X} = \int_{\mathscr{P}} \nabla\boldsymbol{\phi}_m : \mathbf{F}^{-T}J\,\mathrm{d}\mathbf{X}.$$

The Piola identity allows to add to the previous relation the vanishing term

$$\mathrm{Div}\ (J\mathbf{F}^{-T}) \cdot \boldsymbol{\phi}_m,$$

and using the rule for derivation of a product we obtain

$$\int_{\mathscr{P}} \nabla\boldsymbol{\phi}_m : \mathbf{F}^{-T}J\,\mathrm{d}\mathbf{X} + \int_{\mathscr{P}} \mathrm{Div}\ (J\mathbf{F}^{-T}) \cdot \boldsymbol{\phi}_m\,\mathrm{d}\mathbf{X} = \int_{\mathscr{P}} \mathrm{Div}(J\mathbf{F}^{-1}\boldsymbol{\phi}_m)\,\mathrm{d}\mathbf{X}. \tag{1.21}$$

The comparison between equations (1.20) and (1.21) gives the Piola transformation

$$(\mathrm{div}\ \boldsymbol{\phi})_m J = \mathrm{Div}(J\mathbf{F}^{-1}\boldsymbol{\phi}_m). \tag{1.22}$$

Piola transformation of the divergence of a tensor function

Let us now consider the divergence of $\mathbf{T} \in \mathrm{Lin}_2$. Since the divergence of a second order tensor is applied to the rows of the associated linear transformation, relation (1.22) can be modified taking the transpose of the argument

$$(\mathrm{div}\ \mathbf{T})_m J = \mathrm{Div}(J\mathbf{T}_m\mathbf{F}^{-T}). \tag{1.23}$$

1.1.4 Balance of mass

The mass of a body, or of a given subpart of it, is computed through a density function [MH94]. This is a strictly positive function. We use the symbols $\rho_r(\mathbf{X}, t)$ to indicate the mass density in material coordinates and $\rho(\mathbf{x}, t)$ to indicate the mass density in spatial coordinates.

Material and spatial densities are related by the fact that the mass of each subpart of the body has to be same in both descriptions at each time. Let $\mathscr{P} \subset \mathscr{B}_r$ be a subpart of the body and let $\mathscr{P}_t = \boldsymbol{\chi}(\mathscr{P}, t)$ be the image of $\mathscr{P}_t$ through the motion, we impose the condition of equality of the mass with respect to both the configurations by means of the following equality

$$\int_{\mathscr{P}} \rho_r \mathrm{d}\boldsymbol{X} = \int_{\mathscr{P}_t} \rho \mathrm{d}\boldsymbol{x}. \tag{1.24}$$

Changing the variable of integration in the right hand side of (1.24), we obtain

$$\int_{\mathscr{P}} \rho_r \mathrm{d}\boldsymbol{X} = \int_{\mathscr{P}} \rho_m J \mathrm{d}\boldsymbol{X}, \tag{1.25}$$

The arbitrariness of $\mathscr{P}$ allows to convert the integral relation (1.25) into the pointwise equality

$$\rho_r(\mathbf{X}, t) = \rho_m(\mathbf{X}, t) J(\mathbf{X}, t). \tag{1.26}$$

Equation (1.26) is the Piola transformation of a scalar field.

The density ρ obeys the mass conservation law. This law states that the mass of each subpart $\mathscr{P}_t \subset \mathscr{B}_t$ is constant during the motion [MH94]. Hence the following equation has to hold during the motion

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathscr{P}_t} \rho \mathrm{d}\boldsymbol{x} = 0. \tag{1.27}$$

Writing the integral in (1.27) in material coordinates, we obtain an integral that does not depends on time. Hence, we can exchange time-derivative and space-integral; exploiting again the arbitrariness of $\mathscr{P}$, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}(\rho_m J) = 0. \tag{1.28}$$

Since $\rho_r = \rho_m J$, equation (1.28) states that density in material coordinates cannot depend on time. Once a given initial density ρ_0 and an initial deformation $\boldsymbol{\chi}_0$ are given, the material density can be easily computed by $\rho_r = \rho_{0m} J_0$ and the evolution of ρ can be obtained by

$$\rho = \left(\frac{\rho_r}{J}\right)_s.$$

The time derivative in (1.28) can be computed using (1.11). This leads to

$$(\rho')_m J + (\mathrm{grad}\rho)_m \dot{\boldsymbol{x}} J + \rho_m \dot{J} = 0. \tag{1.29}$$

Employing the definition of spatial gradient (1.10) into (1.7), we obtain the following relation for the time derivative of the deformation determinant

$$\dot{J} = J\mathrm{tr}(\dot{\boldsymbol{F}}\boldsymbol{F}^{-1}) = J(\mathrm{div}\mathbf{v})_m. \tag{1.30}$$

Using this latter equality in (1.29) and dividing by J, the conservation of mass can be written as

$$(\rho')_m + (\mathrm{grad}\rho)_m\dot{\mathbf{x}} + \rho_m(\mathrm{div}\mathbf{v})_m = 0. \tag{1.31}$$

Equation (1.31) can be written in spatial coordinates in two possible ways:

$$\rho' + \mathrm{div}(\rho\mathbf{v}) = 0 \tag{1.32}$$

or

$$D_t\rho + \rho\,\mathrm{div}(\mathbf{v}) = 0 \tag{1.33}$$

Incompressibility

For our purpose we consider a body incompressible if its spatial density is constant. This means that equation (1.32) becomes the constraint

$$\mathrm{div}(\mathbf{v}) = 0$$

or, in material coordinates, equation (1.28) becomes

$$\dot{J} = 0.$$

1.1.5 Conservation of linear and angular momentum

In Newtonian mechanics, conservation of linear momentum and angular momentum are the basic postulates. Such equations are written for a system of particles. In continuum mechanics we assume that these two conservation principles should hold true for each subpart $\mathscr{P}_t$ of the body. The existence of a vector field $\mathbf{t}(\mathbf{x}, t, \mathbf{n})$ is postulated which represents the force per unit area exerted on the surface of $\mathscr{P}_t$ oriented with normal $\mathbf{n}$ at $\mathbf{x}$ by the rest of the the body [MH94]. The conservation of linear momentum reads

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathscr{P}_t} \rho\mathbf{v}\mathrm{d}\mathbf{x} = \int_{\partial\mathscr{P}_t} \mathbf{t}(\mathbf{x}, t, \mathbf{n})\mathrm{d}\mathbf{a} + \int_{\mathscr{P}_t} \mathbf{b}\mathrm{d}\mathbf{x} \tag{1.34}$$

i.e. the rate of change of the linear momentum of $\mathscr{P}_t$ equals the force acting on it. In the previous equation, $\mathbf{b}$ represents the contribution of the external volumetric forces. Cauchy proved that the dependence of $\mathbf{t}$ on $\mathbf{n}$ is linear, and hence the stress can be written as $\mathbf{t}(\mathbf{x}, t, \mathbf{n}) = \mathbf{T}(\mathbf{x}, t)\mathbf{n}$, where $\mathbf{T}$ is the so-called Cauchy stress tensor. The linear

dependence on the normal allows to apply the divergence theorem to the right hand side of (1.34) to obtain

$$\int_{\partial \mathscr{P}_t} \mathbf{T}\mathbf{n}\,\mathrm{d}\mathbf{a} = \int_{\mathscr{P}_t} \mathrm{div}\,\mathbf{T}\mathrm{d}\mathbf{x}. \tag{1.35}$$

The time derivative of the linear momentum in(1.34) can be written in two possible ways. The first one can be derived exploiting the time-independence of ρ_r from the following expression:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathscr{P}_t} \rho\mathbf{v}\mathrm{d}\mathbf{x} = \frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathscr{P}} \rho_m\dot{\mathbf{x}}J\,\mathrm{d}\mathbf{X} = \int_{\mathscr{P}} \frac{\mathrm{d}}{\mathrm{d}t}(\rho_r\dot{\mathbf{x}})\mathrm{d}\mathbf{X} = \int_{\mathscr{P}} \rho_r\ddot{\mathbf{x}}\mathrm{d}\mathbf{X} = \int_{\mathscr{P}_t} \rho\mathbf{a}\mathrm{d}\mathbf{x} \tag{1.36}$$

Instead, the second approach is based on the convective derivative of ρ and can be obtained from the following identities:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathscr{P}_t} \rho\mathbf{v}\mathrm{d}\mathbf{x} = \int_{\mathscr{P}} \frac{\mathrm{d}}{\mathrm{d}t}(\rho_m\dot{\mathbf{x}}J)\mathrm{d}\mathbf{X} \tag{1.37}$$

$$= \int_{\mathscr{P}} (\rho')_m\dot{\mathbf{x}}J + (\mathrm{grad}\,\rho)_m\cdot\dot{\mathbf{x}}J + \rho_m\dot{\mathbf{x}}(\mathrm{div}\,\mathbf{v})_m J + \rho_m\ddot{\mathbf{x}}J\,\mathrm{d}\mathbf{X} \tag{1.38}$$

$$= \int_{\mathscr{P}_t} \rho'\mathbf{v} + \mathrm{grad}\,\rho\cdot\mathbf{v} + \rho\,\mathbf{v}\mathrm{div}\,\mathbf{v} + \rho\mathbf{a}\,\mathrm{d}\mathbf{x} \tag{1.39}$$

In (1.38), we used the derivative of the deformation determinant (1.30). Finally, the convective derivative of the velocity (1.14) can be put into (1.39) which provides the following expression for the derivative of the linear momentum

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathscr{P}_t} \rho\mathbf{v}\mathrm{d}\mathbf{x} = \int_{\mathscr{P}_t} (\rho\mathbf{v})' + \mathrm{div}\,(\rho\mathbf{v}\otimes\mathbf{v})\mathrm{d}\mathbf{x} \tag{1.40}$$

The conservation of linear momentum (1.34) can be reformulated as a differential equation using the Green's theorem (1.35) and the formulae (1.36) or (1.40) for the variation of linear momentum. These two different choices give two possible differential formulations that are respectively:

$$\rho\mathbf{a} = \mathrm{div}\,\mathbf{T} + \mathbf{b} \tag{1.41}$$

or

$$(\rho\mathbf{v})' + \mathrm{div}\,(\rho\mathbf{v}\otimes\mathbf{v}) = \mathrm{div}\,\mathbf{T} + \mathbf{b}. \tag{1.42}$$

The first approach is usually employed in solid mechanics while the second in fluid mechanics.

The conservation of angular momentum for a subpart $\mathscr{P}_t$ of a non polar body reads:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathscr{P}_t} \mathbf{x}\times\rho\mathbf{v}\mathrm{d}\mathbf{x} = \int_{\partial\mathscr{P}_t} \mathbf{x}\times\mathbf{T}\mathbf{n}\mathrm{d}\mathbf{a} + \int_{\mathscr{P}_t} \mathbf{x}\times\mathbf{b}\mathrm{d}\mathbf{x}.$$

From this equation, by means of differential and algebraic relations, we obtain that the Cauchy stress tensor is symmetric.

Balance of momentum in material coordinates

The conservation of linear momentum can be written in material coordinates applying the operator $(\cdot)_m$ to equation (1.41) and multiplying by J. This leads to the following equation

$$\rho_m\mathbf{a}_m J = (\mathrm{div}\,\mathbf{T})_m J + \mathbf{b}_m J. \tag{1.43}$$

The left hand side can be rewritten as $\rho_r\ddot{\mathbf{x}}$ (cf. Eq. (1.36)). We also introduce the volumetric forces with respect to the reference configuration as $\mathbf{b}_r = \mathbf{b}_m J$. The first term of the right hand side of (1.43) can be written using the Piola transformation (1.23). The tensor $J\mathbf{T}_m\mathbf{F}^{-T}$ is known as the first Piola-Kirchhoff tensor and it is denoted by $\mathbf{P}$. Thanks to these quantities defined in the reference configuration, Eq. (1.43) can be written as

$$\rho_r\ddot{\mathbf{x}} = \mathrm{Div}\,\mathbf{P} + \mathbf{b}_r. \tag{1.44}$$

The tensor $\mathbf{P}$ is not symmetric and measures the stress acting on the boundary of $\mathscr{P}_t$ of the current configuration at time t with respect to the reference configuration. The first Piola-Kirchhoff tensor can be also written as $\mathbf{P} = \mathbf{F}\mathbf{S}$, where the tensor $\mathbf{S} = J\mathbf{F}^{-1}\mathbf{T}_m\mathbf{F}^{-T}$. $\mathbf{S}$ is known as the second Piola-Kirchhoff tensor and it is symmetric.

The tensor $\mathbf{P}$, and hence $\mathbf{S}$, accounts for the constitutive behavior of the material of the body. They depend only on $\mathbf{F}$ and $\mathbf{X}$. If the stress tensors do not depend on the material coordinates, the material is said to be *uniform*.

A material is said to be *hyperelastic*, if there exists an Helmholtz energy function $\Psi = \Psi(\mathbf{F},\mathbf{X})$ for which

$$\mathbf{P} = \frac{\partial\Psi}{\partial\mathbf{F}}.$$

If Ψ depends only on $\mathbf{F}$, it is called strain-energy function. In this case, the study of objectivity imposes that the energy function may depend on $\mathbf{F}$ only through the Green tensor $\mathbf{C}$. We hence define the energy $\mathscr{W} = \mathscr{W}(\mathbf{C}) = \mathscr{W}(\mathbf{F}^T\mathbf{F}) = \Psi(\mathbf{F})$. According to this last definition, we can write

$$\mathbf{P} = \frac{\partial\mathscr{W}}{\partial\mathbf{F}} \quad \text{and} \quad \mathbf{S} = 2\frac{\partial\mathscr{W}}{\partial\mathbf{C}}. \tag{1.45}$$

Equation (1.44) is a partial differential equation of second order in time and it is non linear in the material coordinates. In order to have to a well defined differential problem, it has to be completed with two initial conditions: the initial position and the initial velocity. Moreover, at each time for each point of the boundary $\partial\mathscr{P}_r$, a boundary condition has to be imposed. The boundary is splitted into two subset, on which Dirichlet or Neumann conditions are imposed. Dirichlet condition means that the boundary is clamped or a fixed displacement is set and Neumann condition means that a boundary stress is set. If external forces are negligible, $\mathbf{b}_r$ can be set to zero. If inertial forces are negligible, $\rho_r\ddot{\mathbf{x}}$ can be omitted so that the problem reads

$$-\operatorname{Div}\mathbf{P} = \mathbf{f}. \tag{1.46}$$

This problem can still be time-dependent because the stress tensor can be a function of $\dot{\mathbf{u}}$ as in case of viscoelastic materials. Moreover boundary conditions can also depend on time.

1.1.6 Constitutive equations

The goal of constitutive theory is to create mathematical models that describe the physical behavior of materials. They allow to relate the kinematics to the continuity equations. Details on constitutive models can be found in several books, i.e. [Bon08, Hol01]. Here, we will just report two energies that we will use in the upcoming chapters of this thesis. The models that we will consider are all hyperelastic, uniform, and isotropic.

The neo-Hookean material law is characterized by the following energy [Bon08]

$$\mathscr{W}(\mathbf{C}) = \frac{\mu}{2}\left[\operatorname{tr}(\mathbf{C}) - 3\right] - \mu\ln(J) + \frac{\lambda}{2}\left[\ln(J)\right]^2 \tag{1.47}$$

where $J = \det(\mathbf{C})^{1/2}$. The material parameters μ and λ are the shear modulus and the second Lamè parameter, respectively. Eq. (1.45) lead to the following expressions of $\mathbf{S}$

$$\mathbf{S}(\mathbf{C}) = 2\frac{\partial\mathscr{W}}{\partial\mathbf{C}} = \mu[\mathbf{I} - \mathbf{C}^{-1}] + \lambda[\ln(J)]\mathbf{C}^{-1}. \tag{1.48}$$

and $\mathbf{P}$

$$\mathbf{P}(\mathbf{F}) = \frac{\partial\mathscr{W}}{\partial\mathbf{F}} = \mu[\mathbf{F} - \mathbf{F}^{-T}] + \lambda[\ln(J)]\mathbf{F}^{-T}. \tag{1.49}$$

In (1.48), the identity tensor $\mathbf{I}$ should be substituted with the inverse $\mathbf{g}^{-1}$ of the metric tensor if generalized coordinated are employed. Notice that $\mathbf{P} = \mathbf{FS}$.

A different material law is the so called Fung potential [Fun90a]. This constitutive equation is widely used in biomechanics and it is characterized by the following energy

$$\mathscr{W}(\mathbf{C}) = \alpha\exp\left[\frac{1}{\alpha}\left(\mu\mathbf{E}:\mathbf{E} + \frac{\lambda}{2}\operatorname{tr}(\mathbf{E})^2\right)\right] \tag{1.50}$$

where

$$\mathbf{E} = \frac{\mathbf{C} - \mathbf{I}}{2}$$

is the Lagrange strain tensor. The parameter α is a further parameter that measures the stiffening of the material. The corresponding second Piola-Kirchhoff tensor is given by

$$\mathbf{S}(\mathbf{E}) = 2\frac{\partial \mathscr{W}}{\partial \mathbf{C}} = \left[2\mu\mathbf{E} + \lambda \mathrm{tr}(\mathbf{E})\mathbf{I}\right] \frac{\mathscr{W}}{\alpha}. \tag{1.51}$$

For $\alpha \to 0$, the previous energy tends to the De Saint Venant-Kirchhoff model. This is characterised by a linear stress-strain relation but includes geometrical non-linearities.

1.2 Poroelasticity

Poroelasticity is a branch of continuum mechanics that studies the behaviour of deformable porous media filled-up with a fluid. A porous material, or medium, is a solid, called matrix, permeated by an interconnected network of pores (voids). If the pores are completely filled with the fluid, the porous media is said to be saturated.

Historically, two different models developed to describe elastic porous materials, namely the Biot Theory (BT) and the approach based on Mixture Theory (MT). In both these approaches the level of investigation is so coarse that there is no geometrical resolution of the porous structure.

Biot [Bio40] proposed the first three-dimensional model to describe poroelasticity. BT was presented as an extension of the mono dimensional study [vT23] realized by von Terzaghi. In this approach, there is no distinction between the two phases but a small cube of a biphasic material is considered as single material. The effect of the fluid was taken into account by means of a macroscopic variable, i.e. the porosity. This additional variable measures the density of the voids. To obtain its distribution, a further continuity equation has to be solved.

In MT, we consider a biphasic mixture consisting of a solid and a fluid phase. The solid phase represents the porous medium, whereas the fluid phase consists of a fluid that completely saturates the pores of the solid and may move throughout it. This model has a more symmetric approach from the perspective of the constituents [PDIS10]. In fact, continuity equations are solved for both the phases. The MT requires the introduction of some closure relations such as Darcy's and Fick's laws and hence the introduction of further parameters, i.e. permeability and diffusivity constants.

In the linearized case, the two models give rise to the same system of equations whose unknowns are the solid displacements and the pore pressure, but the physical coefficients possess different interpretations [Sch03].

The early works where poroelasticity was successfully employed aimed to study consolidation of soils and reservoirs in geomechanics, but a system of this type may also represent the most essential model of those soft biological tissues whose main constituents are a porous solid matrix and an interstitial fluid. An example of such tissues

is articular cartilage, in which the solid consists of cells (chondrocytes), collagen fibres and a matrix of proteoglycans, and the fluid comprises water, ions and various chemical compounds. The latter ones can be either nutrients for the cells or byproducts of the cellular metabolic activity and are known as *constituents of the fluid phase*. From here on, we refer to the fluid phase as the system made of water and all the constituents, but for example in [GAG+03] and [GFW12], the influence and the interactions of the constituents were considered. In the former conservation equations in the form of diffusion-convection equations were introduced while in the second continuity equations are solved for all the constituents. In the bio-mechanical field, poroelasticity has been used, for example, for the modeling of cartilage [FGLR+05], teeth [FGD+11], and heart walls [HvCAH91].

1.2.1 Volume fractions

According to MT, the two continua (the solid and the fluid) are interpenetrable and occupy the same region of space at each time t. For each phase, modified constitutive relations are obtained by averaging equations on a smaller scale. The averaging procedures are often based on volume- and mass-average methods [HG89, BMC00]. These procedures assume the existence of a Representative Elementary Volume (REV), $\Omega(\mathbf{x}) \subset \mathbb{R}^3$, which supplies information about the composition and structure of the mixture at the point $\mathbf{x} \in \mathbb{R}^3$. Our level of investigation is so coarse that a spatial point has to be seen as a REV. The characteristic size of the REV depends on the system under investigation. Two examples of porous materials are reported in Fig. 1.4.

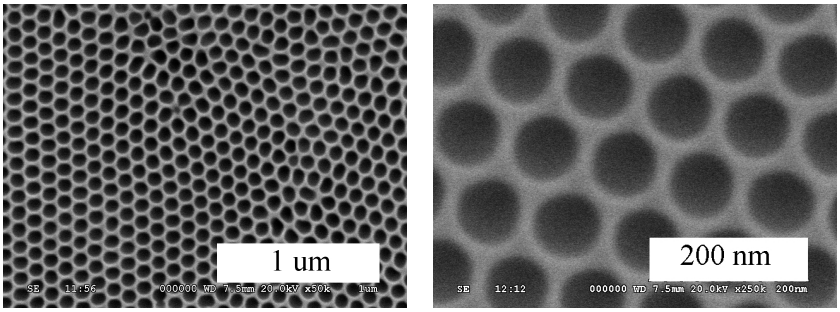


Figure 1.4. Two examples of porous material: sections of porous alumina at different scales. Images from Prof. Hongtao Shi [Hon].

The subvolume $\Omega_\alpha(\mathbf{x}, t)$, where the subscript α indicates the α-constituent, represents the subset of $\Omega(\mathbf{x})$ occupied by the α-th phase at time t. The measure of Ω_α is defined as

$$|\Omega_\alpha(\mathbf{x},t)| = \int_{\Omega(\mathbf{x})} \xi_\alpha(\mathbf{z},t)\,\mathrm{d}\mathbf{z} \tag{1.52}$$

where ξ^α is the characteristic function of the α constituent

$$\xi(\mathbf{z},t) = \begin{cases} 1 & \text{if } \mathbf{z} \in \Omega_\alpha(\mathbf{x},t) \\ 0 & \text{otherwise} \end{cases} .$$

At a macroscopic level, the local composition is described by the volume fractions $\phi_\alpha(\mathbf{x},t) = |\Omega_\alpha(\mathbf{x},t)|/|\Omega(\mathbf{x},t)|$. The abbreviations $\mathscr{P}_s$ and $\mathscr{P}_\ell$ will be sometimes used to denote the solid and the fluid phase, respectively.

If the porous medium is saturated, the condition $\phi_s + \phi_\ell = 1$ applies at all times and all spatial points. The distribution of mass of $\mathscr{P}_\alpha$ in Ω_α is the "true" mass density of $\mathscr{P}_\alpha$, and is denoted here by $\hat{\rho}_\alpha$. The product

$$\rho_\alpha = \phi_\alpha \hat{\rho}_\alpha \tag{1.53}$$

measures the distribution of mass of $\mathscr{P}_\alpha$ in Ω, and is called *apparent* mass density of $\mathscr{P}_\alpha$. The derivation of the poroelasticity system of equations is usually done assuming

C1: that the constituents are incompressible, i.e. constant true densities $\hat{\rho}_\alpha$;

C2: quasi-static conditions: inertia terms are negligible.

For the presentation of the balance laws of a poroelastic material we will use the same notation for motions and tensors employed in the previous section adding a the subscript $_s$ or $_\ell$ to refer to the solid and fluid phase respectively.

1.2.2 Kinematics

In order to describe the kinematics of a poroelastic material, we introduce two different body manifolds: $\mathscr{B}_s$ for the solid phase and $\mathscr{B}_\ell$ for the fluid phase. The former consists of solid particles P_s and the latter of the fluid particles P_ℓ. As done in the previous section, generalized coordinates $\mathbf{X}_s$ are assigned to solid particles P_s by means of a bijective mapping $\boldsymbol{\varkappa}_s$. The reference configuration $\mathscr{B}_r$ is the image $\boldsymbol{\varkappa}(\mathscr{B}_s)$. The motion of the solid phase, also called s-motion, is a time-dependent sequence of deformations defined as

$$\mathscr{B}_r \times \mathscr{I} \ni (\mathbf{X}_s, t) \longmapsto \mathbf{x} = \boldsymbol{\chi}_s(\mathbf{X}_s, t) \in \mathbb{R}^3.$$

For $\mathscr{P}_s$, we can define the deformation gradient $\mathbf{F}_s$, the stretch tensor $\mathbf{C}_s$, and the deformation determinant J_s. The first time derivative $\dot{\boldsymbol{\chi}}_s(\mathbf{X}_s, t)$ and the second time derivative $\ddot{\boldsymbol{\chi}}_s(\mathbf{X}_s, t)$ are respectively the velocity and the acceleration of the solid particle with material coordinates $\mathbf{X}_s$ at time t.

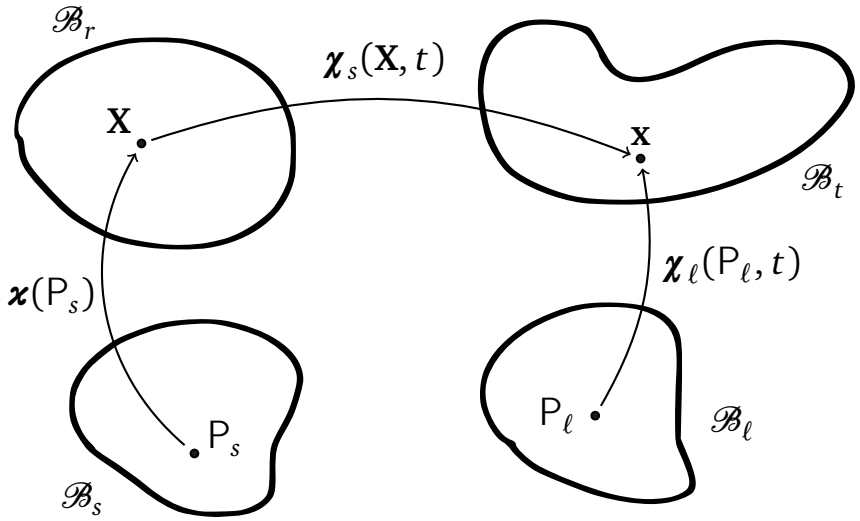


Figure 1.5. Representation of motions for a poroelastic material: solid particles are embedded into $\mathbb{R}^3$ and s-motion is defined in the reference configuration; ℓ-motion embeds fluid particles in $\mathbb{R}^3$. s-motion and ℓ-motion have the same image.

As customary in fluid mechanics no reference configuration is needed for $\mathscr{P}_\ell$. Particle P_ℓ is placed in the Euclidean space by means of a mapping that locates the particle at place $\mathbf{x}$. A smooth motion of $\mathscr{P}_\ell$, the ℓ-motion, is a sequence of mappings defined by

$$\mathscr{B}_\ell \times \mathscr{I} \ni (\mathsf{X}_\ell, t) \longmapsto \mathbf{x} = \boldsymbol{\chi}_\ell(\mathsf{X}_\ell, t) \in \mathbb{R}^3.$$

The motions are sketched in Fig. 1.5. The portion of $\mathbb{R}^3$ occupied by the mixture at time t is the current configuration and it is given by the intersection $\mathscr{B}_t = \boldsymbol{\chi}_s(\mathscr{B}_r, t) \cap \boldsymbol{\chi}_\ell(\mathscr{B}_\ell, t)$.

The velocities $\mathbf{v}_s$ and $\mathbf{v}_\ell$ characterize the standard motion of $\mathscr{P}_s$ and $\mathscr{P}_\ell$. Among the various ways of describing the motion of the phases $\mathscr{P}_s$ and $\mathscr{P}_\ell$, we choose the one based on the set of *standard* velocities

$$\mathscr{V}_{\mathrm{st}} := \{\mathbf{v}_s, \mathbf{w}_{\ell s}\}, \tag{1.54}$$

where $\mathbf{w}_{\ell s} := \mathbf{v}_\ell - \mathbf{v}_s$ is the relative velocity of $\mathscr{P}_\ell$ with respect to $\mathscr{P}_s$. $\mathbf{w}_{\ell s}$ is also known as seepage velocity [EM01]. Another relevant kinematic quantity is the velocity of the centre of mass of the mixture, which is defined by

$$\mathbf{v} := \frac{\rho_s \mathbf{v}_s + \rho_\ell \mathbf{v}_\ell}{\rho}, \quad \rho := \rho_\ell + \rho_s. \tag{1.55}$$

Here, ρ denotes the mass density of the mixture as a whole. The relative velocities $\tilde{\mathbf{v}}_\ell := \mathbf{v}_\ell - \mathbf{v}$ and $\tilde{\mathbf{v}}_s := \mathbf{v}_s - \mathbf{v}$ describe the relative motion of each phase with respect to the motion of the mixture as a whole. For each phase, we denote by $\mathbf{a}_\alpha$ (with $\alpha = \ell, s$) the spatial acceleration of that phase, which is defined by the convective derivative of $\mathbf{v}_\alpha$ with respect to the motion of the phase $\mathscr{P}_\alpha$, i.e.

$$\mathbf{a}_\alpha := \mathbf{v}'_\alpha + \mathrm{grad}(\mathbf{v}_\alpha)\mathbf{v}_\alpha, \quad \alpha = \ell, s. \tag{1.56}$$

The acceleration of the mixture is indicated by $\mathbf{a}$ and is defined as the convective derivative of $\mathbf{v}$ with respect to the motion of the mixture as a whole, i.e.

$$\mathbf{a} := \mathbf{v}' + \mathrm{grad}(\mathbf{v})\mathbf{v}. \tag{1.57}$$

In order to make the Eulerian description consistent with the Lagrangian one, we need to relate the velocities $\mathbf{v}_s$ and $\mathbf{v}_\ell$ with the motions of the solid and fluid phases. To this purpose the following relations has to hold true:

$$\dot{\boldsymbol{\chi}}_s(\mathbf{X_s}, t) = \mathbf{v}_s(\boldsymbol{\chi}_s(\mathbf{X}_s, t), t) \tag{1.58}$$

and

$$\dot{\boldsymbol{\chi}}_\ell(\mathsf{X}_\ell, t) = \mathbf{v}_\ell(\boldsymbol{\chi}_\ell(\mathsf{X}_\ell, t), t). \tag{1.59}$$

They mean that two different velocities coexist at the each point $\mathbf{x}$ of the Euclidean space. Eq. (1.58) is a change of coordinates from the spatial to the material description. On the other hand, equation (1.59) is not a change of description but, in general, a quantity associated with the fluid phase, e.g., the velocity $\mathbf{v}_\ell$, can be expressed in terms of the material coordinates by using the following composition of maps:

$$(\mathbf{v}_\ell)_m(\mathbf{X}_s, t) = \mathbf{v}_\ell(\boldsymbol{\chi}_s(\mathbf{X}_s, t), t) = [\mathbf{v}_\ell(\,\cdot\,, t) \circ \boldsymbol{\chi}_s(\,\cdot\,, t)](\mathbf{X}_s). \tag{1.60}$$

1.2.3 Balance of mass

Since in a poroelastic material two different phases are present, conservation of mass has to be written for each phase and reads

$$\rho'_\alpha + \mathrm{div}(\rho_\alpha \mathbf{v}_\alpha) = 0, \tag{1.61}$$

with $\alpha \in \{s, \ell\}$. Substituting relation (1.53) into (1.61) and assuming incompressibility of both phases, equation (1.61) takes the form

$$\phi'_\alpha + \mathrm{div}(\phi_\alpha \mathbf{v}_\alpha) = 0, \tag{1.62}$$

that is the conservation law for volume fractions. Using the standard velocities (1.54), conservation of the fluid volume fraction reads

$$\phi_\ell' + \mathrm{div}(\phi_\ell \mathbf{w}_{\ell s}) + \mathrm{div}(\phi_\ell \mathbf{v}_s) = 0. \tag{1.63}$$

Conservation of the solid volume fraction can also be written using the total derivative (cf. Eq. (1.29)) and it takes the following form

$$D_s \phi_s + \phi_s \mathrm{div}(\mathbf{v}_s) = 0. \tag{1.64}$$

The subscript $_s$ in the operator D_s is the convective derivative computed with respect to the solid motion. The computation convective derivative consists in writing ϕ_s in material coordinates, computing the total time derivative, and then transforming again to spatial coordinates. Applying the material operator $(\cdot)$ to (1.64) and multiplying by J_s, conservation of mass can be written as

$$J \frac{\mathrm{d}}{\mathrm{d}t} \phi_s(\boldsymbol{\chi}_s(\mathbf{X}, t), t) + J \phi_s(\boldsymbol{\chi}_s(\mathbf{X}, t), t)(\mathrm{div}(\mathbf{v}_s))_m = 0. \tag{1.65}$$

Exploiting identity (1.30), the former equation can be simplified as

$$\frac{\mathrm{d}}{\mathrm{d}t}(J(\phi_s)_m) = 0. \tag{1.66}$$

This means that the solid volume fraction in the reference configuration

$$\Phi_s = J(\phi_s)_m \tag{1.67}$$

cannot depend on time. If an initial deformation $\boldsymbol{\chi}_{s0}(\mathbf{X})$ and an initial volume fraction $\phi_{s0}(\mathbf{x})$ are given, the reference volume fraction can be obtained by

$$\Phi_s(\mathbf{X}) = J_{s0}(\phi_{s0})_m$$

and it remains constant during the motion. If we also assume that the reference configuration is the initial configuration, i.e. $\mathbf{X} = \boldsymbol{\chi}_s(\mathbf{X}, 0)$, then the following condition holds true

$$\Phi_s = \phi_{s0}.$$

In order to obtain a simpler expression for the balance to the fluid volume fraction, we sum up Eq. (1.63) and Eq. (1.62) with $\alpha = s$ to get the following identity:

$$\mathrm{div}(\phi_\ell \mathbf{w}_{\ell s}) + \mathrm{div}(\mathbf{v}_s) = 0. \tag{1.68}$$

In order to write a conservation law in material coordinates, we write the last equation in material coordinates and we multiply by J_s. Using the kinematical identity (1.30), we obtain an evolution equation for the deformation determinant that has the following form

$$-\dot{J}_s = J_s(\mathrm{div}(\phi_\ell \mathbf{w}_{\ell s}))_m. \tag{1.69}$$

Finally, by means of the Piola transform, Eq. (1.69) can be written as

$$-\dot{J}_s = \mathrm{Div}(J_s \phi_{\ell m} \mathbf{F}^{-1} \mathbf{w}_{\ell s}). \tag{1.70}$$

In the previous equation, we used the same symbol for the velocity $\mathbf{w}_{\ell s}$ written in material and spatial coordinates. Finally, multiplying the saturation condition by J_s and using the solution (1.67) of the conservation of the solid volume fraction we obtain

$$(\phi_\ell)_m J_s = J_s - \Phi_s.$$

This last equality can be substituted into (1.70) to obtain

$$-\dot{J}_s = \mathrm{Div}((J_s - \Phi_s)\mathbf{F}^{-1} \mathbf{w}_{\ell s}). \tag{1.71}$$

This last equation is a conservation law written in material coordinates that uses only quantities related to the solid volume fraction.

1.2.4 Conservation of linear momentum

Together with mass balance, also the balance of momentum of the phases $\mathscr{P}_s$ and $\mathscr{P}_\ell$ has to be studied. The Eulerian, local form of these balance laws is given by

$$(\rho_s \mathbf{v}_s)' + \mathrm{div}(\rho_s \mathbf{v}_s \otimes \mathbf{v}_s) - \mathrm{div}(\mathbf{T}_s) = \rho_s \mathbf{m}_s, \tag{1.72}$$

$$(\rho_\ell \mathbf{v}_\ell)' + \mathrm{div}(\rho_\ell \mathbf{v}_\ell \otimes \mathbf{v}_\ell) - \mathrm{div}(\mathbf{T}_\ell) = \rho_\ell \mathbf{m}_\ell, \tag{1.73}$$

where $\mathbf{T}_s$ and $\mathbf{T}_\ell$ denote the Cauchy stress tensors of the solid and fluid phase, $\rho_s \mathbf{m}_s$ and $\rho_\ell \mathbf{m}_\ell$ are the rate of exchange of momentum between the two phases. The forces $\rho_s \mathbf{m}_s$ and $\rho_\ell \mathbf{m}_\ell$ satisfy the condition

$$\rho_s \mathbf{m}_s + \rho_\ell \mathbf{m}_\ell = \mathbf{0}, \tag{1.74}$$

which states that the mixture is closed with respect to momentum. An explanation of the physical meaning of (1.74) in terms of pore scale considerations can be found, for example, in [HG89]. The use of (1.74), and the definition of the relative velocities $\tilde{\mathbf{v}}_\ell$ and $\tilde{\mathbf{v}}_s$ as well as of the accelerations $\mathbf{a}$ and $\mathbf{a}_\ell$ allow for reformulating the balance laws (1.72) and (1.73) in the following way

$$\rho \mathbf{a} - \mathrm{div}(\mathbf{T}) = \rho_s \mathbf{p}_s, \tag{1.75}$$

$$\rho_\ell \mathbf{a}_\ell - \mathrm{div}(\mathbf{T}_\ell) = \rho_\ell \mathbf{m}_\ell, \tag{1.76}$$

where $\mathbf{T}$, which denotes the Cauchy stress tensor of the mixture, is defined by

$$\mathbf{T} := \mathbf{T}_s + \mathbf{T}_\ell - \left[\rho_s \tilde{\mathbf{v}}_s \otimes \tilde{\mathbf{v}}_s + \rho_\ell \tilde{\mathbf{v}}_\ell \otimes \tilde{\mathbf{v}}_\ell\right]. \tag{1.77}$$

Equation (1.75) represents the balance of momentum of the mixture as a whole, and is obtained by adding together (1.72) and (1.73) and applying the definitions (1.55)–(1.57) to the result. Furthermore, substituting the identity

$$\rho \mathbf{a} := \sum_{\alpha=\ell,s} \rho_\alpha \mathbf{a}_\alpha - \sum_{\alpha=\ell,s} \mathrm{div}(\rho_\alpha \tilde{\mathbf{v}}_\alpha \otimes \tilde{\mathbf{v}}_\alpha) \tag{1.78}$$

into (2.49) yields

$$\rho_s \mathbf{a}_s + \rho_\ell \mathbf{a}_\ell - \mathrm{div}(\mathbf{T}_s + \mathbf{T}_\ell) = \mathbf{0}, \tag{1.79}$$

$$\rho_\ell \mathbf{a}_\ell - \mathrm{div}(\mathbf{T}_\ell) = \rho_\ell \mathbf{m}_\ell. \tag{1.80}$$

Finally, neglecting the inertial forces of both phases, the balances of momentum (1.72) and (1.73) become

$$-\mathrm{div}(\mathbf{T}_s + \mathbf{T}_\ell) = \mathbf{0}, \tag{1.81}$$

$$-\mathrm{div}(\mathbf{T}_\ell) = \rho_\ell \mathbf{m}_\ell. \tag{1.82}$$

By means of the Piola transformations of (1.81) and (1.82), the momentum balance laws of the mixture can be written with respect to the reference placement $\mathscr{B}_r$, i.e.

$$-\mathrm{Div}(\mathbf{P}_s + \mathbf{P}_\ell) = \mathbf{0}, \tag{1.83}$$

$$-\mathrm{Div}(\mathbf{P}_\ell) = J \rho_\ell \mathbf{m}_\ell, \tag{1.84}$$

where

$$\mathbf{P}_s := J(\mathbf{T}_s)_m \mathbf{F}^{-T}, \quad \mathbf{P}_\ell := J(\mathbf{T}_\ell)_m \mathbf{F}^{-T} \tag{1.85}$$

denote, respectively, the first Piola-Kirchhoff stress tensors of the solid and fluid phase. In order to close the mathematical problem resulting from (1.71), (1.83) and (1.84), it is necessary to provide information about the stresses $\mathbf{P}_s$ and $\mathbf{P}_\ell$, and the force densities $\rho_\ell \mathbf{m}_\ell$ and $\rho_s \mathbf{m}_s$. These aspects will be discussed in detail in the next chapter when a more general model including growth will be derived.

Chapter 2

Mechanics of volumetric growth

The goal of this chapter is to derive a model for the transport of fluid in a deformable porous medium whose mass and internal structure may vary in time. Following the approach of the previous chapter, we will first show how a growth model modifies the balance equations for elastic continua and then the complete derivation will be done for biphasic continua. The content of this chapter has been published in the work of the author [GGF+12].

2.1 Introduction

The picture sketched in the previous chapter becomes more complex when the structural changes of the tissue are considered besides deformation. In solid mechanics, by *structural changes* we mean growth and re-organization, such as the one of cells and fibrils in living matter. In poroelasticity, we also consider mass exchange between the fluid and the solid phase. Even though these phenomena are all intermingled with each other, a conceptual classification can be found in the biomechanical literature (cf., for example, Fung [Fun90b] and Taber [Tab95]): processes involving mass addition or exchange are referred to as *growth*, while processes involving re-organization are referred to as *remodelling*. All these structural changes contribute also to modify the properties of the tissue (e.g., the stiffness, diffusivity and permeability).

From the point of view of Continuum Mechanics, the structural change of an elastic or poroelastic tissue is modeled by means of a class of second order tensors that describe how the material particles are distributed in the tissue. We denote this tensor by $\mathbf{F}_a$. With the terminology of [EM00], $\mathbf{F}_a$ measures the material inhomogeneities triggered by growth, mass exchange processes, and cellular re-organization. There is, thus, a strong conceptual difference between the standard deformations and those described by $\mathbf{F}_a$: whereas the former ones are related to the gradient of displacement of the body, the latter deformations need not be the gradient of any vector field. Rather, they are primitive entities that define, together with displacements, the parameters that are nec-

essary and sufficient for describing the kinematics of deformable bodies with variable internal structure.

In classical Continuum Mechanics[1], tensor $\mathbf{F}_a$ represents the inelastc part of the deformation gradient tensor, $\mathbf{F}_s$, which describes the overall change of shape of a solid. This inelastic deformation may be due to, for example, plastic deformations, thermal distortions, and damage [Mić09]. Tensor $\mathbf{F}_a$ individuates an evolving relaxed configuration of body elements. The accommodating deformation, which determines the actual configuration of the body from the relaxed one, is denoted by $\mathbf{F}_e$, and obeys the multiplicative decomposition $\mathbf{F}_s = \mathbf{F}_e\mathbf{F}_a$ [Krö59][Krö60][Lee69]. This multiplicative splitting is also called Bilby-Kröner-Lee decomposition. Usually, $\mathbf{F}_e$ is said to be the elastic part of $\mathbf{F}_s$.

Rodriguez et al. [RHD94] used the decomposition of $\mathbf{F}_s$ to study growth mechanics, and identified $\mathbf{F}_a$ with the deformations due to growth. When $\mathbf{F}_a$ is not the gradient of any vector field, it is said to be incompatible. Physically, this represents the case in which grown material points lose their geometric compatibility (this situation usually leads to residual stresses). The rate of inelastc deformation, $\mathbf{L}_a = \dot{\mathbf{F}}_a(\mathbf{F}_a)^{-1}$, is related to the variation of body mass in such a way that that the mass density of the body is constant when measured with respect to the relaxed configuration (cf., for example, [EM00]).

The kinematic entities $\mathbf{F}_s$ and $\mathbf{F}_a$ are the mathematical objects describing the physical processes that influence the transport properties of a tissue (i.e., diffusivity and permeability). To be more specific, $\mathbf{F}_a$ may be decomposed as the product of tensors, which distinguish the inelastc deformations associated with growth from those associated with the exchange of mass between the solid and the fluid phase. The case in which both processes are modelled together has been studied in [GWGM09][GFW11]. In our present contribution, however, we do not consider mass exchange, so that $\mathbf{F}_a$ accounts for the inelastc deformations associated with growth only.

Under the hypothesis of negligible inertial terms and incompressible solid and fluid phases, the unknowns to be determined are given by the displacement field (whose material gradient is $\mathbf{F}_s$), pressure, and $\mathbf{F}_a$. We formulate a boundary value problem for the calculation of pressure and displacements, and discuss how to find an equation for $\mathbf{L}_a$ (tensor $\mathbf{F}_a$ is then found by solving the initial value problem $\dot{\mathbf{F}}_a = \mathbf{L}_a\mathbf{F}_a$, with $\mathbf{F}_a(t_0) = \mathbf{F}_{a0}$). Accepting Darcy's law amounts to say that fluid flow is determined by the hydraulic conductivity, $\mathbf{K}$, and pressure gradient. Therefore, the study of the transport of fluid in a deformable porous medium with variable mass and internal structure reduces to the determination of the influence of $\mathbf{F}_a$ on $\mathbf{K}$ and pressure. In particular, we show that different choices of the initial value $\mathbf{F}_{a0}$ lead to different pressure distributions and displacements. The latter ones, in turn, affect $\mathbf{K}$ and are thus able to influence the capability of the medium of conveying fluid. The alteration of pressure may be relevant for some biomechanical applications in which the health of the cells of a tissue depends

[1] By *classical* we mean here the Continuum Mechanics that studies non-living matter.

on the pressure (for example, this is the case of chondrocytes in articular cartilage). Our study aims to put some of concepts presented in [GZBG07] in a more rigorous framework.

When considering a continuum with growth (or any an elastic effect like plasticity), we intend that the deformations occurred are not reversible but they are the multiplicative combination of inelastc and elastic combination. Under this hypothesis we assume that the deformation gradient is written as

$$\boldsymbol{F}(\boldsymbol{p},t) = \boldsymbol{F}_e(\boldsymbol{p},t)\boldsymbol{F}_a(\boldsymbol{p},t).$$

We moreover denote with J_a and J_e the determinant of $\mathbf{F}_a$ and $\mathbf{F}_e$. Since $\boldsymbol{F}$ is assumed to be invertible so $\mathbf{F}_a$ and $\mathbf{F}_e$ are, and J_a and J_e are both positive.

Among the factors that assess the health of a tissue, an important role is played by the constituents that supply nutriment to the cells and regulate their metabolism. Other relevant constituents are those that can either promote or hinder processes, which can damage the tissue (e.g., growth of tumours). In any of these cases, a mathematical model of a tissue should provide information about the concentration of constituents and the transport processes to which they are subjected. As long as continuum models are concerned (these models do not explicitly track molecular or sub-cellular processes), the evolution of constituents is put in the form of diffusion-advection-reaction equations. Therefore, if the validity of Fick's and Darcy's laws is accepted, it is essential to determine the diffusivity and permeability of the tissue in order to quantify its capability of transporting matter. This capability, on the other hand, depends on geometrical (i.e., geometry of the pore space) and mechanical factors (i.e., deformation and stress), and couplings between them. Indeed, when the tissue deforms, the geometry of the pore space changes and so do the transport properties (diffusivity and permeability). Moreover, different compositions of the fluid phase lead to different hydraulic conductivities. However, we shall neglect this effect in the rest of our study, and we shall focus on the much simpler case of a fluid phase comprising a single constituent only.

2.2 Volumetric growth for monophasic continua

In this section, we will first extend the elasticity model including inhomogeneity due to growth. The final differential problem will consist of a system of three equations whose unknowns are the spatial density $\rho(\mathbf{x},t)$ (or $\rho_r(\mathbf{X},t)$), the deformation and velocity fields $\boldsymbol{\chi}(\mathbf{X},t)$ and $\dot{\boldsymbol{\chi}}(\mathbf{X},t)$, and the tensor $\mathbf{F}_a(\mathbf{X},t)$. In order to have a unique solution, for each of these quantities, an initial distribution has to given. The initial distribution will be denoted by the quantity with a subscript $_0$ and removing the time dependance, e.g. $\rho_0(\mathbf{x})$ for the density.

2.2.1 Kinematics

From the kinematical point of view, we suppose that growth induces inelastc effects on the configurations of the considered body. This is modeled as second order tensor similar to the deformation gradient but it is not gradient by any material field. This will lead to the usual multiplicative decomposition of the deformation gradient used also in plasticity. Material points P belonging to the body manifold $\mathscr{B}$ are mapped in $\mathbb{R}^3$ by a fixed deformation

$$\mathscr{B} \ni \mathsf{P} \to \mathbf{X} = \boldsymbol{\varkappa}(\mathsf{P}) \in \mathbb{R}^3.$$

In the three-dimensional case, this map has the role of assigning coordinates to the material points. The image $\boldsymbol{\varkappa}(\mathscr{B})$ is the reference configuration $\mathscr{B}_r$. The motion is defined as a time-parametrized family of deformations

$$(\mathsf{P}, t) \ni \mathscr{B} \times \mathscr{I} \to \mathbf{x} = \boldsymbol{\chi}(\mathsf{P}, t)$$

or, equivalently, as

$$(\mathbf{X}, t) \ni \mathscr{B}_r \times \mathscr{I} \to \mathbf{x} = \boldsymbol{\chi}(\mathbf{X}, t).$$

The deformation gradient $\mathbf{F}(\mathbf{X}, t)$ is the material gradient of the motion and hence

$$\mathbf{F} = \nabla \boldsymbol{\chi}.$$

The tensor field $\mathbf{F}$ is *induced* by the motion $\boldsymbol{\chi}$ and it is also said to be *compatible*. Its determinant $J(\mathbf{X}, t)$ accounts for the volume change of the particle with material coordinates $\mathbf{X}$. In fact, the volume of a subset $\mathscr{P}_t \subset \mathscr{B}_t$ is measured by

$$\int_{\mathscr{P}_t} 1 \, \mathrm{d}\mathbf{x} = \int_{\mathscr{P}} J \, \mathrm{d}\mathbf{X}. \tag{2.1}$$

The time derivative of (2.1) gives the variation of the volume of $\mathscr{P}_t$. Hence, in material coordinates the volume changes is measured by $\dot{J}$. In order to compute it in spatial coordinates, we have to exploit the formula of the time derivative of a determinant that reads

$$\dot{J} = J \ \mathrm{tr}\,(\dot{\mathbf{F}}\mathbf{F}^{-1}). \tag{2.2}$$

As shown in Sec. 1.1.1, the spatial gradient $\mathbf{L}$ of the velocity $\mathbf{v}$ is related to the deformation gradient by the following identity

$$(\mathbf{L})_m = \dot{\mathbf{F}}\mathbf{F}^{-1}.$$

Following the approach also used to model inhomogeneities such as plasticity and active strain, we assume that growth introduces inelastic effects modeled by a second

order tensor field $\mathbf{F}_a(\mathbf{X}, t)$. This tensor is not induced by any deformation and hence deformations due to growth are not compatible. This tensor describes the inelastic components of the deformation gradient $\mathbf{F}$. We can define the elastic component of the deformation as

$$\mathbf{F}_e(\mathbf{X}, t) = \mathbf{F}(\mathbf{X}, t)\mathbf{F}_a^{-1}(\mathbf{X}, t). \tag{2.3}$$

and hence

$$\mathbf{F} = \mathbf{F}_e\mathbf{F}_a. \tag{2.4}$$

Notice that also $\mathbf{F}_e$ is not compatible, otherwise then also $\mathbf{F}_a$ would have been.

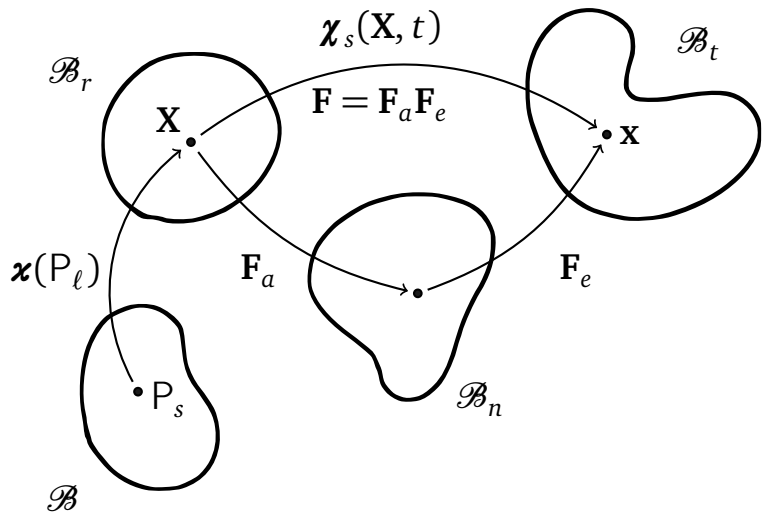


Figure 2.1. Representation of kinematical quantities of an continuum model with growth.

Since $\mathbf{F}$ maps tangent vectors of $\mathscr{B}_r$ into tangent vectors of $\mathscr{B}_t$, with a similar approach we can see $\mathbf{F}_a$ as a map between tangent vectors of $\mathscr{B}$ into tangent vectors of an additional configuration $\mathscr{B}_n$. This latter configuration is known as *natural* or *relaxed* configuration and it is a priori time dependent (see Fig. 2.1). It can be seen as an intermediate configuration that describes the evolution of the body if the mechanical loads are removed [Qui02].

As a consequence of (2.4), the deformation determinant is written as $J = J_eJ_a$, where $J_a = \det(\mathbf{F}_a)$ and $J_e = \det(\mathbf{F}_e)$. Since $\mathbf{F}$ is invertible with positive determinant, so

also J_a and J_e are positive. Since J_a is a determinant, its time derivative (cf. (2.2)) is given by

$$\dot{J}_a = J_a \text{ tr } (\dot{\mathbf{F}}_a \mathbf{F}_a^{-1}). \tag{2.5}$$

With a parallel approach with the standard kinematical definition, we can introduce another tensor

$$\mathbf{L}_a = \dot{\mathbf{F}}_a \mathbf{F}_a^{-1} \tag{2.6}$$

that is the rate of an elastic deformations. Differently from $\mathbf{L}$, this is not a gradient of any velocity field.

If $\dot{J}_a$ is positive, so is tr $\mathbf{L}_a$, and the particle with material coordinates $\mathbf{X}$ is growing. If $\dot{J}_a$ is negative, so is tr $\mathbf{L}_a$, and the particle is being resorbed. If J_a is unitary then there has been no growth but there could have been some local reorganization of the continuum. This would imply that then $\mathbf{L}_a$ is a deviatoric tensor.

The time derivative of the deformation determinant can be obtained by the rule of the derivative of a product

$$\dot{J} = \dot{J}_a J_e + J_a \dot{J}_e$$

and substituting (2.5) into the previous equation, we get

$$\dot{J} = J_a J_e \text{tr}(\mathbf{L}_a) + J_a J_e \text{tr}(\mathbf{L}_e) = J \text{tr } (\mathbf{L}_a + \mathbf{L}_e). \tag{2.7}$$

Comparing equations (2.2) and (2.7), we obtain

$$\text{tr } (\mathbf{L}) = \text{tr } (\mathbf{L}_a) + \text{tr } (\mathbf{L}_e).$$

From (2.6), we obtain an evolution equation for the $\mathbf{F}_a$ that reads

$$\dot{\mathbf{F}}_a = \mathbf{L}_a \mathbf{F}_a, \tag{2.8}$$

where $\mathbf{L}_a$ is the constitutive part. Splitting this last tensor into a deviatoric and isochoric part, Eq. (2.8) can be written as

$$\dot{\mathbf{F}}_a = \big(\text{Dev}(\mathbf{L}_a) + \text{tr}(\mathbf{L}_a)\mathbf{I} \big) \mathbf{F}_a, \tag{2.9}$$

At the right hand side of (2.9), the first term will account for the reorganization and the second term for the changes of volume.

2.2.2 Mass balance

In the previous subsection, we introduced besides the reference and the current configuration also a natural configuration $\mathscr{B}_n$. The natural configuration is more an artificial construction than a real configuration, since no map exists between $\mathscr{B}_r$ and $\mathscr{B}_n$. If we suppose $\mathbf{F}_a$ to be the gradient of a motion $\xi = \boldsymbol{\chi}_a(\boldsymbol{X}, t)$, we could also define a natural

density $\rho_n(\xi, t)$. Hence, since mass is preserved between the natural configuration and the current configuration, we can write

$$\rho(\boldsymbol{\chi}(\mathbf{X}, t), t)J(\mathbf{X}, t) = \rho_n(\boldsymbol{\chi}_a(\mathbf{X}, t), t)J_a(\mathbf{X}, t).$$

that leads to the identity

$$\rho_n(\boldsymbol{\chi}_a(\mathbf{X}, t), t) = \rho(\boldsymbol{\chi}(\mathbf{X}, t), t)J_e(\mathbf{X}, t). \tag{2.10}$$

Since in our setting, no mapping exists between the reference and the natural configuration, and no coordinates exists, we will use (2.10) as an inspiration to define the natural density

$$\rho_n(\mathbf{X}, t) := \rho_m(\mathbf{X}, t)J_e(\mathbf{X}, t) = \rho(\boldsymbol{\chi}(\mathbf{X}, t), t)J_e(\mathbf{X}, t) \tag{2.11}$$

Density ρ_n measures the mass in the natural configuration with respect to the material coordinates.

Hence, we have defined a spatial density $\rho(\mathbf{x}, t)$, a material density $\rho_r(\mathbf{X}, t)$, and a natural density $\rho_n(\mathbf{X}, t)$. These quantities are related by the formulae $\rho_r = \rho_m J$ and $\rho_n = J_e \rho_m$. For a standard continuum, we showed that ρ_r cannot depend on time, but this will not be the case when growth is included.

In order to derive a balance of mass for a such a material, we have to introduce a density for this configuration and we denote it by $\rho_n = \rho_n(\mathbf{X}, t)$. This density gives the changed density of the particle with material coordinates $\mathbf{X}$. The balance of mass for a growing material generalizes the conservation of mass (1.32). In this case the mass is not conserved but a variation is allowed by mass flow and the presence of local sources and sinks of mass. It states that *the variation of mass in a subvolume $\mathscr{P}_t$ is equal to the flow of mass through the boundary $\partial\mathscr{P}_t$ and the volumetric sources*, hence the balance of mass reads

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathscr{P}_t} \rho \,\mathrm{d}\mathbf{x} = \int_{\partial P_t} m(\mathbf{x}, t, \mathbf{n})\mathrm{d}\mathbf{a} + \int_{\mathscr{P}_t} \sigma \,\mathrm{d}\mathbf{x}. \tag{2.12}$$

The surface integral at the right hand side of the previous equation denotes the flux of mass at the boundary of $\mathscr{P}_t$. The mass flow at the boundary depends also on the normal $\mathbf{n}$ and with an argument similar to the Cauchy tetrahedron, we can prove the existence of flux vector $\mathbf{m}$ for which

$$m(\mathbf{x}, t, \mathbf{n}) = \mathbf{m}(\mathbf{x}, t) \cdot \mathbf{n}.$$

Specific forms of $\mathbf{m}$ give rise to heat-like equations for the density. Such a theory requires to include second-gradient effects that we are not going to include in this thesis but further information can be found in [EM00].

The volumetric term in (2.12) is usually assumed to be linear in the density and hence it takes the form $\sigma(\mathbf{x}, t) = \gamma(\mathbf{x}, t)\rho(\mathbf{x}, t)$. In the case at hand, Eq. (2.12) can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathscr{P}_t}\rho\,\mathrm{d}\mathbf{x}=\int_{\mathscr{P}_t}\gamma\rho\,\mathrm{d}\mathbf{x}. \tag{2.13}$$

In a more general setting, the term γ may also be function of the density itself.

We can exchange the derivation and the integral operators passing through the reference configuration. In fact, Eq. (2.14) can be written in material coordinates as

$$\int_{\mathscr{P}}\frac{\mathrm{d}}{\mathrm{d}t}(\rho_m J)\mathrm{d}\mathbf{X}=\int_{\mathscr{P}}\gamma_m\rho_m J\,\mathrm{d}\mathbf{X}. \tag{2.14}$$

where $\mathscr{P}$ is the subset of the reference configuration such that $\mathscr{P}_t=\boldsymbol{\chi}(\mathscr{P},t)$. As done in Sec. 1.1.4, we can expand the left hand side of the last equation and for the arbitrariness of the integration domain, we obtain

$$\rho'+\mathrm{div}\,(\rho\mathbf{v})=\gamma\rho, \tag{2.15}$$

that is the differential form in Lagrangian coordinates of Eq. (2.14). Eq. (2.14) represents the conservation of mass in material coordinates. From the relation between the material and spatial density and exploring the arbitrariness of $\mathscr{P}$ we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_r=\gamma_m\rho_r. \tag{2.16}$$

From this equation, we see that if $\gamma_m<0$ the density of the particle is decreasing while if $\gamma_m>0$ the density is increasing. Finally, substituting in (2.16) the relation between the natural and the material density, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}(\rho_n J_a)=\gamma_m\rho_n J_a. \tag{2.17}$$

and expanding the derivative we get

$$\dot{\rho}_n J_a+\rho J_a\mathrm{tr}\,(\mathbf{L}_a)=\gamma_m\rho_n J_a. \tag{2.18}$$

Imposing the condition

$$\mathrm{tr}\,(\mathbf{L}_a)=\gamma, \tag{2.19}$$

that means that the variation of body mass is compensated for by the rate tr $(\mathbf{L}_a)$, we obtain that the mass density ρ_n is constant in time.

Hence, if the initial distributions ρ_0, $\boldsymbol{\chi}_0$, $\dot{\boldsymbol{\chi}}_0$, and $\mathbf{F}_{a0}$ of the unknowns are given, the constant natural density can be computed and ρ_r can be derived at each time from

$$\rho_r(\mathbf{X},t)=\frac{\rho_n(\mathbf{X})}{J_e(\mathbf{X},t)}=\rho_n(\mathbf{X})\frac{J_a(\mathbf{X},t)}{J(\mathbf{X},t)}.$$

In the approach used in [EM00], the authors assume the existence of a reference crystal, our natural configuration, whose density is constant, and from this hypothesis they prove that relation (2.19) should hold true.

2.2.3 Balance of linear and angular momentum

The balance of linear momentum has to be modified introducing the additional terms due to the source of mass and hence Eq. (1.34) is written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathscr{P}_t}\rho\mathbf{v}\mathrm{d}\mathbf{x}=\int_{\partial\mathscr{P}_t}\mathbf{T}\mathbf{n}\mathrm{d}\mathbf{a}+\int_{\mathscr{P}_t}\mathbf{b}\mathrm{d}\mathbf{x}+\int_{\mathscr{P}_t}\gamma\rho\mathbf{v}\mathrm{d}\mathbf{x}+\int_{\mathscr{P}_t}\rho\mathbf{p}\mathrm{d}\mathbf{x}. \tag{2.20}$$

In the last equation, $\rho_s\mathbf{p}$ is rate of change of momentum due to growth and $\mathbf{b}$ denotes the contribution of the external forces. The left hand side can be simplified writing the integral in the reference configuration as

$$\int_{\mathscr{P}}\rho_m\dot{\mathbf{x}}J\mathrm{d}\mathbf{X}=\int_{\mathscr{P}}\rho_n\dot{\mathbf{x}}J_a\mathrm{d}\mathbf{X}.$$

Using the fact that ρ_n is constant, the time derivative of the right hand side of the last equation can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}(\rho_n\dot{\mathbf{x}}J_a)=\rho_n\frac{\mathrm{d}}{\mathrm{d}t}(\dot{\mathbf{x}}J_a)=\rho_n\ddot{\mathbf{x}}J_a+\rho_n\dot{\mathbf{x}}J_a\mathrm{tr}\,(\mathbf{L}_a)=\rho_m\ddot{\mathbf{x}}J+\rho_m\mathrm{tr}\,(\mathbf{L}_a)\dot{\mathbf{x}}J.$$

From the previous expression, the time derivative of (2.20) is equal to

$$\int_{\mathscr{P}_t}\rho\mathbf{a}\mathrm{d}\mathbf{x}+\int_{\mathscr{P}_t}\rho\mathbf{v}\mathrm{tr}\,(\mathbf{L}_a)\mathrm{d}\mathbf{x}. \tag{2.21}$$

Now using divergence theorem, equality (2.19), and arbitrariness of $\mathscr{P}_t$, the balance (2.20) can be written in differential form as

$$\rho\mathbf{a}=\mathrm{div}\,\mathbf{T}+\mathbf{b}+\rho\mathbf{p}. \tag{2.22}$$

As for the case without growth, from the balance of angular momentum we obtain the algebraic relation $\mathbf{T}=\mathbf{T}^T$. In order to obtain a conservation law in material coordinates, we have to apply the operator $(\cdot)_m$ to equation (2.22) and multiply it by J. It results in

$$J\rho_m\ddot{\mathbf{x}}=J(\mathrm{div}\,\mathbf{T})_m+J\mathbf{b}_m+J\rho_m\mathbf{p}_m. \tag{2.23}$$

Finally, Eq. (2.23) can be written as the following conservation law in material coordinates

$$\rho_r\ddot{\mathbf{x}}=\mathrm{Div}\,(J\mathbf{T}\mathbf{F}^{-T})+\mathbf{b}_r+\rho_r\mathbf{p}_m, \tag{2.24}$$

where $\mathbf{b}_r = J\mathbf{b}_m$ and $\rho_r \mathbf{p}_m$ is the source of additional momentum in material coordinates. The tensor $\mathbf{P} = J\mathbf{T}\mathbf{F}^{-T}$ is the first Piola-Kirchhoff tensor. If we assume that the elastic term has a hyperelastic behavior and hence a strain-energy function

$$\partial \Psi_s(\mathbf{F}_e) \tag{2.25}$$

exists, the Piola-Kirchhoff stress will take the form

$$\mathbf{P} = J_a \mathbf{P}_n \mathbf{F}_a^{-T},$$

where

$$\mathbf{P}_n = \frac{\partial \Psi_n}{\partial \mathbf{F}_e}.$$

Finally, it can be proven by study of dissipation (see Sect. 6.12d) that the source of momentum due to growth is null and hence the momentum equation in material coordinates becomes

$$\rho_r \ddot{\mathbf{x}} = \mathrm{Div}\,(\mathbf{J}_a \mathbf{P}_n \mathbf{F}_a^{-T}) + \mathbf{b}_r. \tag{2.26}$$

2.2.4 Example with neo-Hookean material law

In this section, we present the equation governing growth in an elastic material. Neglecting the terms related to inertia and to the production of extra momentum, we obtain the quasi-static version of (2.22) that reads:

$$-\,\mathrm{div}\,\mathbf{T} = \mathbf{b}. \tag{2.27}$$

Let us suppose again, for the moment, that two maps $\boldsymbol{\chi}_a$ and $\boldsymbol{\chi}_e$ exist between $\mathscr{B}_r$ and $\mathscr{B}_n$, and between $\mathscr{B}_n$ and $\mathscr{B}_t$. This means that we can put a system of natural coordinates in $\mathscr{B}_n$. We call this coordinates ξ. We remind that this is only an hypothesis that we will remove, since our setting is more general.

This hypothesis allows as to write (2.30) in the natural configuration using the Piola transformation. Thus, the momentum equation reads

$$-\,\mathrm{Div}_n(J_e \mathbf{T}_n \mathbf{F}_e^{-T}) = \mathbf{b}_n. \tag{2.28}$$

In (2.30), the Div_n would represent the divergence computed with respect to the coordinates of the natural configuration, and $\boldsymbol{A}_n(\,\cdot\,, t) = A(\boldsymbol{\chi}_e(\,\cdot\,), t)$. The tensor $\mathbf{P}_n = J_e \mathbf{T}_n \mathbf{F}_e^{-T}$ is the first Piola-Kirchhoff stress in the natural configuration and if we assume that the elastic part of deformation is hyperelastic, then $\mathbf{P}_n$ is given by

$$\mathbf{P}_n = \frac{\partial \Psi_n}{\partial \mathbf{F}_e} = \frac{\partial \mathscr{W}_n}{\partial \mathbf{F}_e}.$$

As we have seen in Sect. 1.1.5, the Helmoltz energy function may depend a priori only on the deformation gradients and the coordinates of the configuration, and hence $\Psi_n = \Psi_n(\mathbf{F}_e, \xi)$. Since in general the map $\boldsymbol{\chi}_a$ does not exist and hence no coordinate system can be given in $\mathscr{B}_n$ the strain energy function cannot depend on the point ξ. This explain the restriction, introduced in [EM00], of considering only to uniform materials (also cf. Eq. (2.25)).

Defining the natural second Piola-Kirchhoff tensor $\mathbf{S}_n = \mathbf{F}_e^{-1}\mathbf{P}_n$, equation (2.28) can be written as

$$-\operatorname{Div}_n(\mathbf{F}_e\mathbf{S}_n) = \mathbf{b}_n. \tag{2.29}$$

where $\mathbf{S}_n = \mathbf{S}_n(\mathbf{C}_e)$ can be a function only of the elastic strain $\mathbf{C}_e = \mathbf{F}_e^T\mathbf{F}_e$. If we now use Piola transformation between the natural and the reference configuration and the definition of $\mathbf{F}_e = \mathbf{F}\mathbf{F}_a^{-1}$, we obtain

$$-\operatorname{Div}(J_a\mathbf{F}\mathbf{F}_a^{-1}\mathbf{S}_n\mathbf{F}_a^{-T}) = \mathbf{b}_r. \tag{2.30}$$

In the previous equation, the first Piola Kirchhoff tensor is $\mathbf{P} = J_a\mathbf{F}\mathbf{F}_a^{-1}\mathbf{S}_n\mathbf{F}_a^{-T}$, and if we collect the deformation gradient, we obtain the second Piola-Kirchhoff tensor $\mathbf{S} = J_a\mathbf{F}_a^{-1}\mathbf{S}_n\mathbf{F}_a^{-T}$.

If we suppose that the elastic part of the deformation follows a neo-Hookean law, we can define the energy (1.47) in the natural configuration as

$$\mathscr{W}_n(\mathbf{C}_e) = \frac{\mu_n}{2}\left[\operatorname{tr}(\mathbf{C}_e) - 3\right] - \mu_n \ln(J_e) + \frac{\lambda_n}{2}\left[\ln(J_e)\right]^2. \tag{2.31}$$

It depends only on the elastic deformation tensor $\mathbf{C}_e$. The corresponding $\mathbf{S}_n$ is

$$\mathbf{S}_n = 2\frac{\partial \mathscr{W}_n}{\partial \mathbf{C}_e} = \mu_n[\mathbf{I} - (\mathbf{C}_e)^{-1}] + \lambda_n[\ln(J_e)](\mathbf{C}_e)^{-1}. \tag{2.32}$$

Using now definition (2.3), the natural Piola-Kirchhoff stress can be written as a function of quantities defined in the reference configuration as

$$\mathbf{S}_n = \mu_n[\mathbf{I} - \mathbf{F}_a\mathbf{C}^{-1}\mathbf{F}_a^T] + \lambda_n[\ln\left(\frac{J}{J_a}\right)]\mathbf{F}_a\mathbf{C}^{-1}\mathbf{F}_a^T. \tag{2.33}$$

Hence we can write the explicit form of the Piola-Kirchhoff tensor in the reference configuration as

$$\mathbf{S} = J_a\left(\mu_n\left(\mathbf{C}_a^{-1} - \mathbf{C}^{-1}\right) + \lambda_n \log\left(\frac{J}{J_a}\right)\mathbf{C}^{-1}\right)$$

and

$$\mathbf{P} = J_a\left(\mu_n\left(\mathbf{F}\mathbf{C}_a^{-1} - \mathbf{F}^{-T}\right) + \lambda_n \log\left(\frac{J}{J_a}\right)\mathbf{F}^{-T}\right). \tag{2.34}$$

Eq. (2.30) describes the conservation of momentum for a generic material with inelastic inhomogeneities. This equation has to be coupled with (2.8) that describes the evolution of $\mathbf{F}_a$, and they give rise to the following system of equations

$$\begin{cases} -\mathrm{Div}\,(\mathbf{P}) &= \mathbf{f}_r \\ \dot{\mathbf{F}}_a &= \mathbf{L}_a\mathbf{F}_a \end{cases} \tag{2.35}$$

where $\mathbf{P}$ is given by (2.34) in case of neo-Hookean material law.

2.3 Volumetric growth for poroelastic materials

We will present now the model that will be the subject studied in the next chapters. We will consider a biphasic medium whose solid phase can be subject to volumetric growth. Hence, in the following we will insert the concepts introduced in the previous section in the model derived in Sec. 1.2. The model will be analyzed in details: also performing a study of the dissipation that we omitted in the previous sections. We present the differential problem completed with initial and boundary conditions. Here, we will remind to the reader that the derivation will be based on the following three hypotheses:

D1: incompressibility of the phases,

D2: negligible acceleration for the solid phase, and negligible inertia forces for the fluid,

D3: validity of Darcy's law.

Because of hypothesis 1, we will point out that there will be an analogy between the density for a monophasic continuum and the solid volume fraction of a biphasic continuum.

In the following subsections, we will use the same symbols used for monophasic continua but we will add the subscript $_s$ and $_\ell$ to denote respectively the solid and the fluid phase.

2.3.1 Kinematics

We first remind the basic kinematical definitions. The solid and fluid manifolds will be denoted respectively by $\mathscr{B}_s$ and $\mathscr{B}_\ell$, and their material particles will be denoted by $\mathsf{P}_s \in \mathscr{B}_s$ and $\mathsf{P}_\ell \in \mathscr{B}_\ell$. We assume there exists a time-independent embedding $\boldsymbol{\kappa}_s$ that associates a material coordinates

$$\mathbf{X}_s = \boldsymbol{\kappa}_s(\mathsf{P}_s)$$

to the particle P_s. The image $\mathscr{B}_r = \boldsymbol{\kappa}_s(\mathscr{B}_s)$ is the reference configuration of the solid phase. The motion of the solid phase, also known as s-motion is defined by

$$\mathscr{B}_r \times \mathscr{I} \ni (\mathbf{X}_S, t) \longmapsto \mathbf{x} = \boldsymbol{\chi}_s(\mathbf{X}, t) \in \mathbb{R}^3,$$

where $\mathscr{I} = (0, T_f]$ denotes a time interval. The inverse motion $\mathbf{X} = \boldsymbol{\pi}_s(\mathbf{x}, t)$ associates at each time t the material co-ordinate $\mathbf{X}$ of the particle that occupies the spatial point $\mathbf{x}$. Usually no reference configuration is supposed to exist for fluid particles and hence the ℓ-motion is defined as

$$\mathscr{B}_\ell \times \mathscr{I} \ni (\mathsf{P}_\ell, t) \longmapsto \mathbf{x} = \boldsymbol{\chi}_\ell(\mathsf{P}_\ell, t) \in \mathbb{R}^3.$$

The current configuration $\mathscr{B}_t = \boldsymbol{\chi}_s(\mathscr{B}_r, t)$ is the image of the reference configuration of the solid phase. The deformation gradient is the tangent tensor field at the motion defined as

$$\mathbf{F}_s = \nabla \boldsymbol{\chi}_s.$$

Since we allow the solid phase to grow, besides the tensor field $\mathbf{F}_s$, we also introduce a non compatible tensor field $\mathbf{F}_a(\mathbf{X}_s, t)$ that described the material inhomogeneities due to the change of mass. The elastic deformations are described by the non compatible tensor field

$$\mathbf{F}_e = \mathbf{F}_s \mathbf{F}_a^{-1}.$$

The deformation gradient $\mathbf{F}_s$ is a linear mapping between the tangent space of $\mathscr{B}_r$ and $\mathscr{B}_t$, and $\mathbf{F}_a$ is a linear mapping between the tangent space of the refinance configuration and an intermediate natural configuration $\mathbf{B}_n$.

The same point of Euclidean space is occupied by both solid and fluid particles. At each point $\mathbf{x}$, a velocity $\mathbf{v}_\alpha(\mathbf{x}, t)$ of the phase $\mathscr{P}_\alpha$. The material solid velocity and the spatial solid velocity are related by the formula

$$\dot{\boldsymbol{\chi}}_s(\mathbf{X}, t) = \mathbf{v}_s(\boldsymbol{\chi}_s(\mathbf{X}, t), t) = \left(\mathbf{v}_s\right)_m(\mathbf{X}, t).$$

In particular, notice that

$$\left(\mathbf{v}_\ell\right)_m \neq \mathbf{v}_\ell \circ \boldsymbol{\chi}_\ell.$$

The local distribution of the two phases is given by the volume fractions $\phi_\alpha(\mathbf{x}, t)$. The solid volume fraction is also defined on the reference configuration using the Piola transformation

$$\Phi_s = (\phi_s)_m J.$$

Following the definition (2.11) of density in the natural configuration, we can define a natural density

$$\Phi_{sn}(\mathbf{X},t) = \Phi_{sn} = (\phi_s)_m J_e. \tag{2.36}$$

2.3.2 Balance laws

The Eulerian, local form of the balance of mass of the solid and fluid phase reads

$$\begin{array}{rcrcl} \rho_s' & + & \mathrm{div}(\rho_s \mathbf{v}_s) & = & \rho_s \gamma_s \\ \rho_\ell' & + & \mathrm{div}(\rho_\ell \mathbf{v}_\ell) & = & 0 \end{array} \tag{2.37}$$

where γ_s is the rate at which the mass of the solid phase is produced or depleted. The density $\rho_\alpha = \phi_\alpha \hat{\rho}_\alpha$ is written as the product of the α-th volume fraction and and bulk density. Using the operator

$$D_s\mathbf{A} = \mathbf{A}' + \mathrm{grad}(\mathbf{A}) \cdot \mathbf{v}_s = \left(\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{A})_m \right) \circ \pi,$$

that is the convective derivative of the generic tensor field $\mathbf{A}$ with respect to the motion of the solid phase, equations (2.37) can be written as:

$$\begin{array}{rcrcrcl} D_s\rho_s & + & \rho_s \mathrm{div}(\mathbf{v}_s) & & & = & \rho_s\gamma_s \\ D_s\rho_\ell & + & \rho_\ell \mathrm{div}(\mathbf{v}_s) & + & \mathrm{div}(\rho_\ell \mathbf{w}_{\ell s}) & = & 0 \end{array} \tag{2.38}$$

As done also for the case without growth (see Sect. 1.2.3), writing (2.38) in material co-ordinates and multiplying by J leads to the following form of the mass balance laws of the constituents of the mixture:

$$\begin{array}{rcrcl} \dot{\overline{(J\rho_{sm})}} & & & = & J\rho_{sm}\gamma_{sm}, \\ \dot{\overline{(J\rho_{\ell m})}} & + & \mathrm{Div}(J\rho_{\ell m}\mathbf{F}^{-1}\mathbf{w}_{\ell s}) & = & 0. \end{array} \tag{2.39}$$

By using the definition of apparent mass density and from the incompressibility hypothesis, the previous equations become evolution equations for the volume fractions that read

$$\begin{array}{rcrcl} \dot{\overline{(J\phi_{sm})}} & & & = & J\phi_{sm}\gamma_{sm}, \\ \dot{\overline{(J\phi_{\ell m})}} & + & \mathrm{Div}(J\phi_{\ell m}\mathbf{F}^{-1}\mathbf{w}_{\ell s}) & = & 0. \end{array} \tag{2.40}$$

Since $\Phi_s = J(\phi_s)_m$, Eq. (2.40.1) defines an evolution model for the reference density Φ_s. Notice that differently from the models presented in Chapt. 1, $\Phi_s = \Phi_s(\mathbf{X},t)$ is not constant in time. The definition (2.36) allows to write the reference solid volume fraction as $\Phi_s = \Phi_{sn}J_a$ and putting this equality in (2.40.1), we obtain

$$\mathrm{tr}(\mathbf{L}_a)\Phi_{sn} + \dot{\Phi}_{sn} = \Phi_{sn}\gamma_s. \tag{2.41}$$

Enforcing the condition that the variation of body mass is compensated by the rate $\mathrm{tr}(\mathbf{L}_a)$, we derive that the mass density Φ_{sn} is constant in time. Thus, we arrive at the results

$$\gamma_s = \mathrm{tr}(\mathbf{L}_a), \quad \Phi_{sn} = \Phi_{sn0}, \tag{2.42}$$

where $\Phi_{sn0} = \Phi_{sn0}(\mathbf{X})$ may be a function of material coordinates only. A consequence of (2.41) and ((2.42)) is that the solution to (2.40.1) is given by

$$\phi_s = \frac{\Phi_{sn}}{J_e} = \Phi_{sn}\frac{J_a}{J}. \tag{2.43}$$

This means that the apparent volume fraction of the solid phase, Φ_s, is determined if the constant distribution Φ_{sn0} is assigned, and the volumetric deformations J and J_a are known. We remark that Φ_{sn} is constant and should be regarded as a known quantity of the model.

Following the same steps but without using the incompressibility constraint, from Eq. (2.39.1) we would have obtained

$$\rho_s = \frac{\rho_{sn}}{J_e} = \rho_{sn}\frac{J_a}{J}. \tag{2.44}$$

which involves in this case measures of the density of the solid phase.

A simpler expression of the mass balance can be obtained substituting (2.40.2) from the sum of both (2.40), that leads to

$$\dot{J} = -\mathrm{Div}[J\mathbf{F}^{-1}\mathbf{q}_{\ell s}] + J_a\phi_{sn}\mathrm{tr}(\mathbf{L}_a), \tag{2.45}$$

with $\mathbf{q}_{\ell s} := \phi_\ell \mathbf{w}_{\ell s}$. In the derivation of this equation we used (2.43) and (2.42.1). In summary, (2.43) and (2.45) provide the balances of mass of the solid and the fluid phase, respectively.

2.3.3 Balance of linear momentum

The Eulerian, local form of the balance of momentum of the phases is given by

$$(\rho_s\mathbf{v}_s)' + \mathrm{div}(\rho_s\mathbf{v}_s\otimes\mathbf{v}_s) - \mathrm{div}(\mathbf{T}_s) = \rho_s\mathbf{m}_s + (\rho_s\mathbf{p}_s + \rho_s\gamma_s\mathbf{v}_s), \tag{2.46}$$

$$(\rho_\ell\mathbf{v}_\ell)' + \mathrm{div}(\rho_\ell\mathbf{v}_\ell\otimes\mathbf{v}_\ell) - \mathrm{div}(\mathbf{T}_\ell) = \rho_\ell\mathbf{m}_\ell, \tag{2.47}$$

where, as in Sect. 1.2.4, $\mathbf{T}_s$ and $\mathbf{T}_\ell$ denote the Cauchy stress tensors of the solid and fluid phase, $\rho_s\mathbf{m}_s$ and $\rho_\ell\mathbf{m}_\ell$ are the rate of exchange of momentum between the two phases. As in Eq. (2.20), we also added the term $\rho_s\mathbf{p}_s$ that is the rate of change of momentum due to growth. The forces $\rho_s\mathbf{m}_s$ and $\rho_\ell\mathbf{m}_\ell$ satisfy the condition

$$\rho_s\mathbf{m}_s + \rho_\ell\mathbf{m}_\ell = \mathbf{0}, \tag{2.48}$$

which states that, in the absence of growth (i.e. when $\gamma_s = 0$ and $\mathbf{p}_s = \mathbf{0}$), the mixture is closed with respect to momentum. The use of (1.74), and the definition of the relative velocities $\tilde{\mathbf{v}}_\ell$ and $\tilde{\mathbf{v}}_s$ as well as of the accelerations $\mathbf{a}$ and $\mathbf{a}_\ell$ allow for reformulating the balance laws (1.72) and (1.73) in the following way

$$\rho\mathbf{a} - \operatorname{div}(\mathbf{T}) = \rho_s\mathbf{p}_s + \rho_s\gamma_s\tilde{\mathbf{v}}_s, \tag{2.49}$$

$$\rho_\ell\mathbf{a}_\ell - \operatorname{div}(\mathbf{T}_\ell) = \rho_\ell\mathbf{m}_\ell, \tag{2.50}$$

where $\mathbf{T}$, which denotes the Cauchy stress tensor of the mixture, is defined by

$$\mathbf{T} := \mathbf{T}_s + \mathbf{T}_\ell - \left[\rho_s\tilde{\mathbf{v}}_s \otimes \tilde{\mathbf{v}}_s + \rho_\ell\tilde{\mathbf{v}}_\ell \otimes \tilde{\mathbf{v}}_\ell\right]. \tag{2.51}$$

Equation (2.49) represents the balance of momentum of the mixture as a whole, and is obtained by adding together (1.72) and (1.73) and applying the definitions (1.55)–(1.57) to the result. Furthermore, substituting the identity

$$\rho\mathbf{a} := \textstyle\sum_{\alpha=\ell,s}\rho_\alpha\mathbf{a}_\alpha - \sum_{\alpha=\ell,s}\operatorname{div}(\rho_\alpha\tilde{\mathbf{v}}_\alpha \otimes \tilde{\mathbf{v}}_\alpha) + \rho_s\gamma_s\tilde{\mathbf{v}}_s \tag{2.52}$$

into (2.49) yields

$$\rho_s\mathbf{a}_s + \rho_\ell\mathbf{a}_\ell - \operatorname{div}(\mathbf{T}_s + \mathbf{T}_\ell) = \rho_s\mathbf{p}_s, \tag{2.53}$$

$$\rho_\ell\mathbf{a}_\ell - \operatorname{div}(\mathbf{T}_\ell) = \rho_\ell\mathbf{m}_\ell. \tag{2.54}$$

Finally, neglecting the inertial forces of both phases, the balances of momentum (1.72) and (1.73) become

$$-\operatorname{div}(\mathbf{T}_s + \mathbf{T}_\ell) = \rho_s\mathbf{p}_s, \tag{2.55}$$

$$-\operatorname{div}(\mathbf{T}_\ell) + \rho_\ell\mathbf{m}_\ell = \mathbf{0}. \tag{2.56}$$

By means of the Piola transformations of (2.55) and (2.56), the momentum balance laws of the mixture can be written with respect to the reference placement $\mathscr{B}_r$, i.e.

$$\operatorname{Div}(\mathbf{P}_s + \mathbf{P}_\ell) + J\rho_s\mathbf{p}_s = \mathbf{0}, \tag{2.57}$$

$$\operatorname{Div}(\mathbf{P}_\ell) + J\rho_\ell\mathbf{m}_\ell = \mathbf{0},. \tag{2.58}$$

where we used the first Piola-Kirchhoff stress tensors $\mathbf{P}_s$ and $\mathbf{P}_\ell$ of the solid and the fluid phase respectively. In order to close the mathematical problem resulting from (2.45), (2.57) and (2.58), it is necessary to provide information about the stresses $\mathbf{P}_s$ and $\mathbf{P}_\ell$, and the force densities $\rho_\ell\mathbf{m}_\ell$ and $\rho_s\mathbf{p}_s$.

The change of internal structure of the solid phase is a process whose kinematics are described by the tensor map $\mathbf{F}_a$ and the generalised velocity $\mathbf{L}_a = \dot{\mathbf{F}}_a(\mathbf{F}_a)^{-1}$. The set of generalized forces that perform working on $\mathbf{L}_a = \dot{\mathbf{F}}_a(\mathbf{F}_a)^{-1}$ comprises an internal force, $\mathbf{Z}_n$, which drives the structural evolution, and an external force, $\mathbf{Y}_n$, which models

the interaction of the system with its surrounding environment. Both forces are second-order tensors. It is postulated that they obey the balance law

$$\mathbf{Z}_n = \mathbf{Y}_n, \tag{2.59}$$

The subscript $_n$ means that $\mathbf{Z}_n$ and $\mathbf{Y}_n$ are conceived as forces acting on the natural configuration $\mathscr{B}_n$, although they can be also written with respect to $\mathscr{B}_r$ and $\mathscr{B}_t$ by performing proper transformations. The reader is referred to [DQ02] for an exhaustive explanation of the physical concepts leading to (2.59). Here, we simply state that, in analogy with the balance laws (2.55) and (2.56) (which describe a balance of forces that perform working on the set of standard velocities), also the forces that perform working on the non-standard descriptor $\mathbf{L}_a$ should satisfy a balance law. Some extensions of the results presented in [DQ02] can be found, for example, in [AG07] [FFA11] [AGDM08] [APV10] [GWGM09] [GFW+09] [GFW11].

The internal force-like variables $\rho_\ell \mathbf{m}_\ell$ and $\mathbf{Z}_n$ are responsible for dissipation, and should thus comply with the dissipation inequality that characterises the system under investigation.

2.3.4 Study of the residual dissipation

We introduce the total internal working of the system

$$\mathfrak{W}^{in}(\mathscr{B}_t) = \mathfrak{W}^{in}_{\text{st}}(\mathscr{B}_t) + \mathfrak{W}^{in}_{\text{n-st}}(\mathscr{B}_t), \tag{2.60}$$

where $\mathfrak{W}^{in}_{\text{st}}(\mathscr{B}_t)$ and $\mathfrak{W}^{in}_{\text{n-st}}(\mathscr{B}_t)$ describe, respectively, the working performed by the standard and non-standard forces acting on the system. These two contributions are defined by the following expressions

$$\mathfrak{W}^{in}_{\text{n-st}}(\mathscr{B}_t) := \int_{\mathscr{B}_t} (J_e)^{-1} \mathbf{Z}_n : \mathbf{L}_a, \tag{2.61}$$

$$\mathfrak{W}^{in}_{\text{st}}(\mathscr{B}_t) := \int_{\mathscr{B}_t} \{ -\rho_\ell \mathbf{m}_\ell \cdot \mathbf{w}_{\ell s} + \mathbf{T}_\ell : \text{grad}(\mathbf{w}_{\ell s}) + (\mathbf{T}_s + \mathbf{T}_\ell) : \text{grad}(\mathbf{v}_s) \}. \tag{2.62}$$

In a purely mechanical context, we call dissipation the quantity

$$\int_{\mathscr{B}_t} \mathscr{D} := -\int_{\mathscr{B}_t} \left[\rho_s D_s \Psi_s + \rho_\ell D_\ell \Psi_\ell \right] + \mathfrak{W}^{in}_{\text{st}}(\mathscr{B}_t) + \mathfrak{W}^{in}_{\text{n-st}}(\mathscr{B}_t) \geq 0, \tag{2.63}$$

which is assumed to be non-negative. In terms of the overall Helmholtz free energy density of the system, $\rho\Psi$, the first term on the RHS of (2.63) can be written as

$$-\int_{\mathscr{B}_t} \left[\rho_s D_s \Psi_s + \rho_\ell D_\ell \Psi_\ell \right] = -\mathrm{d}_t \int_{\mathscr{B}_t} \rho\Psi + \int_{\partial \mathscr{B}_t} \mathbf{q}_\Psi \cdot \mathbf{n} + \int_{\mathscr{B}_t} \rho_s \gamma_s \Psi_s, \tag{2.64}$$

where

$$\rho\Psi := \rho_s\Psi_s + \rho_\ell\Psi_\ell, \tag{2.65}$$

$$\mathbf{q}_\Psi := -\left[\rho_s\Psi_s\tilde{\mathbf{v}}_s + \rho_\ell\Psi_\ell\tilde{\mathbf{v}}_\ell\right]. \tag{2.66}$$

Under the hypotheses of hyperelastic solid phase and macroscopically inviscid fluid, the study of the inequality (2.63) yields the following results for the Cauchy stresses $\mathbf{T}_s$ and $\mathbf{T}_\ell$:

$$\mathbf{T}_s = -\phi_s\pi\mathbf{g}^{-1} + \mathbf{g}^{-1}\left(\rho_s\frac{\partial\Psi_s}{\partial\mathbf{F}_e}\right)(\mathbf{F}_a)^{-T}\mathbf{F}^T, \tag{2.67}$$

$$\mathbf{T}_\ell = -\phi_\ell\pi\mathbf{g}^{-1}, \tag{2.68}$$

where $\mathbf{g}$ is the metric tensor associated with $\mathscr{B}_t$

Objectivity

Requiring the invariance of constitutive laws under superimposed rigid motions places further restrictions on the results (2.67) and (2.68). If a rigid motion is impressed, the points $\mathbf{x} \in \mathscr{B}_t$ transform as $\mathbf{x} \mapsto \bar{\mathbf{x}} = \mathbf{R}\,\mathbf{x} + \mathbf{c}$, where $\mathbf{R}$ is a proper orthogonal tensor defining a pure rotation, and $\mathbf{c}$ is a vector defining a pure rigid translation [Ogd84][GFA10]. Consequently, $\mathbf{F}$, $\mathbf{F}_e$ and $\mathbf{F}_a$ transform as follows

$$\mathbf{F} \mapsto \bar{\mathbf{F}} = \mathbf{R}\mathbf{F}, \quad \mathbf{F}_e \mapsto \bar{\mathbf{F}}_e = \mathbf{R}\mathbf{F}_e, \quad \mathbf{F}_a \mapsto \bar{\mathbf{F}}_a = \mathbf{F}_a. \tag{2.69}$$

However, the Helmholtz free energy density Ψ_s has to remain invariant under these transformations. Therefore, Ψ_s may depend on $\mathbf{F}_e$ only through the Cauchy stretch tensor $\mathbf{C}_e = (\mathbf{F}_e)^T\mathbf{F}_e$, which is independent on $\mathbf{R}$. This yields the relation

$$\mathbf{g}^{-1}\left(\rho_s\frac{\partial\Psi_s}{\partial\mathbf{F}_e}\right) = \mathbf{F}_e\left(2\rho_s\frac{\partial\mathscr{W}_s}{\partial\mathbf{C}_e}\right), \tag{2.70}$$

where $\mathscr{W}_s$ is the Helmholtz free energy density of the solid phase written as function of $\mathbf{C}_e$. On the other hand, the Cauchy stresses $\mathbf{T}_s$ and $\mathbf{T}_\ell$ transform as $\bar{\mathbf{T}}_\alpha = \mathbf{R}\mathbf{T}_\alpha\mathbf{R}^T$, with $\alpha = s, \ell$.

Constitutive laws

Furthermore, by invoking the incompressibility of the solid phase and using the definitions (2.43) and (1.85), the first Piola-Kirchhoff stress tensors become

$$\mathbf{P}_s = -J_a\phi_{sn}\pi\mathbf{g}^{-1}\mathbf{F}^{-T} + \mathbf{F}\left[J_a(\mathbf{F}_a)^{-1}\left(2\frac{\partial\mathscr{W}_{sn}}{\partial\mathbf{C}_e}\right)(\mathbf{F}_a)^{-T}\right], \tag{2.71}$$

$$\mathbf{P}_\ell = -(J - J_a\phi_{sn})\pi\mathbf{g}^{-1}\mathbf{F}^{-T}, \tag{2.72}$$

where $\mathscr{W}_{sn} = \phi_{sn}\hat{\rho}_s\mathscr{W}_s$. Since $\mathbf{F}^{-T} = (\mathbf{F}_e)^{-T}(\mathbf{F}_a)^{-T}$, the stress $\mathbf{P}_s$ can be rewritten as

$$\mathbf{P}_s = J_a\mathbf{P}_{sn}(\mathbf{F}_a)^{-T}, \quad \mathbf{P}_{sn} := -\phi_{sn}\pi\mathbf{g}^{-1}(\mathbf{F}_e)^{-T} + \mathbf{F}_e\left(2\frac{\partial\mathscr{W}_{sn}}{\partial\mathbf{C}_e}\right). \tag{2.73}$$

Finally, we call Mandel stress the tensor the quantity

$$\mathbf{M}_{sn} := (\mathbf{F}_e)^t\mathbf{P}_{sn} := -\phi_{sn}\pi(\mathbf{g}_n)^{-1} + \boldsymbol{\eta}^{-1}\mathbf{C}_e\left(2\frac{\partial\mathscr{W}_{sn}}{\partial\mathbf{C}_e}\right), \tag{2.74}$$

with $(\mathbf{F}_e)^t = (\mathbf{g}_n)^{-1}(\mathbf{F}_e)^T\mathbf{g}$ and $\boldsymbol{\eta}$ being the metric tensor of the intermediate configuration.

Darcy's law

The constitutive results (2.71) and (2.72) allow for a simplification of the expression of dissipation. After localisation, we obtain

$$\mathscr{D} = -\{\rho_\ell\mathbf{m}_\ell - \pi\,\mathrm{grad}(\phi_\ell)\}\cdot\mathbf{w}_{\ell s} + (J_e)^{-1}\{\mathbf{M}_{sn} + \mathbf{Z}_n\} : \mathbf{L}_a \geq 0 \tag{2.75}$$

We introduce now the dissipative part of the force-like variables $\mathbf{m}_\ell$ and $\mathbf{Z}_n$, which are given by

$$\rho_\ell\check{\mathbf{m}}_\ell := \rho_\ell\mathbf{m}_\ell - \pi\mathrm{grad}(\phi_\ell), \quad \check{\mathbf{Z}}_n := \mathbf{M}_{sn} + \mathbf{Z}_n, \tag{2.76}$$

and are constrained to satisfy the inequality

$$\mathscr{D} = -\rho_\ell\check{\mathbf{m}}_\ell\cdot\mathbf{w}_{\ell s} + (J_e)^{-1}\check{\mathbf{Z}}_n : \mathbf{L}_a \geq 0. \tag{2.77}$$

Dissipation has to be zero when $\mathbf{w}_{\ell s}$ and $\mathbf{L}_a$ vanish. Substitution of (2.76) into the balance laws (2.56) and (2.59), and use of the constitutive result (2.68) yield

$$\rho_\ell\check{\mathbf{m}}_\ell = \phi_\ell\mathrm{grad}(\pi), \tag{2.78}$$

$$\check{\mathbf{Z}}_n = \mathbf{M}_{sn} + \mathbf{Y}_n. \tag{2.79}$$

Equations (2.78) and (2.79) represent the new form of the force balances that define, respectively, the dynamics of the fluid phase and the internal structure of the solid phase (the dynamics of the mixture as a whole are defined by (2.55) or (2.57)). Before proceeding with the determination of the forces $-\rho_\ell\check{\mathbf{m}}_\ell$ and $(J_e)^{-1}\check{\mathbf{Z}}_n$, a discussion about the study of the dissipation inequality (2.77), and some of its implications, is mandatory.

Let us set $\mathbf{L}_a = \mathbf{0}$ and focus on the pair $(-\rho_\ell\check{\mathbf{m}}_\ell, \mathbf{w}_{\ell s})$. In Biomechanics, it is often assumed that fluid flow obeys Darcy's law (cf., for example, [AW10]). Darcy's model of flow can be retrieved consistently with the study of dissipation (cf., for example, [HG89]) by expressing $-\rho_\ell\check{\mathbf{m}}_\ell$ as a constitutive function of $\mathbf{w}_{\ell s}$ that vanishes when $\mathbf{w}_{\ell s} = \mathbf{0}$. This function is then expanded in Taylor series in a neighbourhood of $\mathbf{w}_{\ell s} =$

$\mathbf{0}$ and, for small velocities, only the first-order term of the expansion is maintained. Therefore, one obtains

$$-\rho_\ell \check{\mathbf{m}}_\ell = \mathbf{A}.\mathbf{q}_{\ell s}, \tag{2.80}$$

where $\mathbf{q}_{\ell s} := \phi_\ell \mathbf{w}_{\ell s}$ is called *specific discharge* and $\mathbf{A}$ is a positive-definite second-order tensor that represents the resistivity of the medium (in this discussion, ϕ_ℓ is assumed to be strictly different from zero). Substitution of this result into (2.78), and inverting $\mathbf{A}$ yield

$$\mathbf{q}_{\ell s} = -\mathbf{K}.\mathrm{grad}(\pi), \tag{2.81}$$

where $\mathbf{K} = \phi_\ell \mathbf{A}^{-1}$ is said to be the hydraulic conductivity of the medium. Equation (2.81) is a simplified form of Darcy's law, in which the buoyancy term is neglected. According to (2.81), $\mathbf{q}_{\ell s}$ vanishes in a permeable medium when the pressure gradient is zero. It is known, however, that Darcy's law may cease to be valid if, for example, the flow starts only when $-\rho_\ell \check{\mathbf{m}}_\ell$ exceeds a certain threshold, which is given by the pressure gradient [Bea72]. In this case, the force $-\rho_\ell \check{\mathbf{m}}_\ell$ need not be smooth at $\mathbf{w}_{\ell s} = \mathbf{0}$.

Growth model

Let us put now $\mathbf{w}_{\ell s} = \mathbf{0}$ and study the pair $((J_e)^{-1}\check{\mathbf{Z}}_n, \mathbf{L}_a)$. In some models of growth mechanics, constitutive laws of the type $\check{\mathbf{Z}}_n = \mathbb{H}_n : \mathbf{L}_a$ have been proposed (cf., for example, [AG07][LSo05][AGDM08][GFW$^+$09][APV10][TEST10][GFW11]), with $\mathbb{H}_n$ being a diagonally symmetric, positive-definite fourth-order tensor. The rate of inelastc deformation $\mathbf{L}_a$ was thus presented in the form $\mathbf{L}_a = \mathbb{G}_n : \check{\mathbf{Z}}_n$, with $\mathbb{G}_n = (\mathbb{H}_n)^{-1}$. Substitution of (2.79) in this relation yields

$$\mathbf{L}_a = \mathbb{G}_n : (\mathbf{M}_{sn} + \mathbf{Y}_n). \tag{2.82}$$

Equation (2.82) follows from the hypothesis that $\check{\mathbf{Z}}_n$ can be assigned as a constitutive function of $\mathbf{L}_a$ that vanishes when $\mathbf{L}_a = \mathbf{0}$. This function is then assumed to be smooth and linearized in a neighbourhood of $\mathbf{L}_a = \mathbf{0}$. For a positive-definite $\mathbb{G}_n$, the formula (2.82) admits the following interpretation: the rate of inelastc deformation, $\mathbf{L}_a$, becomes zero when the external force $\mathbf{Y}_n$ can be tuned in such a way that the sum $(\mathbf{M}_{sn} + \mathbf{Y}_n)$ vanishes. This situation implies that $\mathbf{F}_a$ (which always satisfies the kinematic relation $\dot{\mathbf{F}}_a = \mathbf{L}_a \mathbf{F}_a$) either ceases to evolve in time or remains equal to its initial value. In some biomechanical applications, $\mathbf{Y}_n$ is thought of as the "target stress" that regulates the process with which it is associated (when the target stress is reached, the process ceases). For example, in the model of arterial growth proposed in [OK08], $\mathbf{Y}_n$ is related to the homeostatic stress. On the other hand, if the tensor $\mathbf{Y}_n$ is zero (or negligibly small), the equality $\mathbf{L}_a = \mathbf{0}$ cannot be recovered in general, since the Mandel stress, $\mathbf{M}_{sn}$, is not compensated by any external force. Furthermore, the relation (2.82) could be too restrictive for some applications. In fact, one may relax (2.82) and postulate an evolution law of the type [Mić09]

$$\dot{\mathbf{V}}_a = \langle f \rangle \textstyle\sum_{h=0}^{2}\sum_{k=0}^{2} \beta_{hk}(\mathbf{V}_a, \mathbf{S}_{sn})\left[(\mathbf{V}_a)^h(\mathbf{S}_{sn})^k + (\mathbf{S}_{sn})^k(\mathbf{V}_a)^h\right]. \tag{2.83}$$

Here, $\mathbf{V}_a$ is the symmetric part of $\mathbf{F}_a$ (cf. the decomposition $\mathbf{F}_a = \mathbf{V}_a\mathbf{R}_a$ [Ogd84], where $\mathbf{R}_a$ –the tensor of inelastc rotation– is set equal to the identity), $\mathbf{S}_{sn} := (\mathbf{F}_e)^{-1}\mathbf{P}_{sn}$ is the second Piola-Kirchhoff stress tensor, β_{hk} is a given constitutive function of the invariants of $\mathbf{V}_a$, $\mathbf{S}_{sn}$ and compositions of these tensors, f is said to be dynamic yield function, and the symbol $\langle f \rangle$ equals unity when the inelastc deformation changes with time and equals zero otherwise. Constitutive restrictions on $\mathbf{L}_a$, $\mathbf{S}_{sn}$ and the dynamic yield function are then found by using (2.83) in the computation of the extrema of the inelastc working.

Another method for determining evolution laws is given in [CFS01], where rate-independent plasticity is investigated. The dissipation is defined as a function of $\mathbf{L}_a$ and is assumed to be continuous, but generally non-differentiable, at $\mathbf{L}_a = \mathbf{0}$, while the tensor $\check{\mathbf{Z}}_n$ is constitutively indeterminate at $\mathbf{L}_a = \mathbf{0}$. Within this framework, a maximum-dissipation criterion is formulated and it is proven that the dissipation function is everywhere sub-differentiable and, thus, convex with respect to $\mathbf{L}_a$, and that $\check{\mathbf{Z}}_n$ must belong to the sub-differential of the dissipation function. The evolution of $\mathbf{L}_a$ is determined by introducing a scalar yield criterion through the yield function f and showing that $\mathbf{L}_a$ has to be an element of the sub-differential of f. In the case of a smooth yield function f, it is found that $\mathbf{L}_a$ follows the "normality rule"

$$\mathbf{L}_a = \lambda \frac{\partial f}{\partial \check{\mathbf{Z}}_n}, \tag{2.84}$$

where λ is a Kuhn-Tucker multiplier satisfying the conditions $\lambda \geq 0$, $f \leq 0$, $\lambda f = 0$, and determined by the consistency requirement $\lambda \dot{f} = 0$. Many of the mathematical tools for presenting this theory can be found in [Roc70]. In the context of biological materials, an evolution law for of the type (2.84) was proposed in [PV11] for modelling the reorganization of cells (an inelastc deformation) of a tissue in the presence of growth (e.g., a tumour).

We remark that (2.82)–(2.84) are all plausible ways to determine the evolution of $\mathbf{L}_a$. They are, however, different from each other since they are conceived for modelling different physical situations. On the other hand, the common feature of all these models is that the relations linking the generalised forces $-\rho_\ell \check{\mathbf{m}}_\ell$ and $\check{\mathbf{Z}}_n$ with their power-conjugate generalised velocities $\mathbf{q}_{\ell s}$ and $\mathbf{L}_a$ satisfy a maximum-dissipation principle (cf. [RS04]).

2.3.5 Summary of the model

After having derived the equations describing growth in a poroelastic material, we will now present the complete differential model that governs the behaviour of such kind of media.

From the study of dissipation we derived the Darcy's law (2.81) that can be written in material coordinates as

$$(\mathrm{grad}(\pi))_m = \mathbf{g}^{-1}\mathbf{F}^{-T}\mathrm{Grad}(\pi_m).$$

Substituting this equality into (2.45), using the constitutive results (2.71) and (2.72), and neglecting $\mathbf{p}_s$ in (2.55), we obtain

$$\begin{array}{rcl} -\mathrm{Div}\left[\mathbf{P} - J\mathbf{g}^{-1}\pi\mathbf{F}^{-T}\right] & = & \mathbf{0}, \\ -\dot{J} + \mathrm{Div}\left[\mathbf{K}_r.\mathrm{Grad}(\pi)\right] & = & -J_a\phi_{sn}\mathrm{tr}(\mathbf{L}_a), \\ \dot{\mathbf{F}}_a & = & \mathbf{L}_a\mathbf{F}_a. \end{array} \tag{2.85}$$

In Eq. (2.85.1), $\mathbf{P}$ denotes from here on the constitutive part of $\mathbf{P}_s$ (cf. (2.71)) and it is defined as

$$\mathbf{P} = J_a\mathbf{F}\mathbf{F}_a^{-1}\mathbf{S}_n\mathbf{F}_a^{-T}, \tag{2.86}$$

where $\mathbf{S}_n = \mathbf{S}_n(\mathbf{F}_e) = \mathbf{S}_n(\mathbf{F}_s\mathbf{F}_a^{-1})$ is the second Piola-Kirchhoff stress tensor measured with respect to the natural configuration $\mathscr{B}_n$. From (2.86), we can define also the tensor $\mathbf{S} = J_a\mathbf{F}_a^{-1}\mathbf{S}_n\mathbf{F}_a^{-T}$. For the constitutive part we will assume to work with a neo-Hookean material law presented in Sect. 1.1.6. Following the steps presented in Sect. 2.2.4, the Piola-Kirchhoff tensor have the form

$$\mathbf{S} = J_a\left(\mu_n\left(\mathbf{C}_a^{-1} - \mathbf{C}^{-1}\right) + \lambda_n\log\left(\frac{J}{J_a}\right)\mathbf{C}^{-1}\right)$$

and

$$\mathbf{P} = J_a\left(\mu_n\left(\mathbf{F}\mathbf{C}_a^{-1} - \mathbf{F}^{-T}\right) + \lambda_n\log\left(\frac{J}{J_a}\right)\mathbf{F}^{-T}\right).$$

The tensor $\mathbf{K}_r$ is the material form of the tensor of hydraulic conductivity, i.e., $\mathbf{K}_r = J\mathbf{F}^{-1}\mathbf{K}\mathbf{F}^{-T}$. Tensor $\mathbf{K}$ is taken from [HM90] and adapted to our framework, i.e.

$$\mathbf{K} = k_0\left[\frac{\phi_{s0}}{1-\phi_{s0}}\frac{J-\phi_{sn}J_a}{J_a\phi_{sn}}\right]^{m_0}\exp\left\{\frac{m_1}{2}\left[\frac{J^2-J_a^2}{J_a^2}\right]\right\}\mathbf{g}^{-1}. \tag{2.87}$$

The numbers m_0 and m_1 featuring in (2.87) are material parameters. The formula of the hydraulic conductivity given in (2.87) defines an isotropic tensor. However, other forms of hydraulic conductivity, which account for tissue anisotropy, have been recently proposed in [AW10][FG12a].

To close the problem, $\mathbf{L}_a$ should be supplied by one of the formulae (2.82)–(2.84).

In the absence of growth, the field equations (2.85) were studied in [EM01] in the case of a linear viscoelastic biphasic model for soft tissues.

The unknowns of the equation system (2.85) are the displacement $\mathbf{u}$, the material pressure π, and the tensor $\mathbf{F}_a$. Since only $\mathbf{F}_a$ and $\mathbf{u}$ are derived in time in this equation system, initial conditions have to be provided only for these variables and we denote by

$$\begin{array}{rcl} \mathbf{u}(\mathbf{X},0) & = & \mathbf{u}_0(\mathbf{X}), \\ \mathbf{F}_a(\mathbf{X},0) & = & \mathbf{F}_{a0}(\mathbf{X}). \end{array} \tag{2.88}$$

Equations in (2.85) hold in the internal points of $\mathscr{B}_r$ and have to be completed with conditions prescribed on the boundary $\partial\mathscr{B}_r$. For each unknown, the boundary $\partial\mathscr{B}_r$ is split into a Dirichlet- and a Neumann-type subset. This means that $\partial\mathscr{B}_r$ admits the representations $\partial\mathscr{B}_r = \Gamma_N^{\mathbf{u}} \cup \Gamma_D^{\mathbf{u}}$ and $\partial\mathscr{B}_r = \Gamma_N^{\pi} \cup \Gamma_D^{\pi}$, where $\Gamma_N^{\mathbf{u}}$ and Γ_N^{π} are the subsets of $\partial\mathscr{B}_r$ on which Neumann boundary conditions for the displacements and pressure are prescribed, while $\Gamma_D^{\mathbf{u}}$ and Γ_D^{π} are the subsets on which Dirichlet conditions are supplied. Formally, boundary conditions are written as

$$\begin{cases} \mathbf{u} = \mathbf{u}_b, & \text{on } \Gamma_D^{\mathbf{u}}, \\ \big[-J\mathbf{g}^{-1}\pi\mathbf{F}^{-T} + \mathbf{P}\big]\mathbf{N} = \mathbf{f}_{rb}, & \text{on } \Gamma_N^{\mathbf{u}}, \\ \pi = \pi_b, & \text{on } \Gamma_D^{\pi}, \\ \big[-\mathbf{K}_r \mathrm{Grad}(\pi)\big]\cdot\mathbf{N} = Q_{rb}, & \text{on } \Gamma_N^{\pi}, \end{cases} \tag{2.89}$$

where $\mathbf{N}$ is the unit vector normal to $\partial\mathscr{B}_r$. The surface force $\mathbf{f}_{rb}$ is defined per unit area of the reference boundary $\Gamma_N^{\mathbf{u}}$ and is, thus, generally different from the force $\mathbf{f}_b$ associated with the actual configuration $\mathscr{B}_t$. An analogous argument holds true for the quantities Q_{rb} and Q_b, the latter being the flux prescribed per unit are of the boundary $\partial\mathscr{B}_t$ of the actual configuration. The pairs $\mathbf{f}_{rb}$-$\mathbf{f}_b$ and Q_{rb}-Q_b are related to each other by the formulae [Bon08]

$$\mathbf{f}_{rb} = \mathbf{f}_b J\sqrt{\mathbf{N}\otimes\mathbf{N} : \mathbf{C}^{-1}}, \quad \mathbf{Q}_{rb} = \mathbf{Q}_b J\sqrt{\mathbf{N}\otimes\mathbf{N} : \mathbf{C}^{-1}}. \tag{2.90}$$

Since the displacement $\mathbf{u}$ is a vectorial unknown, we can also prescribe different boundary conditions for its components u^k. Hence, we also suppose that $\partial\mathscr{B}_r$ admits also the representations $\partial\mathscr{B}_r = \Gamma_N^{u^k} \cup \Gamma_D^{u^k}$ with $k = 1, 2, 3$. In this case, boundary conditions for the displacement are written as

$$\begin{cases} u^k = u_b^k, & \text{on } \Gamma_D^{u^k}, \\ (\big[-J\pi\mathbf{F}^{-T} + \mathbf{P}\big]\mathbf{N})^k = f_{rb}^k, & \text{on } \Gamma_N^{u^k}. \end{cases}$$

For each component, a formula similar to (2.90.1) holds true.

Chapter 3

Variational formulations and solution methods

This chapter is devoted to describe the discretization and the solution methods that will be employed for the simulation of poroelasticity systems considered in this thesis. First of all, we introduce some essential concepts of functional analysis and the Finite Element Method (FEM), which will be used for the spatial discretization of the poroelastic model. From a temporal standpoint, a poroelasticity problem is a system of Differential Algebraic Equations (DAEs). Hence, some basic definitions on DAEs are given. Then we describe boundary value problems in non-linear elasticity and Newton's method. Since the application of Newton's method to poroelastic problems gives rise to a sequence of saddle-point systems, we finally present some stability results for these problems and the discretization thereof, i.e. the Mixed Finite Element Method.

3.1 Variational formulation of partial differential equations

In this first section, we present the formal setting of the weak formulation of Partial Differential Equations (PDEs), that is a necessary prerequisite for the introduction of the Finite Element (FE) discretizations of PDEs. Our notation is quite standard and is based on [Sal08, Qua09]. For a more detailed literature, the reader may refer to any PDE textbook [Eva97] or functional analysis textbook.

3.1.1 Function spaces

Let X be a linear space over $\mathbb{R}$. We denote by $\|\cdot\|_X : X \to \mathbb{R}$ a *norm* on X and by $|\cdot|_X : X \to \mathbb{R}$ a *seminorm* on X. Moreover, we denote by $(\cdot,\cdot)_X : X \times X \to \mathbb{R}$ a *scalar product* on X. A scalar product naturally induces a norm, given by

$$\|x\|_X = \sqrt{(x,x)_X}.$$

A norm naturally induces a distance between elements of X. In fact, it is always possible to define a metric in X as

$$d_X(x,y) = \|x-y\|_X.$$

A distance allows to define a topology and, hence, also convergence of sequence in X.

Given an integer $d \geq 1$, let $\Omega \subset \mathbb{R}^d$ be a bounded domain with Lipschitz boundary $\partial\Omega$ and outer normal vector $\mathbf{n}$. Elements of Ω are denoted by $\mathbf{x} = (x_1, x_2, \ldots, x_d)$. We suppose that the boundary can be split into two subsets Γ_D and Γ_N such that $\overline{\Gamma}_D \cup \overline{\Gamma}_N = \partial\Omega$ and $\dot{\Gamma}_D \cap \dot{\Gamma}_N = \emptyset$.

A *multi-index* is a d-tuple $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots \alpha_d)$ of non-negative integers. The length of a multi-index is $|\boldsymbol{\alpha}| = \sum_{i=1}^d \alpha_i$. The $\boldsymbol{\alpha}$-th derivative of order $m = |\boldsymbol{\alpha}|$ is denoted by

$$D^{\boldsymbol{\alpha}} = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}} \cdots \frac{\partial^{\alpha_d}}{\partial x_d^{\alpha_d}}.$$

We use the standard notation of functional analysis to denote the following spaces of scalar functions:

- $C^k(\overline{\Omega})$ is the linear space of functions that are k times continuously differentiable in Ω. It is a Banach space under the norm

$$\|f\|_{C^k} = \max_{\overline{\Omega}} |f| + \sum_{1 \leq |\boldsymbol{\alpha}| \leq k} \max_{\overline{\Omega}} |D^{\boldsymbol{\alpha}} f|.$$

- $L^p(\Omega)$ with $1 \leq p < \infty$ denotes the space of p-summable functions in Ω. It is a Banach space under the norm

$$\|f\|_{L^p} = \left(\int_\Omega |f|^p \right)^{\frac{1}{p}}.$$

- $L^\infty(\Omega)$ denotes the space of essentially bounded functions in Ω. It is a Banach space under the norm

$$\|f\|_{L^\infty} = \operatorname{ess\,sup}_\Omega |f|.$$

- $W^{k,p}(\Omega) = \{f \in L^p(\Omega) : D^{\boldsymbol{\alpha}} f \in L^p(\Omega), \quad \forall \boldsymbol{\alpha} : |\boldsymbol{\alpha}| \leq k\}$ with $1 \leq p \leq \infty$. It is a Banach space under the norm

$$\|f\|_{W^{k,p}} = \|f\|_{L^p} + \sum_{1 \leq |\boldsymbol{\alpha}| \leq k} \|D^{\boldsymbol{\alpha}} f\|_{L^p}.$$

  which is by definition a Sobolev space.

- $H^k(\Omega) = W^{k,2}(\Omega)$. The space $H^p(\Omega)$ is a Hilbert space with respect to the scalar product

$$(f,g)_{H^k} = \int_\Omega f\, g + \sum_{1\le|\boldsymbol{\alpha}|\le k} \int_\Omega (D^\alpha f)(D^\alpha g).$$

- $H^{1/2}(\Gamma_N) = \{u|_{\Gamma_N} : u \in H^1(\Omega)\}$.

In particular, notice that

- $L^2(\Omega) = H^0(\Omega)$;
- $(f,g)_{H^1} = (f,g)_{L^2} + (\nabla f, \nabla g)_{L^2}$.

Two important subspaces of $H^1(\Omega)$ are:

- $H^1_{\Gamma_D}(\Omega) = \{v \in H^1(\Omega) : v|_{\Gamma_D} = 0\}$;
- $H^1_0(\Omega) = H^1_{\Gamma_D}(\Omega)$ if $\Gamma_D = \partial\Omega$.

For the space $H^1_{\Gamma_D}(\Omega)$, the equality $v|_{\Gamma_D} = 0$ has to be intended in a weak sense. In this function space, it is possible to prove the Poincarè-Friedrichs inequality

$$\|f\|_{L^2} \le C_P \|\nabla f\|_{L^2},$$

where C_P is the Poincarè constant that depends on the domain size. This inequality allows to prove that $\|f\|_{H^1}$ is equivalent to $\|\nabla f\|_{L^2}$ and that $(\nabla f, \nabla g)_{L^2}$ is a scalar product in $H^1_{\Gamma_D}(\Omega)$.

For the spaces $W^{k,p}$, a natural semi-norm is defined as

$$|f|_{W^{k,p}} = \sum_{|\boldsymbol{\alpha}|=k} \|D^{\boldsymbol{\alpha}} f\|_{L^p}.$$

The definitions given in this section can be naturally extended to spaces of vector functions with image in $\mathbb{R}^n$. As an example, the space $W^{k,p}(\Omega;\mathbb{R}^n)$ denotes the linear space of functions $\mathbf{v} : \Omega \to \mathbb{R}^n$, whose components belong to $W^{k,p}(\Omega)$.

3.1.2 Operators, functionals and dual

Given two linear function spaces V and W, we denote by $\mathscr{L}(V,W)$ the set of continuous linear operators $L : V \longrightarrow W$. The set $\mathscr{L}(V,W)$ itself is a linear space and a Banach space under the norm

$$\|L\|_{\mathscr{L}(V,W)} = \sup_{\|v\|_V=1} \|Lv\|_W.$$

A linear operator L is bounded if there exists a constant M such that

$$\|Lv\|_W \leq M\|v\|_V \quad \forall v \in V.$$

If $W = \mathbb{R}$, the linear operator $F : V \longrightarrow \mathbb{R}$ is called *linear functional* and the set $V^* = \mathscr{L}(V, \mathbb{R})$ is said *dual* space of V. A linear functional is bounded if and only if it is continuous. Given a functional $F \in V^*$, application of F to a function $v \in V$ will be denoted by Fv, $F(v)$, or $\langle F, v\rangle_{V^* \times V}$.

3.1.3 Bilinear forms

Let V and W be two Banach spaces. A bilinear form a in $V \times W$ is a function

$$a(\cdot, \cdot) : V \times W \longrightarrow \mathbb{R}$$

that is linear with respect to both its arguments. A bilinear form is continuous, or bounded, if there exists an $M > 0$ such that

$$a(v, w) \leq M\|v\|_V \|w\|_W, \quad \forall v \in V, \forall w \in W.$$

If $V = W$, a is said to be a bilinear form on V; this can be seen as a $(2, 0)$ tensor on V. A bilinear form of this kind is

- symmetric, if
$$a(u, v) = a(v, u) \quad \forall u, v \in V;$$
- coercive, if there exists an $\alpha > 0$ such that
$$a(v, v) > \alpha\|v\|_V^2 \quad \forall v \in V, \quad \text{with } v \neq 0.$$

3.1.4 Abstract variational problems

Given a Hilbert space V, a functional $F \in V^*$, and a bilinear form a in V, an *abstract variational problem* is a problem of the kind:

Find $u \in V$ such that

$$a(u, v) = F(v), \quad \forall v \in V. \tag{3.1}$$

A problem of this form is said to be *well-posed* if it has a unique solution that depends continuously on the data. The classical stability result for problem (3.1) is the Lax-Milgram theorem. It states that if

D1: a is bounded with continuity constant M,

D2: a is coercive with coercivity constant α,

D3: F is bounded,

then the variational problem (3.1) has a unique solution and the following stability estimate holds

$$\|u\|_V \leq \frac{1}{\alpha}\|F\|_{V^*}. \tag{3.2}$$

Problems of the form (3.1) are the abstract setting for the *weak*, or *variational*, *formulation* of elliptic Partial Differential Equations (PDEs). A PDE written in weak form does not hold at each point of the domain but it is defined in integral form against certain test functions. This weak integral form, in contrast to the strong one, requires weaker hypotheses on the data and hence allows for a more general definition of a solution. If the data of the problem are regular enough, strong and weak solutions coincide.

As an example, in the next paragraph, we derive the weak formulation of a Laplace equation completed with Dirichlet and Neumann boundary conditions.

Weak formulation of Laplace equation

Let us consider the following differential problem:

$$\begin{cases} -\mathrm{Div}(k\nabla u) & = & f & \text{in } \Omega \\ u & = & 0 & \text{on } \Gamma_D \\ k\nabla u \cdot \mathbf{n} & = & h & \text{on } \Gamma_N \end{cases} \tag{3.3}$$

with $f \in L^2(\Omega)$, $h \in H^{-1/2}(\Gamma_N) := (H^{1/2}(\Gamma_N))^*$, and $k \in L^\infty(\Omega)$. We also suppose that there exists a $k_0 > 0$ such that $k \geq k_0$ a.e. in Ω. Problem (3.3) is written in strong form and the hyphoteses given on the data are not sufficient to ensure the existence of a solution.

In order to obtain the weak form of (3.3), we multiply (3.3.1) by a test function $v \in H^1_{\Gamma_D}(\Omega)$ and integrate over Ω, to obtain

$$-\int_\Omega \mathrm{Div}(k\nabla u)v = \int_\Omega f v.$$

Applying integration by part at the left hand side of the previous equation, we get

$$\int_\Omega k\nabla u \cdot \nabla v - \int_{\partial\Omega} k(\nabla u \cdot \boldsymbol{n})v = \int_\Omega f v. \tag{3.4}$$

Using the additivity of the integral, we can write the boundary integral as

$$\int_{\partial\Omega} k(\nabla u \cdot \boldsymbol{n})v = \int_{\Gamma_D} k(\nabla u \cdot \boldsymbol{n})v + \int_{\Gamma_N} k(\nabla u \cdot \boldsymbol{n})v.$$

The integral over Γ_D vanishes because of the choice of the test function v, while in the integral over Γ_N we can substitute the value of the *co-normal* derivative (3.3.3). Hence, we can write (3.4) as

$$\int_\Omega k\nabla u \cdot \nabla v = \int_\Omega f v + \int_{\Gamma_N} hv. \tag{3.5}$$

Equation (3.5) requires weaker hypotheses: it involves only one derivative of u and, in order for the integrals to be well-defined, it is enough that $u, v \in H^1_{\Gamma_D}(\Omega)$.

Problem (3.3) has been written as an abstract variational problem (3.1), where

- $V = H^1_{\Gamma_D}(\Omega)$,
- $a(u,v) = \int_\Omega k\nabla u \cdot \nabla v$,
- $F(v) = \int_\Omega f v + \int_{\Gamma_N} hv$.

3.2 Finite Element Method

The Finite Element Method (FEM) is the most used discretization method for the numerical solution of physical problems that can be described by PDEs [Qua09, Bra09, Hug09]. FEM belongs to the more general class of Galerkin methods. These latter are based on the *weak formulation* of a considered PDE. The basic idea of Galerkin methods is to formulate the weak problem in a suitable approximation V_h of the space V to which the solution belongs. Before presenting the FEM, we report some general results that hold true for all Galerkin methods.

3.2.1 Galerkin approximation

We consider now a problem of the form (3.1), where V is a Hilbert space. We also suppose that hypotheses **D** hold. A Galerkin method for the numerical approximation of (3.1) consists in finding an approximate solution $u_h \in V_h$, where V_h is a family of subspaces of V that depends on a positive discretization parameter h. We define $N_h = \dim(V_h)$. Hence the Galerkin approximation of (3.1) reads

Find $u_h \in V_h$ such that

$$a(u_h, v_h) \quad = \quad F(v_h) \quad \forall v_h \in V_h \tag{3.6}$$

If we denote with $\{\varphi_i\}$ a basis of V_h, the problem (3.6) can also be written as

Find $u_h \in V_h$ such that

$$a(u_h, \varphi_i) \quad = \quad F(\varphi_i) \quad \text{for } i = 1, 2, \ldots, N_h. \tag{3.7}$$

The approximate solution can be expanded in the basis $\{\varphi_i\}$ as

$$u_h(\boldsymbol{x}) = \sum_{j=1}^{N_h} u_j \varphi_j(\boldsymbol{x}),$$

where u_j are unknown coefficients. Hence, using the bilinearity of a, the equation (3.7) becomes

$$\sum_{j=1}^{N_h} u_j a(\varphi_j, \varphi_i) = F(\varphi_i) \quad \text{with } i = 1, 2, \ldots, N_h. \tag{3.8}$$

We denote by

- $\mathbf{A}$ the matrix whose entries are $\mathbf{A}_{ij} = a(\varphi_j, \varphi_i)$,
- $\underline{f}$ the vector whose elements are $f_i = F(\varphi_i)$,
- $\underline{u}$ the vector whose elements are the coefficients u_j.

The matrix $\mathbf{A}$ is called *stiffness* matrix. It is

- positive definite if a is coercive;
- symmetric if a is symmetric.

Thus, the problem (3.8) can be written as the linear system

$$\mathbf{A}\underline{u} = \underline{f}. \tag{3.9}$$

Analysis of discretized problem

The Galerkin problem admits a unique solution if the stiffness matrix is not singular. This trivially holds if the weak form is coercive. In this case, the stability estimate (3.2) of the continuous problem holds true also for the discrete problem. This is ensured since V_h is a subspace of V. Hence, for the discrete problem, the following bound holds

$$\|u_h\|_V \leq \frac{\|F\|_{V^*}}{\alpha}.$$

Convergence of the discrete solution to the continuous solution for $h \to 0$ is guaranteed by Cea's lemma. It states that

$$\|u - u_h\|_V \leq \frac{M}{\alpha} \inf_{v_h \in V_h} \|u - v_h\|_V. \tag{3.10}$$

3.2.2 Finite element discretization

The FEM is a particular Galerkin method, in which the space V_h is chosen as a subset of a space of polynomial interpolation.

Let $\mathscr{T}_h$ be a face-to-face *triangulation* of the domain Ω, i.e., a generalized coverage of Ω with *elements* that are

- intervals if $d = 1$,
- triangles or quadrilaterals if $d = 2$,
- tetrahedra or hexahedra if $d = 3$.

The elements are required to be non-overlapping, i.e., $\dot{K} \cap \dot{H} = \emptyset$ for two elements $K, H \in \mathscr{T}_h$, with $K \neq H$. If the domain is polytopic, the elements cover the domain, i.e.

$$\Omega = \overline{\dot{\bigcup_{K \in \mathscr{T}_h}} K}.$$

If the domain is not polytopic, the triangulation defines a restricted domain

$$\Omega_h := \overline{\dot{\bigcup_{K \in \mathscr{T}_h}} K} \subset \Omega.$$

We define h_K the diameter of K and $h = \max(h_K)$. We now introduce the interpolation space

$$X_h^r = \{v_h \in C^0(\overline{\Omega}) : v_{h|K} \in \mathbb{P}_r, \forall K \in \mathscr{T}_h\}.$$

The FE discretization consists in choosing in (3.6) the discrete space

$$V_h^r = V \cap X_h^r. \tag{3.11}$$

To distinguish FE discretizations from generic Galerkin discretizations we use the symbol V_h^r. The superscript r denotes the order of polynomials that have been used to approximate functions in V.

Accuracy of FEM depends on the degree r of the interpolation space and on the regularity of the data. The order of convergence can be obtained substituting at the right hand side of (3.10) the discrete solution u_h instead of v_h. Assuming, for example, that $V = H^1_{\Gamma_D}(\Omega)$ and $u \in H^{p+1}(\Omega)$, the interpolation error can be bounded by

$$\|u - u_h\|_{H^1} \leq Ch^s |u|_{H^{s+1}}$$

with $s = \min(p, r)$, where C is a positive constant [QSS00]. Using this bound in (3.10), we obtain the following convergence result

$$\|u - u_h\|_{H^1} \leq \frac{M}{\alpha} Ch^s |u|_{H^{s+1}}.$$

Hence, s is the accuracy of FEM.

3.3 Implicit Euler method for Differential-algebraic equations

Differential-Algebraic Equations (DAEs) are a generalization of systems of ordinary differential equations. For two vector valued functions $\underline{x} = \underline{x}(t) : (0, T_{fin}] \to \mathbb{R}^n$ and $\underline{y} = \underline{y}(t) : (0, T_{fin}] \to \mathbb{R}^m$, we consider a DAE of the following kind:

$$\begin{cases} \dot{J}(\underline{x}) &= F(\underline{x}, \underline{y}) \\ \underline{0} &= G(\underline{x}, \underline{y}) \end{cases} \tag{3.12}$$

with

- $J : \mathbb{R}^n \to \mathbb{R}^n$;
- $F : \mathbb{R}^{n+m} \to \mathbb{R}^n$;
- $G : \mathbb{R}^{n+m} \to \mathbb{R}^m$.

DAEs allow also for more general definitions, for example F and G may also be time dependent, i.e. *non-autonomous*. We focus on the case (3.12), since the poroelastic system takes this form. In the linear poroelastic case, J is the identity.

The variable $\underline{x}$ is the *differential* component of the solution and the variable $\underline{y}$ is the *algebraic* solution. The same terminology applies to F and G respectively. The differential equation (3.12.1) determines the time evolution of the differential component, while the algebraic equation (3.12.2) acts as a constraint on the evolution of both components.

In order to be well-defined, the system (3.12) has to be completed with initial conditions

$$\underline{x}(0) = \underline{x}_0. \tag{3.13}$$

For the numerical integration of the DA problem (3.12) and (3.13), several algorithms have been developed and some relevant ones may be found in [Gea84, BCP95]. Here we observe that, in general, explicit methods are not appropriate for this problem because

- they would require an initial condition on $\underline{y}$ and this is not always available or well-defined;
- the constraint would be violated.

In order to introduce a family of numerical approximations of such a problem, we consider the time interval $\mathscr{I} = (0, t_f]$ and we introduce on it a uniform partition into N sub-intervals (t^{m-1}, t^m) of size τ. We define $\underline{x}^m$ as the approximation of $\underline{x}(t^m)$ at t^m.

Using an implicit Euler scheme [QSS00], the discretized problem at time step t^m reads:

$$\begin{cases} J^m(\underline{x}) - \tau F(\underline{x}^m, \underline{y}^m) &= J^{m-1}(\underline{x}) \\ G(\underline{x}^m, \underline{y}^m) &= \underline{0}. \end{cases} \tag{3.14}$$

Each time step involves the solution of a non-linear problem of the form (3.14), which itself requires the application of an iterative strategy, such as Newton-like methods.

3.4 Non-linear problems

In the previous section, we considered only linear problems. Instead, as shown in the first two chapters, problems arising in continuum mechanics are inherently non-linear. The study of stability of such a problem requires a different formal setting. This is not the focus of this work and hence we refer to [MH94]. The goal of this section is to introduce the weak formulation of non-linear elasticity, also known as the *principle of virtual work*. A detailed treatment of it may be found in [MH94]. Then, we present Newton's algorithm that will be used for the numerical solution of non-linear elasticity. It can be found in several optimization books, for example in [NW06]. Since the derivation of this method requires the differentiability in Banach spaces, we first of all give the definition of Gateaux differential [LP03].

3.4.1 Gâteaux differential

Let V and W be two Banach spaces. A function $f : V \to W$ is said to be Gâteaux differentiable at $\boldsymbol{u} \in V$ in direction $\mathbf{v} \in V$ if there exists a bounded linear operator $T_{\mathbf{u}} \in \mathscr{L}(V, W)$ such that the following equality holds

$$\lim_{\epsilon \to 0} \frac{f(\mathbf{u} + \epsilon \mathbf{v}) - f(\mathbf{u})}{\epsilon} = T_{\mathbf{u}}\mathbf{v}.$$

f is Gâteaux differentiable at $\mathbf{u}$ if and only if the limit above exists for all $\mathbf{v} \in V$ and there exists a bounded linear operator $T_{\mathbf{u}} : \mathbf{v} \to T_{\mathbf{u}}\mathbf{v}$. In this case, we will denote the Gâteaux derivative as $\mathfrak{D}f(\mathbf{u})[\mathbf{v}]$.

If f is differentiable on a subset $\mathscr{U} \subset V$, we say that f is differentiable in $\mathscr{U}$. Thus, the Gâteaux derivative is a map

$$\mathscr{D}f(\,\cdot\,)[\,\cdot\,] : \mathscr{U} \times V \longrightarrow W,$$

that is non-linear with respect to the first argument and linear with respect to second argument. If $W = \mathbb{R}$, the evaluation of $\mathscr{D}f(\,\cdot\,)[\,\cdot\,]$ at a certain $\mathbf{u}$ is the following element of V^*

$$\mathscr{D}f(\mathbf{u})[\,\cdot\,] : V \longrightarrow \mathbb{R}.$$

Gâteaux differential can be naturally extended to second order derivatives as

$$\mathfrak{D}^2 f(\mathbf{u})[\mathbf{v}][\mathbf{h}] = \lim_{\epsilon \to 0} \frac{\mathfrak{D}f(\mathbf{u} + \epsilon \mathbf{h})[\mathbf{v}] - \mathfrak{D}f(\mathbf{u})[\mathbf{v}]}{\epsilon}.$$

Second order derivatives define the following map

$$\mathscr{D}f(\cdot)[\cdot][\cdot] : \mathscr{U} \times V \times V \longrightarrow W.$$

3.4.2 Principle of virtual work

We now present the derivation of the weak form of the elasto-static problem presented in Sec. 1.1.5. In continuum mechanics, the weak form of non-linear elasticity is also called *principle of virtual work*. The boundary problem associated to (1.46) is

$$\begin{cases} -\mathrm{Div}\,\mathbf{P} &= \mathbf{f} & \text{in } \mathscr{B}_r \\ \mathbf{PN} &= \mathbf{t} & \text{in } \Gamma_N \\ \mathbf{u} &= \mathbf{0} & \text{in } \Gamma_D. \end{cases} \tag{3.15}$$

In a purely elastic context, the first Piola-Kirchhoff tensor of a uniform material is a function of the displacement gradient $\mathbf{U} = \nabla\mathbf{u}$ only, and hence $\mathbf{P} = \mathbf{P}(\mathbf{U})$. The solution $\mathbf{u}$ is sought for in a subset $\mathscr{U}$ of V, where $V = W^{1,p}_{\Gamma_D}(\mathscr{B}_r, \mathbb{R}^3)$. To obtain the weak form of (3.15), we scalarly multiply (3.15.1) by a test function $\mathbf{v} \in V$. With the terminology of continuum mechanics, the test function is called *virtual displacement*. Then, we integrate over the domain to get

$$\int_{\mathscr{B}_r} -\mathrm{Div}\,\mathbf{P} \cdot \mathbf{v}\,\mathrm{d}\mathbf{X} = \int_{\mathscr{B}_r} \mathbf{f} \cdot \mathbf{v}\,\mathrm{d}\mathbf{X}.$$

Applying now integration by parts to the left-hand side, we get

$$\int_{\mathscr{B}_r} \mathbf{P} : \mathbf{V}\,\mathrm{d}\mathbf{X} = \int_{\mathscr{B}_r} \mathbf{f} \cdot \mathbf{v}\,\mathrm{d}\mathbf{X} + \int_{\Gamma_N} (\mathbf{PN}) \cdot \mathbf{v}\,\mathrm{d}\mathbf{A} + \int_{\Gamma_D} (\mathbf{PN}) \cdot \mathbf{v}\,\mathrm{d}\mathbf{A},$$

where $\mathbf{V} = \nabla\mathbf{v}$. At the right-hand side, the Dirichlet boundary integral vanishes because of the choice of the test function $\mathbf{v}$. The value $\mathbf{PN}$ on the Neumann boundary is given by condition (3.15.2). If we now define the function $\mathfrak{F}(\mathbf{u}, \mathbf{v})$ as

$$\mathfrak{F}(\mathbf{u}, \mathbf{v}) := \int_{\mathscr{B}_r} \mathbf{P}(\mathbf{U}) : \mathbf{V}\,\mathrm{d}\mathbf{X} - \int_{\mathscr{B}_r} \mathbf{f} \cdot \mathbf{v}\,\mathrm{d}\mathbf{X} - \int_{\Gamma_N} \mathbf{t} \cdot \mathbf{v}\,\mathrm{d}\mathbf{A},$$

the principle of virtual work associated with (3.15) reads

Find $\boldsymbol{u} \in \mathscr{U}$ such that

$$\mathfrak{F}(\mathbf{u}, \mathbf{v}) = 0 \quad \forall \mathbf{v} \in V \tag{3.16}$$

If the material is also hyperelastic (cf. Eq. (1.45)), the tensor $\mathbf{P}$ is the Gâteaux differential of a strain energy function $\mathscr{W}$, and hence the following equality holds

$$\mathfrak{D}\mathscr{W}(\mathbf{U})[\mathbf{V}] = \mathbf{P}(\mathbf{U}) : \mathbf{V}.$$

In this case, if $\Gamma_N = \emptyset$, the problem (3.16) is equivalent to finding in $\mathscr{U}$ a stationary point of the functional

$$I(\mathbf{u}) = \int_{\mathscr{B}_r} f(\mathbf{u}, \mathbf{U}),$$

where f is given by

$$f(\mathbf{u}, \mathbf{U}) = \mathscr{W}(\mathbf{U}) - \mathbf{f} \cdot \mathbf{u}.$$

Under these hypotheses, problem (3.16) can be written as

Find $\boldsymbol{u} \in \mathscr{U}$ such that

$$\mathscr{D}I(\mathbf{u})[\mathbf{v}] \quad = \quad 0 \quad \forall \mathbf{v} \in V.$$

Theorems on existence and uniqueness of the solution of this minimization problem can be found in [Bal76, Tal94].

3.4.3 Newton's method

We describe now the Newton's method for the solution of problems of the form (3.16). The idea of Newton's method is to start from an initial guess $\mathbf{u}^0$ and to generate a sequence $\mathbf{u}^k \in V$ such that

$$\lim_{k \to \infty} \mathbf{u}^k = \mathbf{u}.$$

At the k-th step of the algorithm, the k-th element is obtained from the previous one as

$$\mathbf{u}^k = \mathbf{u}^{k-1} + \mathbf{h}^k,$$

where the increment $\mathbf{h}^k$ is given by the solution of the linearized problem

$$\mathfrak{F}(\mathbf{u}^{k-1}, \mathbf{v}) + \mathfrak{D}_{\mathbf{u}}\mathfrak{F}(\mathbf{u}^{k-1}, \mathbf{v})[\mathbf{h}^k] = 0, \quad \forall \mathbf{v} \in V. \tag{3.17}$$

In an elasticity context, the first derivative of $\mathscr{F}$ corresponds to the second derivative of the strain energy function $\mathscr{W}$. Hence, we define the following bilinear form

$$a^k(\mathbf{h}, \mathbf{v}) := \mathfrak{D}\mathfrak{F}(\mathbf{u}^k, \mathbf{v})[\mathbf{h}] = \mathfrak{D}^2 I(\mathbf{u}^k)[\mathbf{v}][\mathbf{h}].$$

At each step of the Newton's method, the following linear variational problem has to be solved:

Find $\mathbf{h}^k \in V$ such that

$$a^{k-1}(\mathbf{h}^k, \mathbf{v}) \quad = \quad -\mathfrak{F}(\mathbf{u}^{k-1}, \mathbf{v}), \quad \forall \mathbf{v} \in V. \tag{3.18}$$

In a FE context, Eq. (3.18) becomes

Find $\mathbf{h}_h \in V_h$ such that

$$a_h^{k-1}(\mathbf{h}_h^k, \mathbf{v}_h) \quad = \quad -\mathfrak{F}(\mathbf{u}_h^{k-1}, \mathbf{v}_h) \quad \forall \mathbf{v}_h \in V_h, \tag{3.19}$$

where $V_h \subset V$ is a FE space and

$$a_h^k(\mathbf{h}_h, \mathbf{v}_h) := \mathfrak{D}\mathfrak{F}(\mathbf{u}_h^k, \mathbf{v}_h)[\mathbf{h}_h].$$

Convergence

Newton's method is locally second order convergent, i.e. there exists a positive constant M_N such that

$$\|\mathbf{u} - \mathbf{u}^k\|_V \leq M_N \|\mathbf{u} - \mathbf{u}^{k-1}\|_V^2.$$

Quasi-Newton method

In practical cases, the assembling of the linearized discrete problem (3.19) can be very expensive. Hence, it is possible to substitute the bilinear form a_h^k with an approximated one $\hat{a}_h^k$. In this case the method is called a quasi-Newton method.

Stopping criteria

There are two possible strategies to stop Newton's method, namely the control of residuals and the control of increments. In the former, the algorithm stops when

$$\|\mathscr{F}(\mathbf{u}^k, \cdot)\| \leq \epsilon$$

for a given $\epsilon > 0$. In the latter, the algorithm stops if

$$\|\mathbf{u}^k - \mathbf{u}^{k-1}\|_V \leq \epsilon.$$

Line search

An improvement that can be done in the Newton's method is to modify the update of the solution as follow

$$\mathbf{u}^k = \mathbf{u}^{k-1} + \alpha^k \mathbf{h}^k,$$

where $0 < \alpha^k \leq 1$. This choice has been found to be particularly useful in the early steps of the algorithm, when the phenomenon of overshooting may affect or prevent the convergence.

3.5 Saddle-point problems

In this section, we first introduce two model problems whose variational formulation is cast in the framework of saddle-point (SP) problems. The first one is the incompressible linear elasticity. It is the classical example of a SP system in mechanics. The same equations model also the motion of a fluid for large Reynold numbers. In this case, it is called Stokes problem. The second model is the linear poroelasticity. It is the linearized version of the problem considered in this work and derived in Sec. 1.2.

Then, we discuss the stability of the abstract classes to which these two problems belong. Finally, we present the Galerkin discretization of them. In particular, for the FEM, we discuss some stable classes of FE spaces and a particular stabilization for unstable FE spaces.

3.5.1 Variational formulation of saddle-point problems in linear continuum mechanics

The d-dimensional elasticity problem

As seen in the first chapter, the goal of elasto-static is to determine the tridimensional displacement of a continuum under the effect of external traction and forces. In this subsection, we extend equations (1.71) and (1.83) to the d-dimensional case but we assume the hypothesis of small displacement. This assumption allows to write the continuity equation on the reference configuration. Let d be a positive integer and let $\Omega \subset \mathbb{R}^d$ be a bounded domain. Here, Ω represents the reference configuration on which the equations are defined. The displacement is a function

$$\Omega \ni \mathbf{X} \longrightarrow \mathbf{u}(\mathbf{X}) \in \mathbb{R}^d.$$

In $d = 1$, the fact of being in a monodimensional case is emphasized by removing bold letters for vectorial and tensorial quantities, e.g. the displacement is $u(X)$.

Considering only the linear terms, we can define the linearized strain

$$\boldsymbol{\epsilon}(\mathbf{u}) = \frac{\nabla \mathbf{u} + \nabla \mathbf{u}^T}{2}. \tag{3.20}$$

The trace of this second order tensor is the divergence of the vector $\mathbf{u}$, i.e.

$$\mathrm{tr}(\boldsymbol{\epsilon}(\mathbf{u})) = \mathrm{Div}\ \mathbf{u}.$$

We also define the deviator $\mathrm{Dev}(\mathbf{A})$ of a second order tensor $\mathbf{A}$ as

$$\mathrm{Dev}(\mathbf{A}) := \mathbf{A} - \frac{1}{d}\mathrm{tr}(\mathbf{A})\mathbf{I}.$$

We also use the symbol $\mathbf{A}^D$ to denote the deviator of $\mathbf{A}$. The deviatoric part of a tensor has by definition a vanishing trace. In linear elasticity, the stress tensor is usually written as

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}(\mathbf{u}) = 2\mu\boldsymbol{\epsilon}(\mathbf{u}) + \lambda \mathrm{tr}(\boldsymbol{\epsilon}(\mathbf{u}))\mathbf{I},$$

where $\mu > 0$ and λ are the shear modulus and the second Lamé parameter, respectively. We use a different formulation obtained by splitting the deviatoric and the isochoric components of the stress. Hence $\boldsymbol{\sigma}$ takes the form

$$\boldsymbol{\sigma} = 2\mu\boldsymbol{\epsilon}^D(\mathbf{u}) + G\mathrm{tr}(\boldsymbol{\epsilon}(\mathbf{u}))\mathbf{I}, \tag{3.21}$$

where the coefficient $G = 2\mu/d + \lambda$ is the bulk modulus. It is a positive number and measures the resistance of the material to compression and dilation. Using the notation introduced in the first chapter, the linear elasto-static problem reads:

$$\left\{ \begin{array}{rcll} -\mathrm{Div}(\boldsymbol{\sigma}) & = & \mathbf{f} & \text{in } \Omega \\ \mathbf{u} & = & \mathbf{0} & \text{on } \Gamma_D \\ \boldsymbol{\sigma}\mathbf{N} & = & \mathbf{t} & \text{on } \Gamma_N. \end{array} \right. \tag{3.22}$$

Multiplying (3.22.1) by a test function $\mathbf{v} \in V := H^1_{\Gamma_D}(\mathscr{B}_r, \mathbb{R}^d)$ and following the procedure shown in Sec. 3.4.2, we obtain the variational formulation of the linear elastic problem

Find $\mathbf{u} \in V$ such that

$$2\mu \int_\Omega \boldsymbol{\epsilon}^D(\mathbf{u}) : \boldsymbol{\epsilon}^D(\mathbf{v}) \quad + \quad G\int_\Omega \mathrm{tr}(\boldsymbol{\epsilon}(\mathbf{u}))\,\mathrm{tr}(\boldsymbol{\epsilon}(\mathbf{v})) \quad = \quad F(\mathbf{v}) \qquad \forall \mathbf{v} \in V \tag{3.23}$$

where

$$F(\mathbf{v}) = \int_\Omega \mathbf{f}\cdot\mathbf{v} + \int_{\Gamma_N} \mathbf{t}\cdot\mathbf{v}.$$

Problem (3.23) is equivalent to the minimization problem

$$\inf_{\mathbf{v}\in V} J(\mathbf{v})$$

where

$$J(\mathbf{v}) = \mu \int_\Omega \boldsymbol{\epsilon}^D(\mathbf{v})^2 + \frac{1}{2}G\int_\Omega \mathrm{tr}(\boldsymbol{\epsilon}(\mathbf{v}))^2 - F(\mathbf{v}).$$

Saddle-point formulation of incompressible elasticity

Formulation (3.23) is not convenient for the simulation of incompressible materials, i.e. if $G \to \infty$. It is not possible to directly set $G = \infty$. At the continuous level, it is possible

to include the arising constraint Div $\mathbf{u} = 0$ in the definition of the space V. This choice is not suitable for the FE discretization since it is very hard to generate spaces of discrete functions that are divergence free.

To obtain an alternative formulation, we define the pressure as $p = -G\,\mathrm{Div}\,\mathbf{u}$, and hence the stress becomes

$$\boldsymbol{\sigma} = 2\mu\boldsymbol{\epsilon}^D(\mathbf{u}) - p\mathbf{I}, \tag{3.24}$$

Thus, problem (3.23) may be written as

$$\left\{\begin{array}{ccccccc} -\mathrm{Div}(2\mu\boldsymbol{\epsilon}^D(\mathbf{u}) & - & p\boldsymbol{I}) & = & \mathbf{f} & \text{in } \Omega \\ -\mathrm{Div}\,\boldsymbol{u} & - & \frac{1}{G}p & = & 0 & \text{in } \Omega \\ \mathbf{u} & & & = & \mathbf{0} & \text{on } \Gamma_D \\ (2\mu\boldsymbol{\epsilon}^D(\mathbf{u}) & - & p\boldsymbol{I})\boldsymbol{N} & = & \mathbf{t} & \text{on } \Gamma_N. \end{array}\right. \tag{3.25}$$

The weak formulation of this problem is obtained multiplying (3.25.1) by a test displacement $\mathbf{v} \in V$ and (3.25.2) by a test pressure $q \in Q := L^2(\Omega)$ and then integrating over the domain. The weak problem reads

Find $\mathbf{u} \in V$ and $p \in Q$ such that

$$\begin{array}{ccccccc} 2\mu\int_\Omega \boldsymbol{\epsilon}^D(\mathbf{u}) : \boldsymbol{\epsilon}^D(\mathbf{v}) & - & \int_\Omega \mathrm{tr}(\boldsymbol{\epsilon}(\mathbf{v}))\,p & = & F(\mathbf{v}) & \forall \mathbf{v} \in V \\ -\int_\Omega \mathrm{tr}(\boldsymbol{\epsilon}(\mathbf{u}))\,q & - & \frac{1}{G}\int_\Omega p\,q & = & 0 & \forall q \in Q. \end{array} \tag{3.26}$$

Problem (3.26) is equivalent to find the optimality conditions of the saddle-point problem

$$\inf_{\mathbf{v}\in V}\sup_{q\in Q} L(\mathbf{v}, q),$$

where

$$L(\mathbf{v}, q) = \mu\int_\Omega \boldsymbol{\epsilon}^D(\mathbf{v})^2 - \int_\Omega \mathrm{tr}(\boldsymbol{\epsilon}(\mathbf{v}))\,q - \frac{1}{2G}\int_\Omega q^2 - F(\mathbf{v}).$$

The SP problem (3.26) is well-defined also for incompressible materials. In this case the weak form reads

Find $\mathbf{u} \in V$ and $p \in Q$ such that

$$\begin{array}{ccccccc} 2\mu\int_\Omega \boldsymbol{\epsilon}^D(\mathbf{u}) : \boldsymbol{\epsilon}^D(\mathbf{v}) & - & \int_\Omega \mathrm{tr}(\boldsymbol{\epsilon}(\mathbf{v}))\,p & = & F(\mathbf{v}) & \forall \mathbf{v} \in V \\ -\int_\Omega \mathrm{tr}(\boldsymbol{\epsilon}(\mathbf{u}))\,q & & & = & 0 & \forall q \in Q. \end{array} \tag{3.27}$$

Variational formulation of linear poroelasticity

Problem (3.26) has been written in a SP form starting from a minimization problem. The poroelasticity system has inherently a SP structure. It is a time dependent problem and $\mathscr{I}$ denotes a time interval. The unknowns of the problem are the displacement

$$(\Omega \times \mathscr{I}) \ni (\mathbf{X}, t) \longrightarrow \mathbf{u}(\mathbf{X}, t) \in \mathbb{R}^d,$$

and the pressure

$$(\Omega \times \mathscr{I}) \ni (\mathbf{X}, t) \longrightarrow \pi(\mathbf{X}, t) \in \mathbb{R}.$$

The linear poroelastic model, or Biot model, reads:

$$\left\{\begin{array}{rcrcll} -\mathrm{Div}\,(\boldsymbol{\sigma} & - & \pi\mathbf{I}) & = & \mathbf{f} & \text{in } \Omega \times \mathscr{I} \\ -\mathrm{Div}\,(\dot{\mathbf{u}}) & + & \mathrm{Div}\,(k\nabla p) & = & 0 & \text{in } \Omega \times \mathscr{I} \\ (\boldsymbol{\sigma} & - & \pi\mathbf{I})\boldsymbol{N} & = & \mathbf{t} & \text{on } \Gamma_N^{\mathbf{u}} \times \mathscr{I} \\ & & \mathbf{u} & = & \mathbf{0} & \text{on } \Gamma_D^{\mathbf{u}} \times \mathscr{I} \\ & & (k\nabla p)\cdot\mathbf{N} & = & 0 & \text{on } \Gamma_N^{\pi} \times \mathscr{I} \\ & & \pi & = & 0 & \text{on } \Gamma_D^{\pi} \times \mathscr{I} \\ & & \mathbf{u}(\cdot, 0) & = & \mathbf{u}_0(\cdot) & \text{in } \Omega. \end{array}\right. \tag{3.28}$$

The permeability k is assumed to be constant and $\boldsymbol{\sigma}$ is defined as in (3.21).

Defining the spaces $V := H^1_{\Gamma_D^{\mathbf{u}}}(\mathscr{B}_r, \mathbb{R}^3)$ and $W := H^1_{\Gamma_D^{\pi}}(\mathscr{B}_r)$, the weak formulation of (3.28) reads

At each t, find $\mathbf{u} \in V$ and $\pi \in W$ such that

$$\begin{array}{rcrcll} \int_\Omega \boldsymbol{\sigma} : \boldsymbol{\epsilon}(\mathbf{v}) & - & \int_\Omega \mathrm{tr}(\boldsymbol{\epsilon}(\mathbf{v}))\,\pi & = & F(\mathbf{v}) & \forall \mathbf{v} \in V \\ -\int_\Omega \mathrm{tr}(\boldsymbol{\epsilon}(\mathbf{u}))\,q & - & \tau \int_\Omega k\nabla\pi\cdot\nabla q & = & 0 & \forall q \in W. \end{array} \tag{3.29}$$

In order to apply the implicit Euler method to this DAE, we introduce a temporal grid on $\mathscr{I}$. The superscript m denotes the approximation of quantity A at time t^m, i.e. $A^m(\mathbf{X}) \simeq A(\mathbf{X}, t^m)$. At each time step m, the following semi-discrete in time problem has to be solved

Find $\mathbf{u} \in V$ and $p \in W$ such that

$$\begin{array}{rcrcll} \int_\Omega \boldsymbol{\sigma}^m : \boldsymbol{\epsilon}(\mathbf{v}) & - & \int_\Omega \mathrm{tr}(\boldsymbol{\epsilon}(\mathbf{v}))\,\pi^m & = & F(\mathbf{v}) & \forall \mathbf{v} \in V \\ -\int_\Omega \mathrm{tr}(\boldsymbol{\epsilon}(\mathbf{u}^m))\,q & - & \tau \int_\Omega k\nabla\pi^m\cdot\nabla q & = & G^m(q) & \forall q \in W \end{array} \tag{3.30}$$

with

$$\boldsymbol{\sigma}^m : \boldsymbol{\epsilon}(\mathbf{v}) = 2\mu\boldsymbol{\epsilon}^D(\mathbf{u}^m) : \boldsymbol{\epsilon}^D(\mathbf{v}) + G\mathrm{tr}(\boldsymbol{\epsilon}(\mathbf{u}^m))\,\mathrm{tr}(\boldsymbol{\epsilon}(\mathbf{v}))$$

and

$$G^m(q) = -\int_\Omega \mathrm{tr}(\boldsymbol{\epsilon}(\mathbf{u}^{m-1}))q.$$

3.5.2 Basic definitions for saddle-point problems

In this subsection, we summarize some essential stability results for SP problems. A detailed discussion can be found in [BF91, BBF08, Qua09]. Let V and Q be two Hilbert spaces. We consider two continuous bilinear forms a and b defined as

F1: $a : V \times V \to \mathbb{R}$, such that $|a(u,v)| < M_a\|u\|_V\|v\|_V$,

F2: $b : V \times Q \to \mathbb{R}$, such that $|b(u,q)| < M_b\|u\|_V\|q\|_Q$,

The bilinear form a defines two linear operators A and A^T such that

$$V \ni \mathbf{u} \longrightarrow A\mathbf{u} = a(\mathbf{u}, \cdot) \in V^*$$

and

$$V \ni \mathbf{u} \longrightarrow A^T\mathbf{u} = a(\cdot, \mathbf{u}) \in V^*.$$

Using duality symbol, we also write

$$\langle A\mathbf{u}, \mathbf{v}\rangle_{V^*\times V} = \langle \mathbf{u}, A^T\mathbf{v}\rangle_{V\times V^*} = a(\mathbf{u}, \mathbf{v}), \quad \forall \mathbf{u}, \mathbf{v} \in V.$$

The bilinear form b defines two operators $B : V \to Q^*$ and $B^T : Q \to V^*$, such that

$$\langle B\mathbf{u}, q\rangle_{Q^*\times Q} = \langle \mathbf{u}, B^T q\rangle_{V\times V^*} = b(\mathbf{u}, q), \quad \forall \mathbf{u} \in V, \forall q \in Q.$$

Given two functionals $F \in V^*$ and $G \in Q^*$, we define the following problem

Find $\mathbf{u} \in V$ and $p \in Q$ such that:

$$\begin{cases} a(\mathbf{u},\mathbf{v}) & +b(\mathbf{v},p) & = & F(\mathbf{v}) & \forall \mathbf{v} \in V \\ b(\mathbf{u},q) & & = & G(q) & \forall q \in Q. \end{cases} \tag{3.31}$$

Using the operators previously introduced, problem (3.31) can also be written as

$$\begin{cases} A\mathbf{u} & +B^T p & = & F & \text{in } V^* \\ B\mathbf{u} & & = & G & \text{in } Q^*. \end{cases} \tag{3.32}$$

3.5.3 Stability analysis of saddle-point problems

We will now discuss the stability of problem (3.31) and of some extensions of it. The results and theorems reported may be found in [BF91, BBF08]. We define the affine space

$$V^G = \{\mathbf{v} \in V : b(\mathbf{v}, q) = G(q) \quad \forall q \in Q\}.$$

If $G = 0$, the symbol V^0 coincides with $\ker B$. The image of B is defined as

$$\text{Im } B = \{G \in Q^* : \text{ there exists a } \mathbf{v} \in V \text{ for which } G(q) = b(\mathbf{v}, q), \forall q \in Q\}.$$

Proposition 1

The most general existence proposition for (3.31) can be obtained assuming **F1**, **F2**, and that

F3: $G \in \text{Im } B$;

F4: a is coercive on the kernel of B, i.e., there exists a constant $\alpha_0 > 0$ such that

$$a(\mathbf{v}, \mathbf{v}) \geq \alpha_0 \|\mathbf{v}\|_V^2 \quad \forall \mathbf{v} \in V^0.$$

Under these hypotheses, the first component $\mathbf{u}$ of the solution $(\mathbf{u}, p)$ of (3.31) can be proved to exist and be unique. No stability bound can be proved on $\|\mathbf{u}\|$.

Proposition 2

The existence and stability of p depend on the properties of the operator B. They can be proved if the range of B is closed, i.e.

F5: there exists a $\beta > 0$ such that

$$\sup_{q \in Q} \frac{b(\mathbf{v}, q)}{\|q\|_Q} \geq \beta \|\mathbf{v}\|_V, \quad \forall \mathbf{v} \in (V^0)^\perp. \tag{3.33}$$

The symbol $(V^0)^\perp$ denotes the orthogonal space to V^0. This hypothesis means that the operator B is bounded from below in $(V^0)^\perp$. This implies that $\text{Im } B$ is a closed set and hence B is an isomorphism between V^0 and Q^*. This hypothesis allows to extend the classical results of the finite dimensional case to the infinite dimensional one. The similar result holds true also for the mapping B^T between $(Q^0)^\perp$ and V^*. Notice that the $\ker B$ may be non-trivial even if the $\ker B^T$ is so and vice-versa.

Hence, supposing that hypotheses **F1**, **F2**, **F3**, **F4**, and **F5** hold, problem (3.31) admits a unique solution $(\mathbf{u}, p)$ that depends continuously on the data if $\ker B^T = \{0\}$. Otherwise, p is defined up to elements in the kernel.

Condition **F5** suggests that the coercivity of a on $\ker B$ is a too strong requirement. A more weak condition is

F4b: there exists a constant $\alpha_0 > 0$ such that

- $\inf_{\mathbf{v}\in\ker B} \sup_{\mathbf{u}\in\ker B} \frac{a(\mathbf{u},\mathbf{v})}{\|\mathbf{u}\|_V \|\mathbf{v}\|_V} \geq \alpha_0$;
- $\inf_{\mathbf{v}\in\ker B} \sup_{\mathbf{u}\in\ker B} \frac{a(\mathbf{v},\mathbf{u})}{\|\mathbf{u}\|_V \|\mathbf{v}\|_V} \geq \alpha_0$.

These hypotheses imply that the linear operator A is bijective on the kernel of B. Condition **F4b** can replace **F4** in order to obtain the same result with a more general condition.

3.5.4 Extensions of saddle-point problems

In some applications, such as almost incompressible elastic continua and biphasic materials, formulation (3.31) is not sufficient and a more general one has to be considered. First, we introduce a bilinear form $d : Q \times Q \to R$ such that it is

F6: continuous on Q with constant M_d;

F7: positive semi-definite on Q;

F8: symmetric.

Hence, we consider the following generalization of (3.31):

Find $\mathbf{u} \in V$ and $p \in Q$ such that:

$$\begin{cases} a(\mathbf{u},\mathbf{v}) & +b(p,\mathbf{v}) & = & F(\mathbf{v}) = & \forall \mathbf{v} \in V \\ b(\mathbf{u},q) & -\epsilon d(p,q) & = & G(q) & \forall q \in Q, \end{cases} \tag{3.34}$$

where $0 \leq \epsilon \leq 1$. We also assume that

F9: a is positive semi-definite on V.

Proposition 3

Hypotheses **F1**, **F2**, **F4b**, **F5**, **F6**, **F7**, **F8**, **F9** ensure a unique solution of problem (3.34) if the set $M = \ker B^T \cap \ker D$ is empty. If M is not empty, the solution is defined up to an element of M.

Proposition 4

We consider now a variation of (3.34) that is particularly relevant for the study of the stability of the linear poroelastic model. In this extension, the second component p of the solution is sought for in a Hilbert space W continuously embedded in Q. Given $G \in Q^*$, we consider the following weak problem

Find $\mathbf{u} \in V$ and $p \in W$ such that:

$$\begin{cases} a(\mathbf{u},\mathbf{v}) & +b(p,\mathbf{v}) & = & F(\mathbf{v}) = & \forall \mathbf{v} \in V \\ b(\mathbf{u},q) & -\epsilon d(p,q) & = & G(q) & \forall q \in W, \end{cases} \tag{3.35}$$

where

$$d(\cdot,\cdot) : W \times W \to \mathbb{R}$$

is continuous and coercive on W.

For this problem it is possible to prove existence, uniqueness, and the following stability bound of the solution

$$\|\mathbf{u}\|_V + \|p\|_Q + \epsilon\|p\|_W \leq K\left(\|F\|_{V^*} + \frac{\|G\|_{Q^*}}{\epsilon}\right),$$

where K depends on the coercivity and continuity constant of the bilinear forms. This stability estimate is not independent on ϵ. In fact, for $\epsilon \to 0$ the norm of p is bounded in Q but it may be unbounded in W.

Proposition 5

A last saddle-point problem considered is

Find $\mathbf{u} \in V$ and $p \in W$ such that:

$$\begin{cases} a(\mathbf{u},\mathbf{v}) & +b(p,\mathbf{v}) & = & F(\mathbf{v}) & \forall \mathbf{v} \in V \\ c(\mathbf{u},q) & -\epsilon d(p,q) & = & G(q) & \forall q \in W, \end{cases} \tag{3.36}$$

where b and c are two weak forms defined on $V \times Q$. Proposition 3 is still valid for this problem, provided that

- b and c respect condition **F5**;
- condition **F4b** is substituted by
  - $\inf_{\mathbf{u}\in\ker B} \sup_{\mathbf{v}\in\ker C} \frac{a(\mathbf{u},\mathbf{v})}{\|\mathbf{u}\|_V \|\mathbf{v}\|_V} \geq \alpha_0$;

- $\inf_{\mathbf{v}\in\ker C}\sup_{\mathbf{u}\in\ker B}\frac{a(\mathbf{u},\mathbf{v})}{\|\mathbf{u}\|_V\|\mathbf{v}\|_V}\geq\alpha_0$.

3.5.5 Galerkin discretization of saddle-point problems

To obtain the Galerkin approximation of (3.31), we introduce two finite dimensional subspaces $V_h\subset V$ and $Q_h\subset Q$ of dimensions $N=\dim(V_h)$ and $M=\dim(Q_h)$. The discrete problem reads

Find $\mathbf{u}_h\in V_h$ and $p_h\in Q_h$ such that:

$$\begin{cases} a(\mathbf{u}_h,\mathbf{v}_h) & +b(\mathbf{v}_h,p_h) & = & F(\mathbf{v}_h) & \forall\mathbf{v}_h\in V_h \\ b(\mathbf{u}_h,q_h) & & = & G(q_h) & \forall q_h\in Q_h. \end{cases} \tag{3.37}$$

As in the continuous case, the weak forms in (3.37) define four operators:

- $A_h:V_h\to V_h^*$;
- $A_h^T:V_h\to V_h^*$;
- $B_h:V_h\to Q_h^*$;
- $B_h^T:Q_h\to V_h^*$.

The space Q_h^* is larger than Q^* since $Q_h\subset Q$. In general, in order to directly use the stability theorems presented above, the dual space Q_h^* is assumed to be a subspace of Q^* [BF91].

We denote by

$$\{\boldsymbol{\varphi}_j\}_{j=1}^N,\quad \{\phi_k\}_{k=1}^M \tag{3.38}$$

a basis of V_h and Q_h, respectively. Hence, the discrete solution can be written as

$$\mathbf{u}_h=\sum_{j=1}^M u_j\boldsymbol{\varphi}_j\,,\qquad p_h=\sum_{l=1}^N p_l\phi_l\,, \tag{3.39}$$

Thus, we can define the algebraic representations of the weak forms $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ as $\mathbf{A}\in\mathbb{R}^{N\times N}$ and $\mathbf{B}\in\mathbb{R}^{N\times M}$, respectively. Their entries are given by

$$A_{ij}=a(\boldsymbol{\varphi}_j,\boldsymbol{\varphi}_i)\qquad B_{kj}=b(\boldsymbol{\varphi}_j,\phi_k). \tag{3.40}$$

Then, we introduce the vectors $\underline{f}$ and $\underline{g}$ whose components are given by

$$f_i=F(\boldsymbol{\varphi}_i)\,,\qquad g_k=G(\phi_k)\,. \tag{3.41}$$

The problem (3.37) is hence equivalent to the algebraic problem

$$\begin{pmatrix} \mathbf{A} & \mathbf{B}^T \\ \mathbf{B} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} = \begin{pmatrix} \underline{f} \\ \underline{g} \end{pmatrix}, \tag{3.42}$$

where the vectors $\underline{u}$ and $\underline{p}$ collect the coefficients of the solution $\mathbf{u}_h$ and $\mathbf{p}_h$ with respect to the bases (3.39).

Stability of (3.37) depends on the kernel V_h^0 of B_h. In fact, a priori this set is not a subset of V^0. Hence, Proposition 1 in Subsec. 3.5.3 applies to this problem with different constants α_h and β_h. The first component of solution $(\mathbf{u}_h, p_h)$ exists and it is unique and bounded by the data of the problem, but the second component p_h is defined up to functions in the kernel of B_h^T. In this case, B_h is not the discrete counterpart of B. The discrete and the continuous operators coincide only if $B_h V_h \subset Q_h^*$.

Elements in the kernel of B_h^T that are not in the kernel of B^T are called *spurious modes*. The presence of spurious modes affects the quality of the discrete solution $(\mathbf{u}_h, p_h)$ and its convergence to the continuous solution $(\mathbf{u}, p)$.

A condition that ensures that $V_h^0 \subset V^0$ is the so-called inf-sup or Ladyženskaja-Babuška-Brezzi (LBB) condition [BF91]. It is a necessary condition and it is a classical result in the study of SP problems. It simply reads:

$$\beta_h \geq \beta. \tag{3.43}$$

Thus, the constant of the discrete condition **F5** has to be larger than the continuous condition **F5**. Condition (3.43) also implies that

$$\ker B_h^T = \ker B^T \cap Q_h \subset \ker B^T.$$

Discrete spaces that respect this condition are said to be inf-sup *stable* or *compatible*. The LBB condition also ensures convergence.

3.5.6 Mixed FE discretization and stabilization for linear incompressible elasticity

Problem (3.27) can be cast in the setting of (3.31) assuming that

$$a(\mathbf{u}, \mathbf{v}) = \int_\Omega \boldsymbol{\epsilon}^D(\mathbf{u}) : \boldsymbol{\epsilon}^D(\mathbf{v}),$$

$$b(\mathbf{u}, q) = \int_\Omega (\mathrm{Div}\, \mathbf{u}, q)_{L^2},$$

$$F(\mathbf{v}) = \int_\Omega \mathbf{f} \cdot \mathbf{v} + \int_{\Gamma_N} \mathbf{t} \cdot \mathbf{v},$$

and

$$G(q) = 0.$$

The function spaces are $V = H^1_{\Gamma_D}(\Omega, \mathbb{R}^3)$ and $Q = L^2(\Omega)$. If $\Gamma_D = \partial\Omega$, the kernel of B^T is given by the constant function on Ω [BF91]. In fact, only the gradient pressure appears in the equations and thus p is defined up to a constant function. In this case, uniqueness of the pressure component of the solution is obtained by setting

$$Q = L^2_0(\Omega) := \left\{ q \in L^2(\Omega) : \int_\Omega q = 0 \right\}.$$

Using the definition (3.11), the FE discretization is obtained choosing as discrete subspaces V_h^r and Q_h^σ. The superscripts r and σ denote the order of approximation of the two spaces. In literature, the couples are also denoted by $\mathbb{P}_r$-$\mathbb{P}_\sigma$ in case of triangular or tetrahedral elements, or $\mathbb{Q}_r$-$\mathbb{Q}_\sigma$ in case of quadrilateral or hexahedral elements. For FE discretizations, we always consider the basis functions (3.38) as Langrangian bases. Hence, the discretization of (3.27) can be written as (5.27). In general FE discretizations are not LBB stable. There are two strategies to overcome this problem:

- choosing r and σ that are inf-sup stable;
- introducing auxiliary stabilization terms.

Possible choices of stable spaces are reported in [Qua09]. In particular, we remind that the choice $\mathbb{P}_k$-$\mathbb{P}_{k-1}$ is stable for $k \geq 2$. These couples of FE are called Taylor-Hood elements. The Lagrangian nodes $\mathbb{P}_2$-$\mathbb{P}_1$ of this couple are shown in Fig. 3.1.

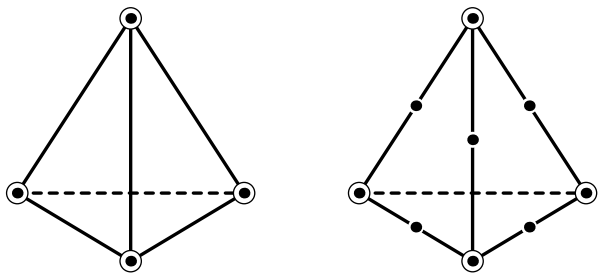


Figure 3.1. DOFs for the employed P1-P1 couple (left) and for Taylor-Hood P2-P1 elements (right). The full circles indicate DOFs for displacement, while the empty circles indicate the DOFs for pressure.

A widely used stabilization method has been presented in [BJ84] and it is named after Brezzi and Pitkäranta. This method is also related to the stability of poroelasticity system that will be discussed in the next chapter. The idea is to change the discrete problem to

Find $\mathbf{u}_h \in V_h^l$ and $p_h \in Q_h^l$, such that:

$$\left\{ \begin{array}{lclcll} a(\mathbf{u}_h, \boldsymbol{\varphi}_i) & + & b(\boldsymbol{\varphi}_i, p_h) & = & F(\boldsymbol{\varphi}_i) & i = 1 \dots M, \\ b(\mathbf{u}_h, \phi_k) & - & d_h(p_h, q_h) & = & G(\phi_k) & k = 1 \dots N, \end{array} \right. \tag{3.44}$$

where the weak form d_h is given by

$$d_h(p_h, q_h) = \alpha \sum_{K \in \mathcal{T}_h} h_K^2 \int_K \nabla p_h \cdot \nabla q_h,$$

with $\alpha > 0$ to be chosen. This stabilization renders the problem not consistent because of the presence of d_h. The stabilization vanishes for $h \to 0$ and the method is convergent. If the problem is regular enough, the accuracy of the method is the order of the FE spaces for both variables.

Now, the analysis for a fixed h has to be done considering it as problem of the form (3.36). This new formulation allows to choose spaces of the same order. In order to see that, we can study the algebraic problem

$$\begin{pmatrix} \mathbf{A} & \mathbf{B}^T \\ \mathbf{B} & -\mathbf{D} \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} = \begin{pmatrix} \underline{f} \\ \underline{g} \end{pmatrix}. \tag{3.45}$$

If we eliminate the displacement vector $\underline{u}$, we obtain a system

$$\mathbf{S}\underline{p} = \mathbf{B}\mathbf{A}^{-1}\underline{f} - \underline{g},$$

with $\mathbf{S} = \mathbf{D} + \mathbf{B}\mathbf{A}^{-1}\mathbf{B}^T$. In case of an unstable FE pair, the matrix $\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^T$ is positive semi-definite but singular and its kernel is given by the spurious modes. The matrix $\mathbf{D}$ is positive semi-definite and its kernel is composed only of constant vectors. This ensures that $\mathbf{S}$ is invertible.

3.5.7 Mixed FE for poroelasticity problem

The FE discretization of the linear poroelasticity system has been presented in [EM01]. We first introduce the two bilinear forms of the deviatoric and the isochoric components. They read:

$$a_1(\mathbf{u}, \mathbf{v}) = 2 \int_\Omega \text{Dev}\,(\boldsymbol{\epsilon}(\mathbf{u})) : \text{Dev}\,(\boldsymbol{\epsilon}(\mathbf{v}))\, \mathrm{d}\mathbf{X},$$

and

$$a_2(\mathbf{u}, \mathbf{v}) = \int_\Omega \text{tr}\,(\boldsymbol{\epsilon}(\mathbf{u}))\text{tr}\,(\boldsymbol{\epsilon}(\mathbf{v}))\, \mathrm{d}\mathbf{X}.$$

Hence, the semi-discrete problem (3.30) can be cast in the general form (3.36) by setting:

$$a(\mathbf{u},\mathbf{v}) = \mu a_1(\mathbf{u},\mathbf{v}) + G a_2(\mathbf{u},\mathbf{v}),$$

$$b(\mathbf{u},q) = \int_\Omega \mathrm{Div}\,(\mathbf{u}) q \,\mathrm{d}\mathbf{X},$$

$$d(\pi,q) = \int_\Omega \nabla\pi \cdot \nabla q \,\mathrm{d}\mathbf{X}.$$

The two spaces are defined as $V := H^1_{\Gamma^{\mathbf{u}}_D}(\mathscr{B}_r,\mathbb{R}^3)$ and $W := H^1_{\Gamma^{\pi}_D}(\mathscr{B}_r)$. The FE discretization can be obtained by choosing two subspaces V^r_h and W^σ_h. Hence, expanding with respect to the Lagrangian bases of these two spaces, we can derive the algebraic formulation

$$\begin{pmatrix} \mathbf{A} & -\mathbf{B}^T \\ -\mathbf{B} & -\mathbf{D} \end{pmatrix} \begin{pmatrix} \underline{u^m} \\ \underline{\pi^m} \end{pmatrix} = \begin{pmatrix} \underline{f} \\ \underline{h} \end{pmatrix}. \tag{3.46}$$

Analysis of the discretized problem can be performed as in Proposition 4. We observe that this problem is stable with the choices $\mathbb{P}_k$-$\mathbb{P}_{k-1}$ and $\mathbb{P}_k$-$\mathbb{P}_k$. Using Taylor-Hood FE spaces, we can directly use Proposition 4 and prove the existence and uniqueness of the solution. Instead, the choice of FE spaces of the same order is similar to the one presented for the Brezzi-Piktäranta stabilization (3.44). The difference is that here the weak form d does not depend on the grid size.

Chapter 4

Stability and stabilization of the discrete poroelastic system

In this chapter, we derive a stability limit that relates the spatial and temporal discretization step-sizes with the mechanical parameters of the poroelastic model. In particular, we extend previously known results from the one-dimensional case to the two- and three-dimensional space. We define a novel Péclet number for the discrete poroelasticity system in the multi-dimensional cases. This Péclet number is obtained by means of numerical experiments, because the simplifications that can be done in the simple one-dimensional study are not possible. A condition on this number allows to ensure the stability of the system. Finally, we discuss an important stabilization method, namely the Fluid Pressure Laplacian (FPL) stabilization. This stabilization depends on a numerical parameter that has to be chosen. Again, the Péclet number provides an optimal value for the numerical parameter of the FPL stabilization. The concepts presented here have been published in [FGK].

4.1 Introduction

In the field of poroelasticity, it is well-known that numerical oscillations may occur if the spatial (h) and the temporal (τ) discretization parameters are not properly chosen depending on the physical parameters of the problems [ML92, VV81, AGLR08, MPK98, PP11, GLV03, Wan03]. The physical parameters are the solid material moduli, i.e (μ, λ), (μ, G), or (E, ν), and the permeability k.

These oscillations are also called *wiggles* [GL81] or, more generally, numerical instabilities. First of all, we want to recall that, even if the term "instabilities" is widely used, it is not totally correct. Indeed, as also pointed out in [Qua09] concerning reaction- and convection-dominated elliptic problems, the discrete problem (3.44) is stable in the sense that it has a unique solution that depends continuously on the data. Here the term "unstable" refers to a solution that is not free of unphysical oscillations.

In the poroelasticity problem, wiggles have been observed in particular for materials with small permeability [HOL12] and when studying the convergence in time of some temporal integration methods [FGT01]. These numerical instabilities are noticeable both when FEM and finite difference methods are used for the spatial discretization. They disappear when the spatial mesh is refined. In this chapter, we focus on the stability of FE discretizations.

For this method, several conditions that depend on the discretization and physical parameters have been derived. These bounds are usually referred to as *stability* condition, *accuracy* condition, and *monotonicity* condition. In general they take the following form

$$\frac{h^2}{C\tau k\,\mathfrak{B}} < 1, \tag{4.1}$$

where $\mathfrak{B}$ is a function of the mechanical parameters of the solid phase. Usually it is written as $\mathfrak{B} = \mathfrak{B}(\mu, \lambda)$. In this work, we consider $\mathfrak{B}$ as a function of μ and G after performing the splitting (3.21) between the deviatoric and the isochoric parts. The use of shear and bulk moduli allows to obtain simpler and more general forms of $\mathfrak{B}$.

The numerical constant C depends on the dimension of the problem and on the FE spaces that will be used to approximate the continuous spaces V and Q. If necessary, we will write it as $C_\beta^{FE_d}$. The superscript FE_d denotes the couple of FE spaces and the dimension of the problem and the subscript β is an integer that denotes the maximum order of derivative that is free from numerical instabilities.

The condition (4.1) induces a lower bound for τ and an upper bound for h, i.e.

$$\tau > \frac{h^2}{Ck\mathfrak{B}(\mu, G)} \quad \text{and} \quad h < \sqrt{C\tau k\mathfrak{B}(\mu, G)}. \tag{4.2}$$

The importance of this limit is twofold.

- For the physical meaning of the solution; for example, in consolidation problems an external load is applied on the porous material and then a monotone positive increase in pressure is observed. If the mesh resolution is not fine enough wiggles are responsible for a change of sign in the solution.

- For the convergence of solvers based on hierarchy of grids, e.g. multigrid (MG) and domain decomposition (DD) methods: for what concerns multilevel solution methods, condition (4.1) has to be respected on all the grids involved in the solution process. If (4.1) is not met on all the levels of the MG algorithm, it may fail in convergence or it may not converge with an optimal rate.

The stability limit (4.1) is well-understood in the one-dimensional case. It has been related to a lack of the discrete maximum principle of an equivalent diffusion-reaction problem. The maximum principle is a property of some elliptic problems that ensures that the solution is a positive function if the external sources are positive functions. The Galerkin discretization of a problem may not satisfy the maximum principle even if it

holds for its continuous counterpart. For reaction-diffusion problems, this happens when the reaction coefficient dominates the diffusion coefficient by orders of magnitude. In this case a fine resolution of the mesh is required in order to ensure that the maximum principle holds also for the discrete problem.

In [PP11], it has been shown that the continuous poroelasticity system can be simplified to a heat equation with diffusion coefficient kG. The semi-discretization in time of this problem results in a reaction-diffusion problem with coefficient τkG. Hence, if the temporal time step is too small, the reaction dominates the diffusion and wiggles may show up.

In [GLV03], the same result is obtained for the discrete poroelasticity problem. There, the Schur complement of the system is shown to be equivalent to a discretized reaction-diffusion problem with diffusion coefficient proportional to τ. Even in this case, if τ is too small numerical oscillations occur.

In the two-dimensional case, some other results have been proposed but no proof was provided.

In [PP11], wiggles have been related to a choice of discrete spaces that do not fulfill the LBB condition. As discussed in Subsec. 3.5.4, the discrete problem is LBB stable independently on the parameter τ. In [GLV03], it was observed that compatible FE couples give larger C and hence allow for smaller time steps.

In order to prevent wiggles in poroelasticity problems, two possible strategies are available: choosing a spatial step-size that is fine enough or introducing stabilization methods. A detailed résumé of the latter can be found in [PP11] and in [Wan03]. In this work, we focus on the FPL method. It consists in the introduction of an artificial permeability that is proportional to h. It is also called numerical permeability and it is summed to the physical permeability. Numerical permeability tends to k for $h \to 0$.

4.1.1 Scaled system

n this chapter we focus on the stability of the poroelastic problem semi-discretized in time (3.30). This problem is a particular form of the equation (3.35):

Find $\mathbf{u} \in V$ and $\pi \in W$ such that:

$$\begin{cases} \mu a_1(\mathbf{u},\mathbf{v}) + G a_2(\mathbf{u},\mathbf{v}) & +b(\pi,\mathbf{v}) & = & F(\mathbf{v}) & \forall \mathbf{v} \in V \\ b(\mathbf{u},q) & -\tau k d(\pi,q) & = & G(q) & \forall q \in W. \end{cases} \tag{4.3}$$

The bilinear forms and the functionals are defined in Subsec. 3.5.7. The bilinear forms a_1 and a_2 are related to the deviatoric and isochoric components of a linear elasticity tensor. The weak form b is related to the divergence and d to the Laplacian. After FE discretization, the variational problem (4.3) can be written as the following linear system:

$$\begin{pmatrix} \mu\mathbf{A}_1 + G\mathbf{A}_2 & -\mathbf{B}^T \\ -\mathbf{B} & -\tau k\mathbf{D} \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{\pi} \end{pmatrix} = \begin{pmatrix} \underline{f} \\ \underline{h} \end{pmatrix}. \tag{4.4}$$

As also used in the previous chapter, a bold capital letter represents the algebraic counterpart of the weak form with the same letter in lower case.

In (4.3), there are three physical parameters μ, G, and k. It is possible to set two of them to one by means of the following change of variables

$$u := u/\mu, \quad \pi := \mu k \pi, \quad G := G/\mu \tag{4.5}$$

or

$$u := u/G, \quad \pi := Gk\pi, \quad \mu := \mu/G. \tag{4.6}$$

4.1.2 Assembly

The FE method transforms a weak problem into an algebraic linear system. The construction of the stiffness matrix is called *assembling* and it is a crucial task in FE softwares. This is the phase where geometrical and functional information are merged together.

After having created a grid, functional data structures are created depending on the type of the considered FE space. In our simulation we focus on linear ($\mathbb{P}_1$), bilinear ($\mathbb{Q}_1$), quadratic ($\mathbb{P}_2$), and biquadratic ($\mathbb{Q}_2$) FE spaces. For linear and bilinear FEs, nodes are defined on the vertices of a grid. For quadratic and biquadratic FEs, nodes are defined on the vertices and on the mid-point of the edges of a grid. Each node corresponds to a Lagrangian basis function and hence to a row of the stiffness matrix.

The assembling procedure is done looping over the elements and assembling a local stiffness matrix on the considered elements. Then the local matrices are summed up to the global stiffness equation. The assembling of the local stiffness matrix is usually done by transforming the integrals on the considered element to a reference element. On the reference element, basis functions and their derivatives do not depend on the geometry of the element.

In the rest of the chapter, we use the following standard:

- lower case non-bold letters refer to weak forms, e.g. d,
- capital bold letters refer to the global stiffness matrices, e.g. $\mathbf{D}$,
- lower case bold letters with subscript K refer to the local stiffness matrices on the element K, e.g. $\mathbf{d}_K$,
- lower case bold letters without subscript refer to the local stiffness matrices on the reference element, e.g. $\mathbf{d}$.

4.1.3 M-matrices

We report some linear algebra results that are necessary for the study of the oscillations.

Definition 1 *A matrix* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *is said to be an M-matrix if*

- $\mathbf{A}_{ii} > 0 \quad 1 \le i \le n,$
- $\mathbf{A}_{ij} \le 0 \quad 1 \le i, j \le n, i \ne j,$
- $\mathbf{A}^{-1}_{ij} > 0.$

An M-matrix $\mathbf{A}$ satisfies the discrete maximum principle (DMP), i.e., if $\mathbf{A}\underline{u} \le \underline{0}$ then $\underline{u} \le \underline{0}$.

Sufficient conditions that ensure a matrix $\mathbf{A}$ is an M-matrix are given by the following propositions.

Proposition 1 *A non-singular matrix* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *is an* M*-matrix if the following conditions are satisfied*

1) $\mathbf{A}_{ii} > 0$

2) $\mathbf{A}_{ij} \le 0 \quad i \ne j$

3) $\mathbf{A}_{ii} \ge \sum_{j=1}^{i-1} |\mathbf{A}_{ij}| + \sum_{j=i+1}^{n} |\mathbf{A}_{ij}| \quad \forall i$

The condition 3 means that the matrix is diagonally dominant.

Proposition 2 *A matrix* $\mathbf{B}$ *deriving from FE assembling is an* M*-matrix, if all the local matrices* $\mathbf{b}_K$ *with* $K \in \mathcal{T}_h$ *are* M*-matrices.*

4.2 Maximum principle

The Laplace equation

$$-\Delta u = 0$$

satisfies the so-called *maximum principle*, i.e. its solution cannot have a maximum or a minimum in an interior point of its domain unless the solution is constant [Sal08]. This theorem directly derives from the mean value property of the solution of the Laplace equation.

A similar result is also true for the non-homogeneous case, i.e. the Poisson equation

$$\begin{cases} -\Delta u &= f \quad \text{in } \Omega \\ u &= 0 \quad \text{on } \partial\Omega. \end{cases} \tag{4.7}$$

For (4.7) it is possible to prove that, if $f \geq 0$ a.e. in Ω, then for every $\mathbf{X} \in \Omega$

$$u(\mathbf{X}) \geq \min_{\partial\Omega} u.$$

In case $f \leq 0$, a similar result holds for the maximum of u on the boundary.

The maximum principle may also be extended to more general elliptic problems, e.g. diffusion-reaction, diffusion-convection, or bi-Laplacian problems. In this section, we focus on diffusion-reaction problems of the following form

$$\begin{cases} -\nabla \cdot k\nabla u + \sigma u &= f \quad \text{in } \Omega \\ u &= 0 \quad \text{on } \Gamma_D \\ -\mu\nabla u &= h \quad \text{on } \Gamma_N. \end{cases} \tag{4.8}$$

We also assume the following hypotheses on the coefficients:

G1: $k > 0$;

G2: $\sigma \in L^\infty(\Omega)$ and $\sigma > 0$;

G3: $f \in L^2(\Omega)$ and $h \in L^2(\Gamma_N)$.

Hence, introducing the weak form

$$B(u,v) := kd(u,v) + \int_\Omega \sigma\, u\, v, \tag{4.9}$$

the linear functional

$$F(v) = \int_\Omega f\, v + \int_{\Gamma_N} h\, v, \tag{4.10}$$

and the function space $V := H^1_{\Gamma_D}(\Omega)$, the variational formulation of (4.8) reads

$$\text{Find } u \in V \text{ s.t. } B(u,v) = F(v) \quad \forall v \in V. \tag{4.11}$$

The variational problem (4.11) admits a unique solution that depends continuously on the data of the problem. In fact, it is possible to apply the Lax-Milgram theorem (see Sec. 3.1.4) with the following continuity and coercivity constants

$$M_B = k + \|\sigma\|_{L^\infty}, \tag{4.12}$$

and

$$\alpha = k/(1 + C_P^2). \tag{4.13}$$

Moreover, problem (4.11) satisfies also the maximum principle property. Actually, it is possible to prove the following theorem [Sal08].

Theorem 1 *Let us assume hypotheses* **G**. *If $f \geq 0$ a.e. in Ω and $h \geq 0$ then*

$$\inf_{\Omega} u \geq \inf_{\partial\Omega} u. \tag{4.14}$$

Introducing a suitable FE space $V_h \subset V$, the discrete formulation of (4.11) reads

$$\text{Find } u_h \in V_h \text{ s.t. } B(u_h, v_h) = F(v_h) \quad \forall v_h \in V_h. \tag{4.15}$$

The FE formulation satisfies the same stability property as its continuous counterpart but u_h does not inherit the maximum principle property. This fact is strictly related to the convergence constant of the FEM. As discussed in Subsec. 3.2.1, the Ceà lemma gives the following convergence estimate

$$\|u - u_h\|_V \leq \frac{M_B}{\alpha} \inf_{v_h \in V_h} \|u - v_h\|_V,$$

where M_B and α are defined in (6.8) and (4.13). This estimate on the error may become meaningless if the ratio between M_B and α is large. This happens in particular when $\|\sigma\|_{L^\infty} > k$. In this case a very high resolution of the grid, and hence a small h, may be needed in order to reach an acceptable discretization error. As we will show in the next subsection, this is strictly related to the maximum principle for the discrete solution.

Respecting the discrete maximum principle is particularly important for differential problems involving positive physical quantities. For example, for the simulation of absolute temperatures or concentrations, having non-negative solutions is important for the physical interpretation. Moreover, when the discrete problem satisfies the maximum principle property, it is possible to prove that the discrete solution converges also uniformly to the solution of the continuous problem [CR73]. In the next subsection, we derive for a one and two-dimensional reaction-diffusion problem a lower bound for h that ensures a discrete maximum principle.

4.2.1 The discrete maximum principle for reaction-diffusion problems

The lack of the DMP introduces wiggles in the approximated solution u_h of (4.15). In this section, we consider this problem on $\Omega = (0,1)^d$ with $d = 1, 2$ and discretizations with $\mathbb{Q}_1$ and $\mathbb{Q}_2$ FEs. We assume that σ is a given positive constant. Hence, the weak form in (4.15) reads

$$B(\mathbf{u}_h, \mathbf{v}_h) = kd(\mathbf{u}_h, \mathbf{v}_h) + \sigma m(\mathbf{u}_h, \mathbf{v}_h),$$

with $m(\mathbf{u}_h, \mathbf{v}_h) = (\mathbf{u}_h, \mathbf{v}_h)_{L^2}$. The bilinear form d is the diffusion term and the bilinear form m is the reaction term. The application of the Galerkin FEM to (4.15) gives rise to the following linear system

$$\mathbf{B}\underline{u} = \underline{f}. \tag{4.16}$$

The matrix $\mathbf{B}$ can be written as the sum of the discrete counterparts of the diffusion and reaction terms, i.e.

$$\mathbf{B} = k\mathbf{D} + \sigma\mathbf{M}.$$

In [Qua09], the monodimensional case discretized with linear FEs was studied. The author showed the existence of a bound on the spatial mesh-size h in order to avoid numerical instabilities. This condition is based on the Péclet number that is defined as

$$\mathbb{P}e = \frac{\sigma h^2}{6\mu}. \tag{4.17}$$

The condition $\mathbb{P}e < 1$ ensures that no oscillations occur. This condition has been derived studying the rows of the linear system as a difference equation.

If (4.17) holds, oscillations do not show up in the solution and in its derivatives. For higher order FEs, a different Péclet number can be defined. Moreover in this case, even if the corresponding Péclet number is smaller than one, unphysical wiggles may be still observed in the first or second derivative. For this reason, we define a more general Péclet number

$$\mathbb{P}e_{\beta}^{FE_d} = \frac{\sigma h^2}{C\mu}. \tag{4.18}$$

The superscript FE_d denotes the FE space and the dimension. The subscript β is an integer that denotes up to which order of derivative the solution is free of wiggles. The only difference between $\mathbb{P}e$ and $\mathbb{P}e^{FE_d}$ is in the numerical constant C. For this reason we denote it as $C_{\beta}^{\mathrm{FE}_d}$.

$\mathbb{P}_1$ FE in 1D

We now consider linear FE and a uniform mesh of step-size h. In $1D$, the diffusion and mass matrices on the reference element are respectively

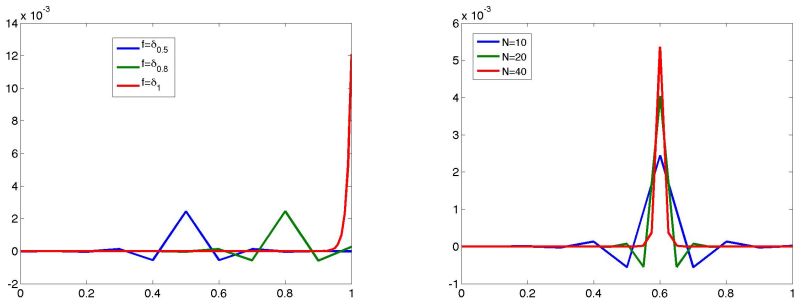
$$\mathbf{d} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \text{ and } \mathbf{m} = \frac{1}{6}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}. \tag{4.19}$$

Thus, the local matrix associated to the reaction-diffusion operator is given by

$$\mathbf{d}_K = \frac{k}{h}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + \frac{\sigma h}{6}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}. \tag{4.20}$$

The global stiffness matrices $\mathbf{D}$ and $\mathbf{M}$ have the following tridiagonal forms

$$\mathbf{D} = \frac{k}{h}\begin{pmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{pmatrix} \tag{4.21}$$



(a) (b)

Figure 4.1. Plot of Green's functions for three different mesh nodes (a). Effect of mesh refinement on the Green's function at node in 0.6.

and

$$\mathbf{M} = \frac{\sigma h}{6}\begin{pmatrix} 2 & 1 & & & \\ 1 & 4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 4 & 1 \\ & & & 1 & 2 \end{pmatrix}. \tag{4.22}$$

The solution of the discrete problem is studied in terms of difference equations. The i-th row of the linear system (4.16) reads

$$\left(\frac{\sigma h}{6} - \frac{k}{h}\right) u_{i-1} + \left(\frac{2\sigma h}{3} + \frac{2k}{h}\right) u_i + \left(\frac{\sigma h}{6} - \frac{k}{h}\right) u_{i+1} = f_i.$$

Dividing the previous equation by $h/6$, we can write it as a function of $\mathbb{P}e = \mathbb{P}e^{\mathbb{P}_1}$ as follows

$$(\mathbb{P}e - 1)u_{i-1} + (2\mathbb{P}e + 1)u_i + (\mathbb{P}e - 1)u_{i+1} = \frac{6}{h} f_i.$$

The solution of this difference equation is positive when

$$\mathbb{P}e < 1. \tag{4.23}$$

The same result is obtained applying Proposition 1 to the local stiffness matrix $\mathbf{b}_K$.

As a numerical example, we consider a case with $k = 1$, $\sigma = 6500$, and $f = \delta_X$. The solution of the problem with a Dirac delta centered at X as right hand side is the Green's function at X. Green's functions are

H1 positive,

H2 increasing over $(0, X)$ and decreasing over $(X, 1)$,

H3 convex on $(0, 1)\backslash\{X\}$.

Their second derivative is positive when defined.

Using a mesh with $N = 10$, the Péclet number is larger than 1. In Fig. 4.1a, we show the Green's functions for three different nodes with $N = 10$. In this case, the Péclet number is $\mathbb{P}e = 10.8\overline{3}$ and we observe that the solution has negative values. In Fig. 4.1b, we see the effect of mesh refinement on the positivity of the solution. We used three different meshes $N = 10, 20, 40$ corresponding to $\mathbb{P}e = 10.833, 2.708, 0.677$. We observe that as soon as the condition (4.23) is respected the solution becomes positive and no oscillations are observed in the first and second derivative.

$\mathbb{P}_2$ FE in 1D

In the case of linear FEs, we have shown that the lack of the DMP is connected to the M-matrix property of the stiffness matrix. When using P_2 FEs, this property is not related to the positivity of the solution. In fact, the diffusion and mass matrices on the reference element are

$$\mathbf{d} = \frac{1}{8}\begin{vmatrix} 7 & -8 & 1 \\ -8 & 16 & -8 \\ 1 & -8 & 7 \end{vmatrix}, \text{ and } \mathbf{m} = \frac{1}{30}\begin{vmatrix} 4 & 2 & -1 \\ 2 & 16 & 2 \\ -1 & 2 & 4 \end{vmatrix}. \tag{4.24}$$

Proposition 1 does not apply to this problem for two reasons:

- $\mathbf{d}_{13}$ and $\mathbf{d}_{31}$ are positive;
- $\mathbf{d}$ is not diagonally dominant.

Hence, even for large k the matrix is not an M-matrix. This does not prevent the global matrix from having positive entries in the inverse, in fact Proposition 1 provides only a sufficient condition.

A necessary bound on h can be derived numerically. We study the effect of the mesh refinement on the solution (Fig. 4.2), on the first derivative (Fig. 4.3), and on the second derivative (Fig. 4.4) for a reaction-diffusion problem with $\sigma = 10000$. In the one-dimensional case, the violation of the DMP directly resulted in negative values in the solution and oscillations in the derivatives. In these numerical tests we observed three different ranges depending on the mesh refinement N. First, there are negative values in the second derivative ($N = 21$), then oscillations in the first derivative ($N = 18$), and finally the violation of the maximum principle in the solution ($N = 15$).

Our idea is to perform numerical tests on (4.15) **with** $k = 1$ **in order to identify for which** σ **the discrete Green's function respects: H1; H1 and H2; or H1 and H2 and H3 . This allows to identify the value of** $\mathbb{P}e_{\beta}^{\mathbf{FE}_d}$**.**

From these numerical experiments, we identified three different $\mathbb{P}e_{\beta}^{\mathbb{P}2_1}$, for $\beta = 0, 1, 2$ that are

$$\mathbb{P}e_0^{\mathbb{P}_{21}} = \frac{\sigma h^2}{40k}, \quad \mathbb{P}e_1^{\mathbb{P}_{21}} = \frac{\sigma h^2}{26.1k}, \quad \mathbb{P}e_2^{\mathbb{P}_{21}} = \frac{\sigma h^2}{20k}.$$

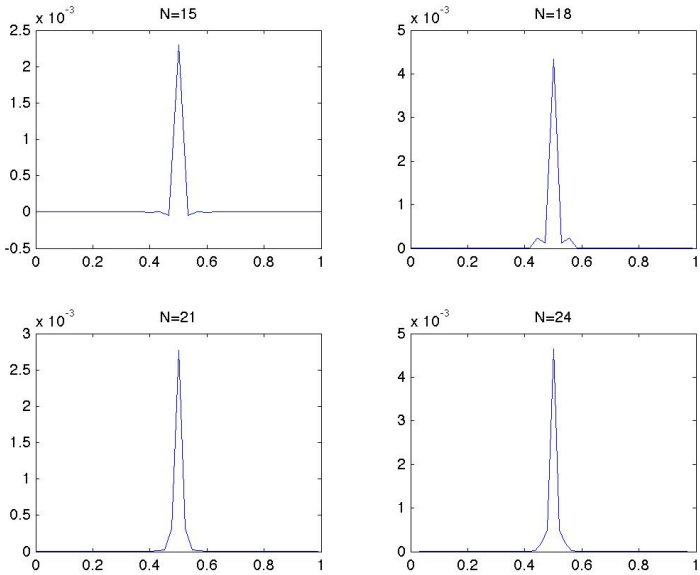


Figure 4.2. The Green's function at node $X = 0.5$ with $\sigma = 10000$ for several mesh resolutions.

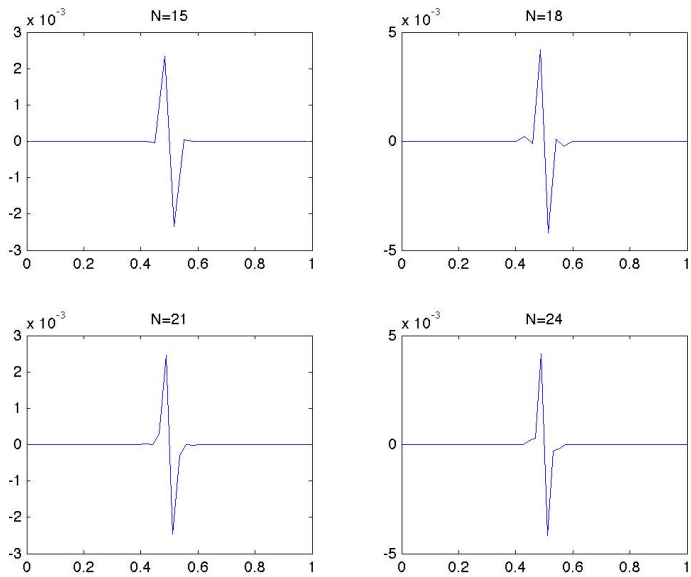


Figure 4.3. The first derivative of the Green's function at node $X = 0.5$ with $\sigma = 10000$ for several mesh resolutions.

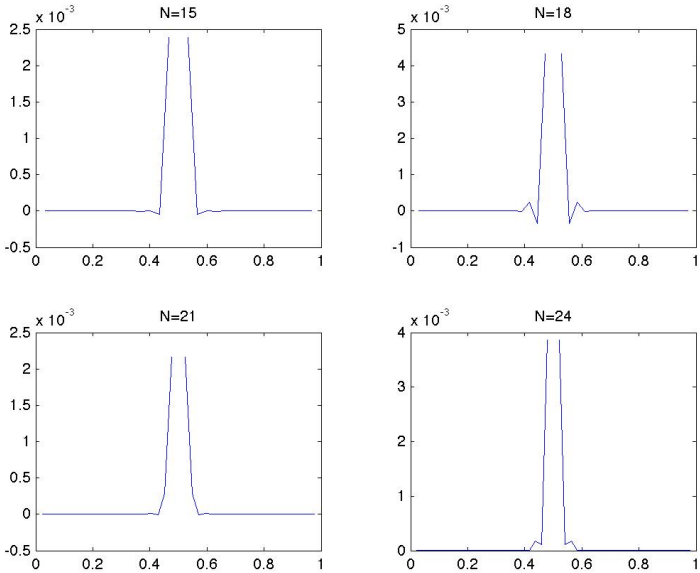


Figure 4.4. The second derivative of the Green's function at node $X = 0.5$ with $\sigma = 10000$ for several mesh resolutions.

If $\mathbb{P}e_{\beta}^{FE_d}$ is smaller then 1, the corresponding β derivative of the solution does not exhibit an unphysical behaviour. The same phenomenon has been observed in other numerical examples where the solution is C^{∞}, e.g. the numerical example in [Qua09]. Finally, we point out that the reaction coefficient $\sigma = 40h^2$ is the one for which we have the transition from positive to negative values of the entries in position $\mathbf{D}_{i,i+1}$.

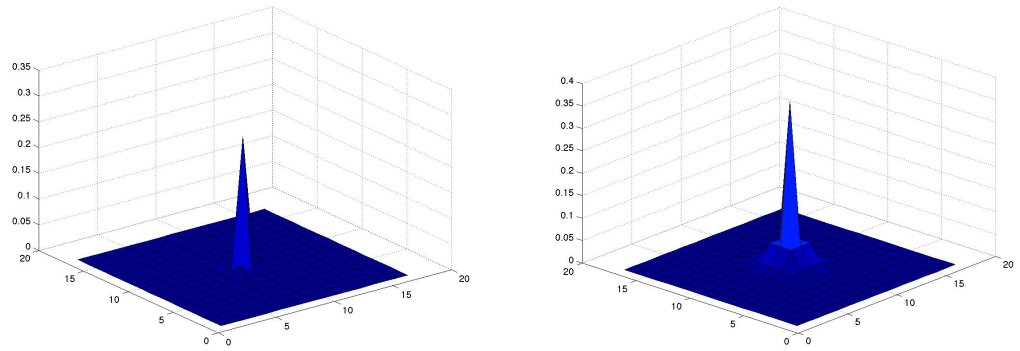
$\mathbb{Q}_1$ FE in 2D

Even in 2D, diffusion-reaction problems present instabilities. For $\mathbb{Q}_1$ FEs, the two local diffusion and mass matrices are

$$\mathbf{d} = \frac{1}{6}\begin{vmatrix} 4 & -1 & -1 & -2 \\ -1 & 4 & -2 & -1 \\ -1 & -2 & 4 & -1 \\ -2 & -1 & -1 & 4 \end{vmatrix}, \tag{4.25}$$

and

$$\mathbf{m} = \frac{1}{36}\begin{vmatrix} 4 & 2 & 2 & 1 \\ 2 & 4 & 1 & 2 \\ 2 & 1 & 4 & 2 \\ 1 & 2 & 2 & 4 \end{vmatrix}. \tag{4.26}$$



(a) *(b)*

Figure 4.5. Lack of maximum principle (left) and oscillations in the gradient (right) in the Green's function for a bidimensional reaction-diffusion problem.

In this case, the M-matrix considerations could be applied because the extra-diagonal entries are negative. In fact, as long as

$$\mathbb{P}e_0^{\mathbb{Q}_{12}} = \frac{\sigma h^2}{3k}$$

is smaller than one, no negative values are observed in the solution. This coefficient coincides with the transition from positive to negative of the entries in the positions $\mathbf{d}_{1,2}$, $\mathbf{d}_{1,3}$, $\mathbf{d}_{2,1}$, $\mathbf{d}_{2,4}$, $\mathbf{d}_{3,1}$, $\mathbf{d}_{3,4}$, and $\mathbf{d}_{4,2}$, $\mathbf{d}_{4,3}$ of the "local" reaction-diffusion operator. In Fig. 4.5a, we observe that even if the solution is positive there are still unphysical oscillations. In fact, the Green's functions should radially decrease from the position where the point source is placed but we observe that this is not the case. From numerical experiments, we found that the Péclet number that avoids this behaviour is

$$\mathbb{P}e_1^{\mathbb{Q}_{12}} = \frac{\sigma h^2}{0.8168k}.$$

4.2.2 Stability of semi-discrete formulations for parabolic problems

The application of a time integration scheme for the solution of the heat equation

$$m(u_t, v) = -d(\nabla u, \nabla v)$$

may also present unphysical oscillations [Har04]. This phenomenon can be observed if the time step of the discrete scheme is reduced without accordingly changing the mesh step-size. It is not related to the A-stability of the method but also it shows up for unconditionally stable schemes. This degradation is again related to the semidiscrete

formulation of the heat equation and to a violation of the DMP. The time discretization of heat equation gives

$$\tau d(\nabla u^m, \nabla b) + m(u^m, v) = m(u^{m-1}, v).$$

Thus, we notice that each single time step consists in the solution of a diffusion equation with diffusion constant τ (or reaction constant $1/\tau$). A critical time step for the onset of spatial oscillations may hence be derived studying the Péclet number of the semidiscrete problem. In the monodimensional heat equation, the constraint is

$$\frac{h^2}{6\tau} < 1.$$

This condition may also be generalized to different FE spaces and dimensions following the approach presented in the previous subsection.

4.2.3 Stabilization of reaction-diffusion problems

We discuss two stabilizations that may be applied to singularly perturbed reaction-diffusion problems in order to ensure the DMP.

A first stabilization technique is the so called *mass lumping* [HRZ76]. This strategy consists in approximating the mass matrix $\mathbf{m}$ with a diagonal matrix $\mathbf{m}^L$ where $m^L_{ii} = \sum_j \mathbf{m}_{ij}$. In the $\mathbb{P}_1$ case, this renders the stiffness matrix an M-matrix and hence ensures the maximum principles. This strategy also works for high order FE in 1D. It has also been shown that it coincides to apply a reduced order integration scheme for the inexact integration of the form m [Qua09].

A second stabilization procedure consists in modifying the diffusion constant k in order to have a Péclet number small than 1. The idea is to introduce a numerical permeability $k_h = k(1+D)$, where the constant $D = D(h)$ vanishes for $h \to 0$. This kind of stabilization renders the system inconsistent. We show this procedure for a generic case. We write the numerical bilinear form $B(\mathbf{u}, \mathbf{v}) = k_h d(\mathbf{u}, \mathbf{v}) + \sigma m(\mathbf{u}, \mathbf{v})$. The Péclet number of this scheme is

$$\mathbb{P}e_h = \frac{\sigma h^2}{Ck(1+D)}.$$

The Péclet number of the stabilized problem is smaller than one if $D = \mathbb{P}e$. In fact, in this case $\mathbb{P}e_h < 1$. The idea of this stabilization is to use the numerical diffusion k_h if $\mathbb{P}e > 1$, otherwise we use the physical diffusion constant k.

4.3 Stability of one-dimensional poroelastic system

In this section we focus on the linear one-dimensional poroelastic model. In particular we consider the following example:

$$\begin{cases} -(\sigma - \pi)_X = 0 & \text{in } (0,1)\times\mathscr{I} \\ -\dot{u}_X + (k\pi_X)_X = 0 & \text{in } (0,1)\times\mathscr{I} \\ u(X,0) = 0 & \text{in } (0,1) \\ u(0,t) = 0 & \text{in } \mathscr{I} \\ u_X(1,t) = t & \text{in } \mathscr{I} \\ \pi_X(0,t) = 0 & \text{in } \mathscr{I} \\ \pi(1,t) = 0 & \text{in } \mathscr{I} \end{cases} \tag{4.27}$$

In 1D, the displacement is a scalar function and hence we did not use bold letters. Since the deviatoric component is null, the stress takes the simple form

$$\sigma = Gu_X. \tag{4.28}$$

The system (4.27) is the Terzaghi problem [Bio40]. It can describe a column of soil supporting an external load t. For this particular case, an analytical solution is available.

It shows that in the simulation of biphasic materials unphysical oscillations may appear. As in the reaction-diffusion problem, the oscillations vanish for $h \to 0$. In Figs. 4.6 and 4.7, the solution of the Terzaghi problem is reported for $G = k = 0.1$ and $\tau = 0.1$.

In Fig. 4.6, the spatial step-size is $h = 0.1$. With this setting of parameters, the nodes of the displacements do not move downward as expected but they show an oscillatory behaviour, see Fig. 4.6a. In Fig. 4.6b, we also notice that the pressure is not monotone.

In Fig. 4.7, the solution with $h = 0.01$ is reported. The refinement removed the oscillations both in the displacement and in the pressure.

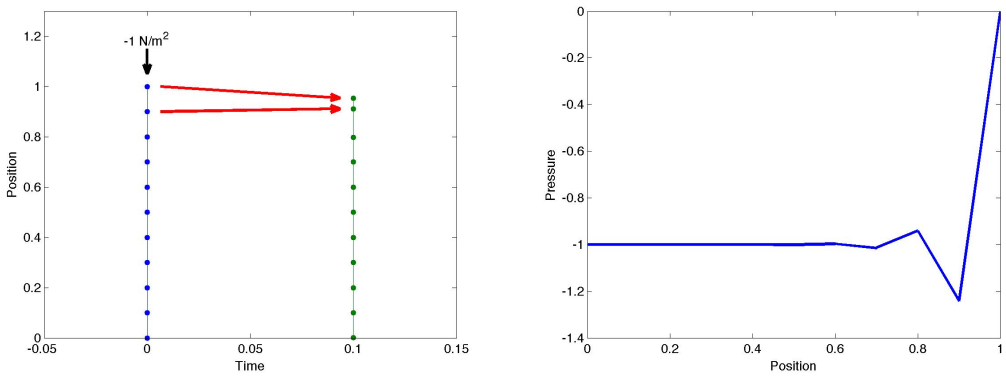
In this one-dimensional case, this phenomenon has been explained both at the continuous and at the discrete level. We will report its motivations in the following sections.

4.3.1 Analytical motivation

In the one-dimensional case, the poroelastic system can be simplified to a heat equation for the pressure. From the continuity equation (4.27.1), we know that

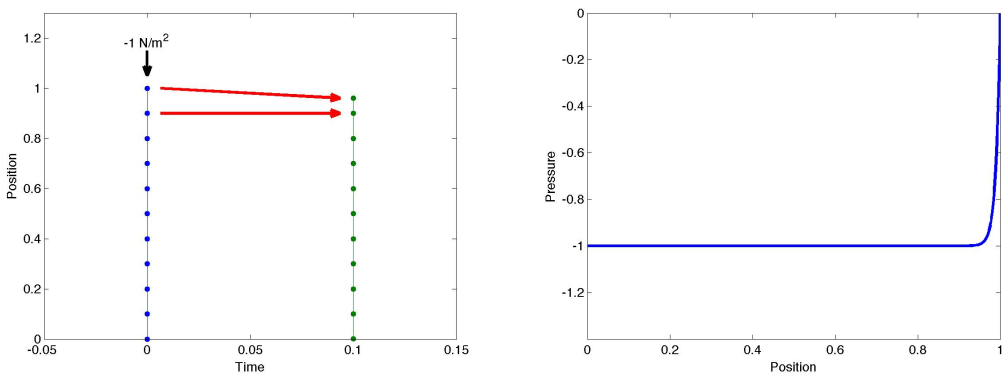
$$\sigma - \pi = \text{const.}$$

Hence, substituting the definition of stress (4.28) in this expression and taking the time derivative, we find



(a) *(b)*

Figure 4.6. Displacement (left) and pressure (right) of the Terzaghi problem in case of instability.



(a) *(b)*

Figure 4.7. Displacement (left) and pressure (right) of the Terzaghi problem with correct mesh refinement.

$$\dot{u}_X = \frac{1}{G}\dot{\pi}.$$

Thus, we can substitute the expression above in (4.27.2) and obtain

$$\frac{1}{G}\dot{\pi} = (k\pi_X)_X.$$

The continuity equation in 1D is thus equivalent to a diffusion problem. Hence, this kind of problems may be affected by the instability presented in Subsec. 4.2.2 and the discretization parameters h and τ cannot be arbitrarily chosen.

4.3.2 Algebraic motivation

In the one-dimensional case, the deviatoric term vanishes and the matrix arising from FE discretization is

$$\begin{vmatrix} G\mathbf{A}_2 & -\mathbf{B}^T \\ -\mathbf{B} & -\tau k\,\mathbf{D} \end{vmatrix}. \tag{4.29}$$

The Schur complement of this matrix is

$$\mathbf{S} = \tau k\,\mathbf{D} + \frac{1}{G}\mathbf{B}\mathbf{A}_2^{-1}\mathbf{B}.$$

The properties of this matrix has been studied in [AGLR08] employing both $\mathbb{P}_2 - \mathbb{P}_1$ and $\mathbb{P}_1 - \mathbb{P}_1$ FE spaces.

The authors showed that, using Taylor-hood elements, the i-th row of $\mathbf{M}_S = \mathbf{B}\mathbf{A}_2^{-1}\mathbf{B}$ is such that

$$M_s\underline{p} = \frac{1}{6}p_{i-1} + \frac{2}{3}p_i + \frac{1}{6}p_{i+1}.$$

It coincides with a mass matrix assembled on the same mesh using $\mathbb{P}_1$ FEs. The Schur complement is hence the discretization of a reaction-diffusion problem, with diffusion parameter $k\tau$ and reaction coefficient G^{-1}. In order to avoid instabilities in the discrete solution of (4.27), the following bound should hold

$$\frac{h^2}{6Gk\tau} < 1.$$

Using the $\mathbb{P}_1 - \mathbb{P}_1$ FEs, the matrix M_S is such that

$$M_s\underline{p} = \frac{1}{4}p_{i-1} + \frac{1}{2}p_i + \frac{1}{4}p_{i+1}.$$

It is a matrix derived by the assembling of the following reference matrix

$$\mathbf{d} = \frac{1}{4}\begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix}. \tag{4.30}$$

Using the local maximum principle argument, we can derive the following stability bound

$$\frac{h^2}{4Gk\tau} < 1.$$

This ensures that the Schur complement is an M-matrix.

4.3.3 Effect on the solver

A further effect that depends on the time discretization parameter is the condition number of the matrix (4.29) [FGT01]. In order to quantify this effect, we performed numerical experiments where we studied the number of iterations necessary to reach the convergence of the linear solver BiCGStab with a tolerance of 10^{-12}. We used a mesh with $h_1 = 0.1$ and chose a reference time step $\tau_0 = h_1^2/(4Gk)$. In each subplot of Fig. 4.8, we reported the number of iterations for $\tau_i = \tau_0 2^i$ with $i = \{-8, -7, \ldots, 2\}$. In each subplot, a different mesh refinement has been considered, i.e. $h_k = h_1/2^k$ with $k = 2, 3, 4$. We observe that the minimum of iterations is reached when i and k are chosen such that

$$\frac{h_i^2}{4Gk_j\tau} = 1.$$

This is the transition value beyond which instabilities show up.

In Fig. 4.9, the same experiment with $\mathbb{P}_2 - \mathbb{P}_1$ pair has been considered. Here, we set $\tau_0 = h_1^2/(6Gk)$. The minimum of linear solver iterations is reached when i and k are chosen such that

$$\frac{h_i^2}{6Gk_j\tau} = 1.$$

4.4 Stability of two-dimensional poroelastic system

In the bidimensional case, the Schur complement is more complicated and takes the form

$$\mathbf{S} = \tau k \mathbf{D} + \mathbf{M}_S,$$

with $\mathbf{M}_S = \mathbf{B}\mathbf{A}^{-1}\mathbf{B}^T$. The deviatoric component does not vanish and hence a free parameter exists even when applying the changes of variables (4.5) or (4.6).

When $G = \mu/d$, the matrix $\mathbf{A}$ is the Schur complement of the Stokes problem. In the fluid-dynamics community, it is well-known that $\mathbf{M}_S$ is spectrally equivalent to a mass matrix assemble on the pressure space. Differently from the Stokes case, the matrix $\mathbf{A}$, and hence the matrix $\mathbf{M}_S$, depends on both the mechanical parameters μ and G.

In the two- and three-dimensional cases, the poroelasticity system cannot be reduced to an equivalent elliptic problem. In this section we derive an empirical stability

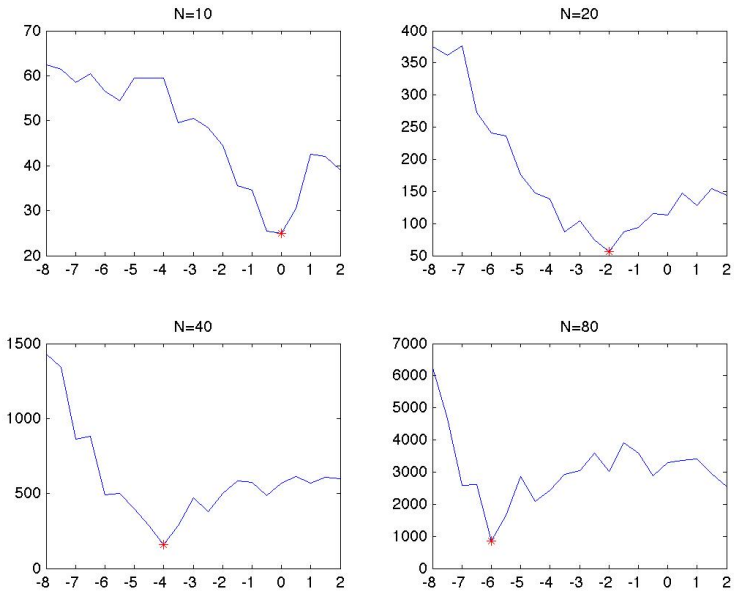


Figure 4.8. Comparison of the effect of τ on the number of BiCGStab iterations for $\mathbb{P}_1 - \mathbb{P}_1$ elements.

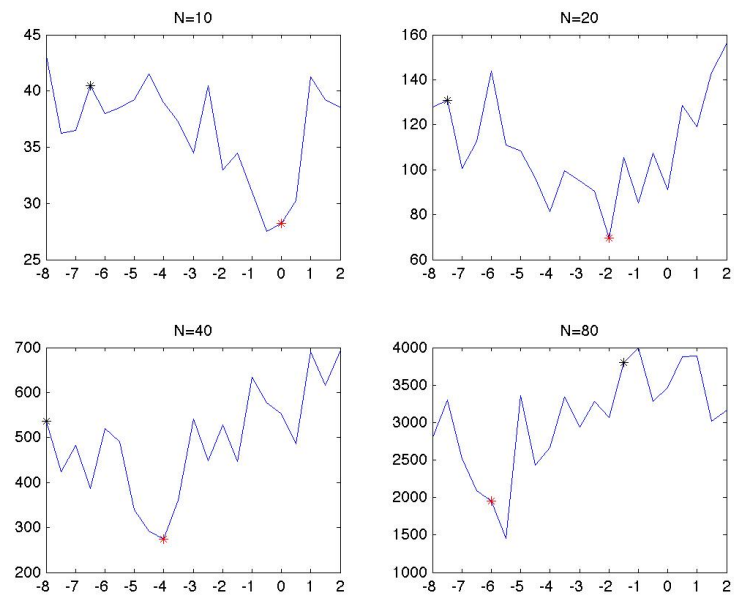


Figure 4.9. Comparison of the effect of τ on the number of BiCGStab iterations for $\mathbb{P}_2 - \mathbb{P}_1$ elements.

limit by means of numerical experiments on $\mathbf{M}_S$. This will be done extending the DMP. For a reaction-diffusion problem, a discretization that respects the DMP does not have negative values in the solution. **Our hypothesis is to find for which value of the mesh-size the minimum of the discrete Green's function stabilizes.** We will perform three different experiments.

1. We set $G = 1$ and study the matrix $\mathbf{M}_S$ for $h \to 0$ and for different mesh-size. We show that the norm $\|(\mathbf{M}_S - \mathbf{M})/h^2\|_{\max} \to 0$ for $\mu \to 0$. Hence, for $G \gg \mu$, the stability limit for this problem coincides with the one of the associated reaction-diffusion problem.

2. We set $\mu = 1$ and $G = 0$. We study Green's functions of the Schur complement $\mathbf{S}$. In particular, we study the minimum of the solution for different mesh step-size. We show that the minimum of the solution stabilizes at a mesh size that is the half of the stability limit of the equivalent reaction-diffusion problem.

3. We perform the computation at item 2 with μ and G different from zero. We show that the two derived stability rules sum up.

The matrix norm $\|\cdot\|_{\max}$ is the so-called uniform norm that corresponds to the maximum of the absolute values of its entries:

$$\|\mathbf{A}\|_{\max} = \max_{i,j} |\mathbf{A}_{ij}|.$$

Since in our numerical experiments we have to explicitly compute $\mathbf{A}^{-1}$, the considered meshes are quite coarse. The considered domain is $\Omega = (0,1)^2$. Mesh resolution is the same on both directions. The edge of the domain is divided into N intervals. The elasticity matrix $\mathbf{A}$ has to be invertible and hence we have to set boundary conditions on it. In our experiments, we consider four different boundary conditions that are reported in Fig. 4.10. We will identify them by

a) minimal boundary conditions,

b) unconfined compression,

c) confined compression,

d) rigid top-bottom sides.

Note that in case a) the minimum number of boundary conditions that make $\mathbf{A}$ invertible are imposed. In the two-dimensional case, $\mathbf{A}$ has three null eigenvalues that correspond to the three rigid motions of the body. Hence, three constraints have to be imposed to render this matrix invertible.

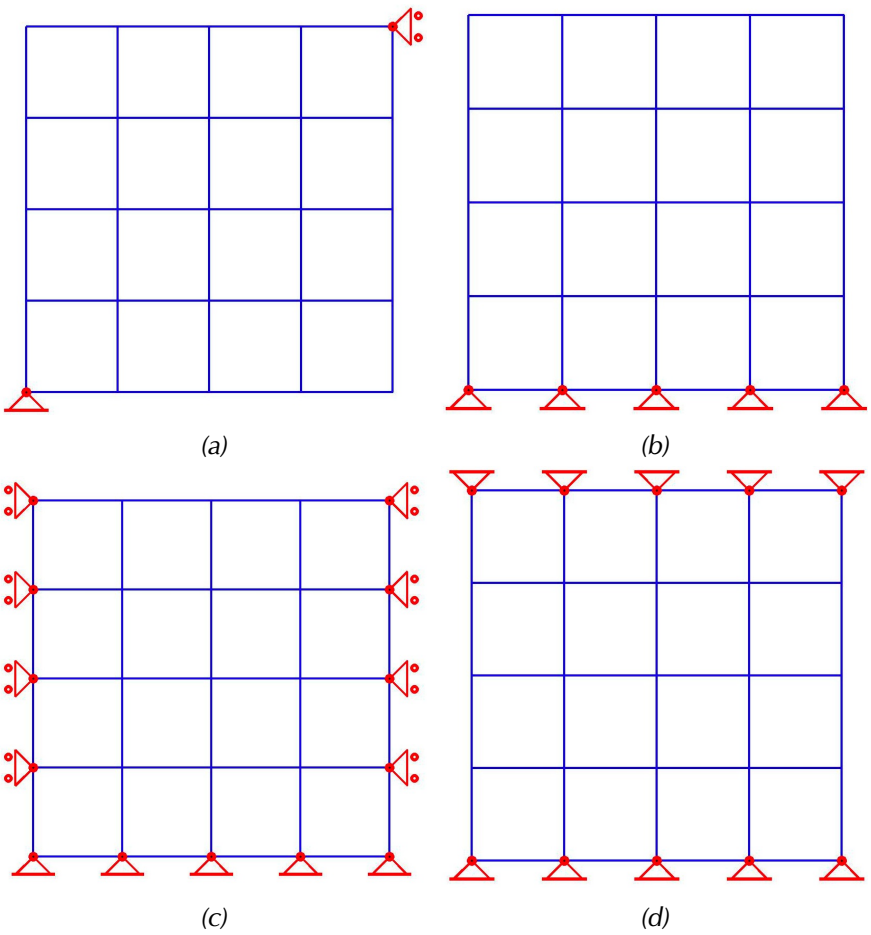


Figure 4.10. Representation of the computational domain $\Omega = (0,1)^2$ with $N = 4$: minimal boundary conditions (4.10a), unconfined compression (4.10b), confined compression (4.10c), and rigid top-bottom sides (4.10d).

4.4.1 Effect of bulk modulus

In the first numerical example, we perform the change of variables (4.5). We set $G = k = 1$ and we have left only a free parameter μ. We study the effect of mesh refinements and shear modulus on the absolute norm $N^2\|\mathbf{M} - \mathbf{M}_S\|_{\max}$.

In both cases a) and b), the matrix $\mathbf{M}_S$ tends to the pressure Schur matrix for $\mu \to 0$. They are reported in Fig. 4.11. We can observe how the norm does not depend on the mesh refinement. A similar behaviour has been obtained also for case c) and d).

In other numerical experiments, we observed that

- for $\mathbb{P}_k - \mathbb{P}_{k-1}$ couples, the matrix $\mathbf{M}_S$ tends to the mass matrix assembled on the $\mathbb{P}_{k-1}$ space for $\mu \to 0$;
- for $\mathbb{Q}_k - \mathbb{Q}_{k-1}$ couples, the same behaviour have been observed but only for structured meshes.

In the $\mathbb{Q}_1 - \mathbb{Q}_1$ case, the matrix $\mathbf{M}_S$ is not a mass matrix. Driven by the monodimensional example, we studied the convergence of $\mathbf{M}_S$ to a different matrix $\mathbf{M}_R$. It is a sort of mass matrix assembled with the following reference matrix

$$\mathbf{m}_R = \frac{1}{16}\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 4 \end{vmatrix}. \tag{4.31}$$

Norm $\|\mathbf{M}_S - \mathbf{M}_R\|_{\max}$ is reported in Tab. 4.1. We can observe a similar behaviour as in the $\mathbb{Q}_2 - \mathbb{Q}_1$ case.

For the $\mathbb{Q}_2 - \mathbb{Q}_1$ case, if $\mu \ll 1$ the stability of the system can be studied as the one of the equivalent reaction-diffusion problem. For example, in this case the stability limit reads

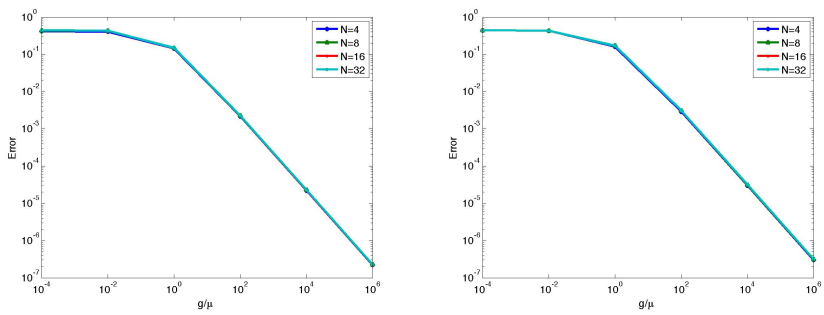


Figure 4.11. Error between Schur complement and pressure mass matrix for several values of μ. Left: Boundary conditions a). Right: boundary conditions b).

μ	$N=8$	$N=16$	$N=32$	$N=64$
10^4	4.4×10^{-1} (0.30)	4.4×10^{-1} (0.36)	4.4×10^{-1} (0.38)	4.4×10^{-1} (0.39)
10^2	4.3×10^{-1} (0.62)	4.3×10^{-1} (0.63)	4.3×10^{-1} (0.63)	4.3×10^{-1} (0.63)
1	1.6×10^{-1} (0.95)	1.7×10^{-1} (0.95)	1.7×10^{-1} (0.95)	1.7×10^{-1} (0.95)
10^{-2}	2.9×10^{-3} (0.99)	3.1×10^{-3} (0.99)	3.1×10^{-3} (0.99)	3.1×10^{-3} (0.99)
10^{-4}	3.0×10^{-5} (1.00)	3.1×10^{-5} (1.00)	3.2×10^{-5} (1.00)	3.2×10^{-5} (1.00)
10^{-6}	3.0×10^{-7} (1.00)	3.1×10^{-7} (1.00)	3.2×10^{-7} (1.00)	3.2×10^{-7} (1.00)

Table 4.1. $\mathbb{Q}_1-\mathbb{Q}_1$ case: $\|N^2(\mathbf{M}_R-\mathbf{M}_S)\|_{\max}$ and, in parenthesis, the sum of the entries of M_S.

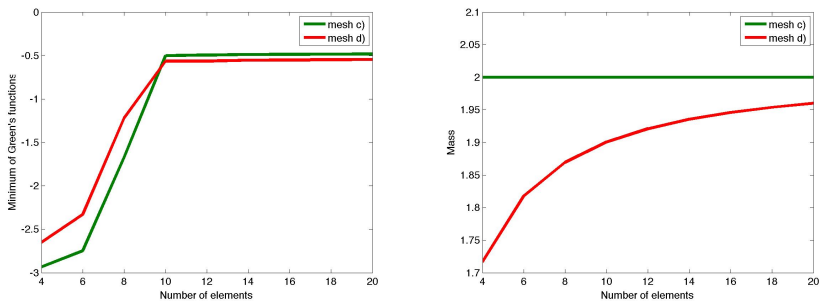
$$\frac{h^2}{3\tau kG}<1. \tag{4.32}$$

4.4.2 Effect of shear modulus

We study the effect of the shear modulus on the Schur complement $\mathbf{S}$. We perform the change of variables (4.5) and, hence, we can set $\tau=\mu=k=1$. To study its effect independently on the bulk modulus we set $G=0$. If the bulk modulus is null, boundary conditions a) cannot be applied. Thus, we focus on cases c) and d). In these cases, the DMP argument of $\mathbf{S}$ cannot be applied.

In order to understand the effect of μ on the oscillations, we compute a Green's function of the Schur complement and we study its minimum as a function of mesh refinement. In the reference case, we set $N=10$ and hence $\tau=2/(3\cdot10^2)$. The coefficient 3 derives from the stability of reaction-diffusion problem.

From Fig. 4.12 we observe that oscillations disappear for $N=10$. This test has been



(a) *(b)*

Figure 4.12. Left: minimum of Green's function of $\mathbf{S}$ at the node $(0.5,0.5)$ with $G=0$; it stabilizes for $N=10$. Right: sum of the entries of $\mathbf{M}_S$.

realized for several mesh refinements and it provided the same results. We can conclude that the problem is stable if

$$\frac{h^2}{3\tau k \frac{\mu}{2}} < 1. \tag{4.33}$$

4.4.3 Combined effect of shear and bulk moduli

We now study the minimum of the Green's function at point $(0.5, 0.5)$ for a general case with μ and G different from zero. We assume that the contribution to the stability rule of these parameters is linear. Thus, we suppose that the problem is stable if

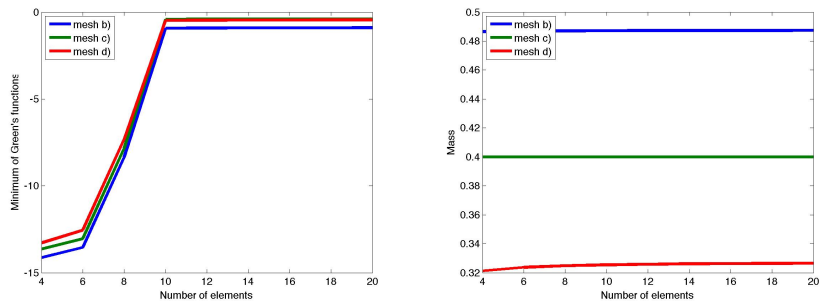
$$\frac{h^2}{3\tau k \left(\frac{\mu}{2} + G\right)} < 1. \tag{4.34}$$

Here, we set a reference value of $N = 10$ and we set $\tau = 1/\left((\mu/2 + G)\, 10^2\right)$. In Fig. 4.13, the minimum of a Green's function of the Schur complement as a function of mesh refinement is reported. Our hypothesis is confirmed from the numerical experiments. The value $N = 10$ is the one for which the minimum of the Green's function is stable.

4.4.4 Perfusion experiment

In this section we compare the limit presented in [VV81], in the following called V&V and the one we have derived. We consider a perfusion test, i.e. a poroelastic medium to which we apply two different pressures at the boundary. The bottom side is clamped. The boundary conditions are reported in Fig. 4.14.

The domain is the unit square and the physical parameters are $\mu = 0.3137$ Pa m^{-2}, $G = 0.6703$ Pa m^{-2}, and $k = 3.6454$ m^4 Pa^{-1} s^{-1}.



(a) *(b)*

Figure 4.13. Left: minimum of Green's function of $\mathbf{S}$ at the node $(0.5, 0.5)$ with $G = 2$; it stabilizes for $N = 10$. Right: sum of the entries of $\mathbf{M}_S$.

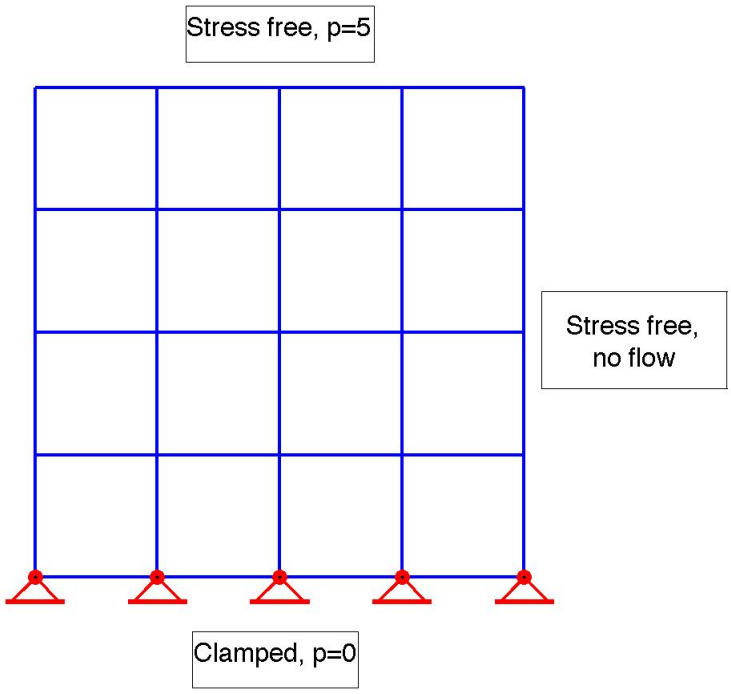


Figure 4.14. Boundary conditions for the perfusion test.

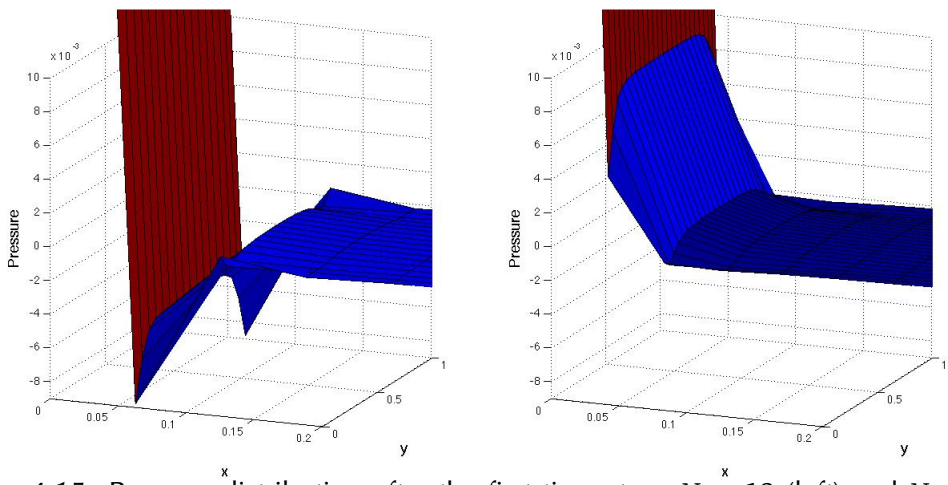


Figure 4.15. Pressure distribution after the first time step: $N = 13$ (left) and $N = 17$ (right).

We employ a time step $\tau = 2.5 \times 10^{-4}$s. The stability limit according to V&V gives a critical value $N_V = 13$. Our approach gives a value $N_F = 17$.

In Fig. 4.15 the numerical results with the two values of N are reported. In the plot on the left, we can observe that the resolution given by the V&V limit is not enough to detect the boundary layer, while with the limit presented in our work no instabilities occur.

4.5 Summary of stability and stabilization

By means of numerical tests we derived a stability condition that prevents oscillations in the pressure of the discretized poroelasticity problem. We obtained the condition

$$\frac{h^2}{C\tau k\,\mathfrak{B}(\mu, G)} < 1. \tag{4.35}$$

The left hand side of this inequality is the Péclet number of the poroelasticity problem. For Taylor-Hood FEs, the constant $C = C_\beta^{FE_d}$ is the stability constant in the Péclet number of the reaction-diffusion problem build on the pressure space. For $\mathbb{P}_1 - \mathbb{P}_1$ and $\mathbb{Q}_1 - \mathbb{Q}_1$ FEs, C has to be derived studying the stability of the problem $\tau k\mathbf{D} + \mathbf{M}_R$.

The function $\mathfrak{B}$ takes the form

$$\mathfrak{B}^{FE} = D\mu + G.$$

The constant $D = D^{FE}$ depends on the FE discretization. The constants C and D are reported in Tab. 4.2 for several FEs and dimensions. For triangular and tetrahedral elements C depends on the geometry of the element. In the table, the value is reported for an equilateral triangle in 2D and for a regular tetrahedron in 3D.

FE	d	β	C	D
$\mathbb{P}_2 - \mathbb{P}_1$	1		6	0
$\mathbb{P}_1 - \mathbb{P}_1$	1		4	0
$\mathbb{P}_2 - \mathbb{P}_1$	2		8	1
$\mathbb{P}_1 - \mathbb{P}_1$	2		6	1
$\mathbb{Q}_2 - \mathbb{Q}_1$	2	0	3	1/2
$\mathbb{Q}_2 - \mathbb{Q}_1$	2	1	0.8169	1/2
$\mathbb{Q}_1 - \mathbb{Q}_1$	2	0	8/3	1/2
$\mathbb{P}_2 - \mathbb{P}_1$	3		12	1
$\mathbb{P}_1 - \mathbb{P}_1$	3		48/5	1
$\mathbb{Q}_2 - \mathbb{Q}_1$	3	0	1.74	1/2
$\mathbb{Q}_2 - \mathbb{Q}_1$	3	1	unstable	
$\mathbb{Q}_1 - \mathbb{Q}_1$	3	0	1.66	1/2

Table 4.2. Stability coefficients for condition (4.35).

Thanks to the derived condition, we can extend the FPL stabilization to the multi-dimensional case. If condition (4.35) is not respected on a mesh, the permeability is changed accordingly to the following formula

$$k_h = k\left(1 + \frac{h^2}{C_\beta^{FE_d}\tau\,\mathfrak{B}^{FE}}\right).$$

This ensures that the Péclet number of the stabilized problem is small than one.

Chapter 5

Solution algorithms and numerical approximation

In this chapter we present a solution method for the non-linear poroelasticity system including inelastic phenomena. We first introduce the weak formulation of this problem. Then we present the solution method based on: implicit Euler scheme, quasi-Newton method, and FE discretization. In order to validate our method and the code we have implemented, we compare the results of a non-linear confined compression case against the simulation obtained with a monodimensional code. Finally, we test our parallel multigrid implementation for the discrete poroelastic system. Part of the content of this chapter has been published in [GGF+12].

5.1 Introduction

The coupled system (2.85) is a non-linear system of PDEs coupled with a tensorial ODE. Eq. (2.85.1) is a stationary PDE and it depends non-linearly on the displacement and on the tensor $\mathbf{F}_a$. Eq. (2.85.2) is an evolution PDE for the deformation determinant J and it is a Poisson equation for the pressure π. Eq. (2.85.3) is a first-order ODE for the tensor $\mathbf{F}_a$. It is linear with respect to the unknown $\mathbf{F}_a$ but the tensor $\mathbf{L}_a$ may also depend on $\mathbf{u}$ and π.

From a mechanical standpoint, Eq. (2.85.1) describes the conservation of linear and angular momentum. The pressure π is an additional variable that allows to impose the incompressibility of the constituents. Eq. (2.85.2) is the constraint that derives from the conservations of mass of both constituents.

Instead, from a temporal standpoint, system (2.85) is a DAEs system. The incompressibility constraint (2.85.2) governs the time evolution and it is the differential component of the system. This equation is the differential part of the system. The momentum equation (2.85.1) is its algebraic part and it acts as a constraint on the time evolution.

The differential problem describing the poroelasticity model is the ensemble of the equations (2.85), initial conditions (2.88), and boundary conditions (2.89). For the numerical analysis of this problem, no strategy has been presented yet in literature. To develop a suitable solution method, we observe that there are several points that have to be considered:

D1: the coupling between evolution equation of growth and the mechanical balance laws;

D2: the inherent non-linearity of the elastic problem;

D3: the saddle-point structure of the poroelastic problem;

D4: the DA structure of the poroelastic problem;

The coupling between the evolution of $\mathbf{F}_a$ (**D1**) and balance laws is similar to one treated in elasto-plasticity. The classical solution procedure for inhomogeneity due to plastic flow is the return mapping algorithm [SH98]. If $\mathbf{L}_a$ is a function of the current displacement and pressure, an explicit treatment of (2.85.3) may not be sufficient and hence the return mapping algorithm has to be adapted to growth problem. We will not consider this in our solution strategy. We introduce the hypothesis that the growth tensor is a given function of space and time. This means that its evolution may be studied independently on the other variables. Hence, at each time step growth can be considered as already occurred and, hence, $\mathbf{F}_a$ can be considered as data of the problem.

Time integration (**D3**) and algorithms for saddle-point problems (**D4**) are strictly connected. For both of them a coupled and decoupled approach exists. Coupled methods are usually referred to as *monolithic methods*. Decoupled methods are usually referred to as *Uzawa methods* for the solution of the saddle-point systems and *splitting schemes* for time integration.

Coupled approaches consist in the blind application of a solution method, or time integration schemes, to a system of equations without considering the mathematical properties of its subparts. After spatial discretization, they require the solution of a single large system of equations. Instead, the decoupled approaches are based on the mathematical (and physical) properties of the differential operators appearing in the equations of the system. They are based on an external loop in which at each iteration smaller linear systems are solved. The decoupled approach results particularly efficient in the linear case, when symmetric and positive definite matrices have to be inverted and efficient preconditioners are known.

A comparison between monolithic and splitting schemes for time integration has been presented in [MHE10]. Splitting scheme are based on the linearity of the differential operators and hence they are not suitable for non-linear problems. Hence, we consider monolithic time integration.

Since we are going to use the FEM for the spatial discretization of poroelasticity problem, we have first to introduce its weak formulation. This can be seen also as an extension of the principle of virtual work for non-linear poroelastic problem.

In the following sections of this chapter, we will present a discretization and solution method based on the following steps:

1. weak formulation of the differential problem;
2. implicit Euler scheme for the time integration;
3. quasi-Newton method for the solution of the non-linear problem;
4. finite element method for the spatial discretization.

The poroelasticity model derived in the first two chapters depends only on quantities associated to the solid phase that were denoted by the subscript s. In the following it will be neglected for the sake of simpler notation. The test functions used in the weak formulation will be denoted by $\mathbf{v}$ and q. The increment functions used in the quasi-Newton's method will be denoted by $\mathbf{h}$. The second order tensors corresponding to the derivative of $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{h}$ will be denoted by the capital corresponding letter: $\mathbf{U} = \nabla\mathbf{u}$, $\mathbf{V} = \nabla\mathbf{v}$, and $\mathbf{H} = \nabla\mathbf{h}$.

5.2 Weak formulation of the field equations

In this section, we introduce the weak form of the equations in (2.85). Eq. (2.85.1) is a vectorial equation. To obtain its weak form, we multiply by a test function $\mathbf{v} = \mathbf{v}(\mathbf{X}) \in V$. The space V will be briefly discussed later. For now we suppose that test functions admit one weak derivative and that $\mathbf{v}|_{\Gamma_D^{\mathbf{u}}} = \mathbf{0}$. The scalar multiplication of Eq. (2.85.1) by $\mathbf{v}$ gives a scalar function at each point $\mathbf{X} \in \mathscr{B}_r$. Integrating this scalar product over the domain, we get

$$-\int_{\mathscr{B}_r} \mathrm{Div}\,(\mathbf{P} - J\pi\mathbf{F}^{-T}) \cdot \mathbf{v}\,\mathrm{d}\mathbf{X} = 0. \tag{5.1}$$

In the field of continuum mechanics, the function $\mathbf{v}$ is called *virtual velocity*. It is usually denoted by $\boldsymbol{v}_v$ and its gradient is denoted by $\dot{\mathbf{F}}_v = \mathrm{Grad}(\boldsymbol{v}_v)$. Applying integration by parts to Eq. (5.1), we obtain

$$\int_{\mathscr{B}_r} (\mathbf{P} - J\pi\mathbf{F}^{-T}) \colon \nabla\mathbf{v}\,\mathrm{d}\mathbf{X} - \int_{\Gamma_D^{\mathbf{u}}} (\mathbf{P} - J\pi\mathbf{F}^{-T})\mathbf{N}\cdot\mathbf{v}\mathrm{d}\mathbf{A} - \int_{\Gamma_N^{\mathbf{u}}} (\mathbf{P} - J\pi\mathbf{F}^{-T})\mathbf{N}\cdot\mathbf{v}\mathrm{d}\mathbf{A} = 0. \tag{5.2}$$

In the last equality, we exploited additivity of the boundary integral to separate the integration over the Dirichlet and the Neumann boundary. Since the test functions are

chosen such that they vanish on the Dirichlet boundary, the integral over $\Gamma_D^{\mathbf{u}}$ is zero. The boundary condition (2.89.2) provides the value of the Piola-Kirchhoff stress tensor in the normal direction. Substituting this into (5.2), we obtain

$$\int_{\mathscr{B}_r} (\mathbf{P} - J\pi\mathbf{F}^{-T}) : \mathbf{V} \, \mathrm{d}\mathbf{X} = \int_{\Gamma_N^{\mathbf{u}}} \mathbf{f}_{rb} \cdot \mathbf{v} \, \mathrm{d}\mathbf{A}, \tag{5.3}$$

that is the weak form of momentum equation. Substituting (2.86) into the last equation, we can express the dependance of $\mathbf{P}$ on $\mathbf{F}$ and $\mathbf{F}_a$. In this case the weak form reads

$$\int_{\mathscr{B}_r} (J_a \mathbf{F}\mathbf{F}_a^{-1}\mathbf{S}_n\mathbf{F}_a^{-T}) : \mathbf{V} \, \mathrm{d}\mathbf{X} - \int_{\mathscr{B}_r} J\pi\mathbf{F}^{-T} : \mathbf{V} \, \mathrm{d}\mathbf{X} = \int_{\Gamma_N^{\mathbf{u}}} \mathbf{f}_{rb} \cdot \mathbf{v} \, \mathrm{d}\mathbf{A}, \tag{5.4}$$

where we remind that the second Piola-Kirchhoff stress tensor in the natural configuration can be function only of $\mathbf{F}_e$, and hence $\mathbf{S} = \mathbf{S}(\mathbf{F}\mathbf{F}_a^{-T})$. From Eq. (5.3), we see that the conjugate variable to the first Piola stress tensor is the gradient of a virtual velocity. Since the first Piola stress tensor is $\mathbf{P} = \mathbf{F}\mathbf{S}$, the total stress in (5.3) can also be written as $\mathbf{F}(\mathbf{S} - J\pi\mathbf{C}^{-1})$. Thanks to the properties of the tensor scalar product and the trace, the following identities hold true

$$\begin{aligned} \mathbf{F}(\mathbf{S} - J\pi\mathbf{C}^{-1}) : \mathbf{V} &= \mathrm{tr}\,(\mathbf{F}(\mathbf{S} - J\pi\mathbf{C}^{-1})\mathbf{V}^T) \\ &= \mathrm{tr}\,((\mathbf{S} - J\pi\mathbf{C}^{-1})\mathbf{V}^T\mathbf{F}) \\ &= (\mathbf{S} - J\pi\mathbf{C}^{-1}) : \mathbf{F}^T\mathbf{V}. \end{aligned} \tag{5.5}$$

Since $\mathbf{S}$ and $\mathbf{C}^{-1}$ are symmetric tensor, the last equality can also be written as

$$(\mathbf{S} - J\pi\mathbf{C}^{-1}) : \mathbf{E}_l$$

with

$$\mathbf{E}_l = \frac{\mathbf{F}^T\mathbf{V} + \mathbf{V}^T\mathbf{F}}{2}. \tag{5.6}$$

The tensor $\mathbf{E}_l$ is the directional derivative of the Green-Lagrange strain tensor

$$\mathbf{E} = \frac{\mathbf{C} - \mathbf{I}}{2}.$$

Using the definition (5.6), the weak form (5.3) can be written as

$$\int_{\mathscr{B}_r} (\mathbf{S} - J\pi\mathbf{C}^{-1}) : \mathbf{E}_l \, \mathrm{d}\mathbf{X} = \int_{\Gamma_N^{\mathbf{u}}} \mathbf{f}_{rb} \cdot \mathbf{v} \, \mathrm{d}\mathbf{A}. \tag{5.7}$$

From (5.4), we see that the conjugate variable to the second Piola-Kirchhoff stress is the linearized Green-Lagrange strain.

In order to obtain the weak form of the conservation of mass, we multiply (2.85.2) by a test function $q = q(\mathbf{X}) \in W$. Functions in W are supposed to have at least one weak

derivative and to vanish on Γ_D^π. In the mechanical jargon, they are virtual pressures. After integrating over $\mathscr{B}_r$, we obtain

$$-\int_{\mathscr{B}_r} \dot{J} q \, \mathrm{d}\mathbf{X} + \int_{\mathscr{B}_r} \mathrm{Div}\,(\mathbf{K}_r \nabla \pi) q \, \mathrm{d}\mathbf{X} = -\int_{\mathscr{B}_r} J_a \phi_{sn} \mathrm{tr}(\mathbf{L}_a) q \, \mathrm{d}\mathbf{X}. \tag{5.8}$$

Applying integration by parts to the divergence term in (5.8), we get

$$\begin{aligned}\int_{\mathscr{B}_r} \mathrm{Div}\,(\mathbf{K}_r \nabla \pi) q \, \mathrm{d}\mathbf{X} = &-\int_{\mathscr{B}_r} (\mathbf{K}_r \nabla(\pi) \cdot \nabla q \, \mathrm{d}\mathbf{X} \\ &+ \int_{\Gamma_D^{\mathbf{u}}} (\mathbf{K}_r \nabla \pi) \cdot \mathbf{N} q \, \mathrm{d}\mathbf{A} \\ &+ \int_{\Gamma_N^{\mathbf{u}}} (\mathbf{K}_r \nabla \pi) \cdot \mathbf{N} q \, \mathrm{d}\mathbf{A}\end{aligned}$$

where we have already separated the boundary integral over the Dirichlet and the Neumann boundary. As done for the momentum equation, the integral over the Dirichlet boundary vanishes because of the choice of the test function. The value of the function over the Neumann boundary can be evaluated from (2.89.4). This leads to the following weak formulation of the fluid flow equation

$$-\int_{\mathscr{B}_r} \dot{J} q \, \mathrm{d}\mathbf{X} - \int_{\mathscr{B}_r} (\mathbf{K}_r \nabla \pi) \cdot \nabla q \, \mathrm{d}\mathbf{X} = -\int_{\Gamma_N^\pi} Q_{rb} q \, \mathrm{d}\mathbf{A} - \int_{\mathscr{B}_r} J_a \phi_{sn} \mathrm{tr}(\mathbf{L}_a) q \, \mathrm{d}\mathbf{X}. \tag{5.9}$$

In Eq. (2.85.3), no spatial derivatives appear and no weak formulation is needed. It is an evolution equation for the tensor $\mathbf{F}_a \in \mathrm{Lin}^+$. This is the space of tensor with positive determinant.

5.2.1 Definition of the functional spaces

The function spaces V and W have been discussed only for the linear case [ML92, Sho00]. They are the Hilbert spaces $V = H^1_{\Gamma_D^{\mathbf{u}}}(\mathscr{B}_r, \mathbb{R}^3)$ and $W = H^1_{\Gamma_D^{\mathbf{u}}}(\mathscr{B}_r)$.

The function spaces of the non-linear case are

$$V = W^{1,s}_{\Gamma_D^{\mathbf{u}}}(\mathscr{B}_r, \mathbb{R}^3) \quad \text{and} \quad W = W^{1,s^*}_{\Gamma_D^{\mathbf{u}}}(\mathscr{B}_r),$$

with $s, s^* \in \mathbb{R}$. The two unknowns of (5.10) are hence the functions

$$\mathscr{I} \in t \longrightarrow \mathbf{u}(\,\cdot\,, t) \in V,$$

and

$$\mathscr{I} \in t \longrightarrow \mathbf{q}(\,\cdot\,, t) \in W.$$

5.2.2 Weak problem

Following the notation used in [ML92], the weak form of (2.85) reads

For each $t \in \mathscr{I}$, find $\mathbf{u}(\cdot, t) \in \mathscr{U}$, $\pi(\cdot, t) \in \mathscr{U}^*$, and $\mathbf{F}_a(\cdot, t) \in \mathrm{Lin}^+$ such that:

$$\left\{\begin{array}{lcll} \int_{\mathscr{B}_r} (J_a \mathbf{F}\mathbf{F}_a^{-1}\mathbf{S}_n\mathbf{F}_a^{-T}) - \int_{\mathscr{B}_r} J\pi\mathbf{F}^{-T} : \mathbf{V}\,\mathrm{d}\mathbf{X} & = & \int_{\Gamma_N^{\mathbf{u}}} \mathbf{f}_{rb} \cdot \mathbf{v}\,\mathrm{d}\mathbf{A} & \forall \mathbf{v} \in V \\ -\dfrac{\mathrm{d}}{\mathrm{d}t}\displaystyle\int_{\mathscr{B}_r} Jq\,\mathrm{d}\mathbf{X} - \int_{\mathscr{B}_r} (\mathbf{K}_r \nabla\pi) \cdot \nabla q\,\mathrm{d}\mathbf{X} & = & \int_{\mathscr{B}_r} J_a \phi_{sn} \mathrm{tr}(\mathbf{L}_a) q\,\mathrm{d}\mathbf{X} & \forall q \in W \\ \dot{\mathbf{F}}_a & = & \mathbf{L}_a\mathbf{F}_a & \text{on } \mathscr{B}_r, \end{array}\right. \tag{5.10}$$

where $\mathscr{U} \subset V$ and $\mathscr{U}^* \subset W$.

A simpler problem can be obtained if we consider the very particular case in which the tensor of anelastic deformation is kept constant (i.e. $\mathbf{F}_a$ is constant and known from the outset). Hence, since no anelastic evolution occurs, the tensor $\mathbf{L}_a$ is zero. This means that anelastic deformations have already taken place, namely, the tissue has already grown and/or remodelled. Physically, this can be rephrased by saying that the tissue grows and remodels over a time scale much larger than the scale over which fluid flows and elastic deformations take place. In this case, since $\mathbf{F}_a$ is known, the weak problem can be reformulated as

For each $t \in \mathscr{I}$, find $\mathbf{u}(\cdot, t) \in \mathscr{U}$ and $\pi(\cdot, t) \in \mathscr{U}^*$ such that:

$$\left\{\begin{array}{lcll} \int_{\mathscr{B}_r} (J_a \mathbf{F}\mathbf{F}_a^{-1}\mathbf{S}_n\mathbf{F}_a^{-T}) - \int_{\mathscr{B}_r} J\pi\mathbf{F}^{-T} : \mathbf{V}\,\mathrm{d}\mathbf{X} & = & \int_{\Gamma_N^{\mathbf{u}}} \mathbf{f}_{rb} \cdot \mathbf{v}\,\mathrm{d}\mathbf{A} & \forall \mathbf{v} \in V \\ -\dfrac{\mathrm{d}}{\mathrm{d}t}\displaystyle\int_{\mathscr{B}_r} Jq\,\mathrm{d}\mathbf{X} - \int_{\mathscr{B}_r} (\mathbf{K}_r \nabla\pi) \cdot \nabla q\,\mathrm{d}\mathbf{X} & = & 0 & \forall q \in W \end{array}\right. \tag{5.11}$$

This problem will used to present time integration and linearization of the poroelastic system. In particular, notice that if $\mathbf{F}_a$ is the identity tensor then the problem results in the standard non-linear poroelasticity problem.

5.3 Time discretization

From a temporal standpoint, the system (2.85), or its weak counterpart (5.11), is a system of differential-algebraic equations (DAEs). In fact, while possible time dependencies can be found in the problem data (i.e. external forces, boundary conditions), time derivative features explicitly only in Eq. (2.85.2). Hence, while the unknown of the momentum equation is the displacement, the time evolution of the *differential variable* $\mathbf{u}$ is governed by the mass conservation (2.85.2) through the deformation determinant

J. Eq. (2.85.1) acts as a constraint for the time evolution of the *algebraic variable* π.

In order to introduce a family of numerical approximations of such a problem, we consider the time interval $\mathscr{I} = (0, t_f]$ and we introduce on it a partition into N subintervals (t^{m-1}, t^m) of size $\tau_i = t^m - t^{m-1}$. In case of uniform discretization, time-step size will be denoted by $\tau = 1/N$ and a node will be given by $t^m = \tau m$ with $m = 0, 1, \ldots N$. The symbol $\mathbf{A}^m = \mathbf{A}^m(\mathbf{X})$ denotes the approximation of a generic tensor function $\mathbf{A}(\mathbf{X}, t)$ at node t^m, i.e. $\mathbf{A}^m \approx \mathbf{A}(\mathbf{X}, t^m)$.

We approximate the time derivative of J with a finite difference and we use an implicit scheme [QSS00]. Thus, we write

$$\dot{J}^m \approx \frac{J^m - J^{m-1}}{\tau}.$$

Hence, using an implicit Euler scheme, the approximation of the system of equations (5.4) and (5.9) can be written in the following way

For each $t \in \mathscr{I}$, find $\mathbf{u}(\cdot, t) \in \mathscr{U}$, $\pi(\cdot, t) \in \mathscr{U}^*$, and $\mathbf{F}_a(\cdot, t) \in \mathrm{Lin}^+$ such that:

$$\left\{ \begin{array}{lllll} \int_{\mathscr{B}_r} (J_a \mathbf{F}^{\mathbf{m}} \mathbf{F}_a^{-1} \mathbf{S}_n^m \mathbf{F}_a^{-T}) & - \int_{\mathscr{B}_r} J^m \pi^m (\mathbf{F}^{-T})^m : \mathbf{V} \, \mathrm{d}\mathbf{X} & = & \int_{\Gamma_N^{\mathbf{u}}} \mathbf{f}_{rb}^m \cdot \mathbf{v} \, \mathrm{d}\mathbf{A} & \forall \mathbf{v} \in V \\ - \int_{\mathscr{B}_r} J^m q \, \mathrm{d}\mathbf{X} & - \tau \int_{\mathscr{B}_r} (\mathbf{K}_r^m \nabla \pi^m) \cdot \nabla q \, \mathrm{d}\mathbf{X} & = & - \int_{\mathscr{B}_r} J^{m-1} q \, \mathrm{d}\mathbf{X} & \forall q \in W \end{array} \right. \tag{5.12}$$

Formulation using second Piola-Kirchhoff tensor

An alternative form of Eq. (5.11) can be derived based on (5.7). In this case the weak form of the discretization in time of the momentum equation reads

$$\int_{\mathscr{B}_r} \mathbf{S}^m : \mathbf{E}_l^m \, \mathrm{d}\Omega - \int_{\mathscr{B}_r} J^m \pi^m (\mathbf{C}^m)^{-1} : \mathbf{E}_l^m \, \mathrm{d}\Omega \quad = \quad \int_{\Gamma_N^{\boldsymbol{u}}} \boldsymbol{\nu}_\nu \cdot \boldsymbol{f}_{rb}^m \, \mathrm{d}A = 0, \tag{5.13}$$

where $\mathbf{C}^m = (\mathbf{F}^m)^T \mathbf{F}^m$, $\mathbf{S}^m := (\mathbf{F}^m)^{-1} \mathbf{P}^m$, and $\mathbf{E}_l^m$ is given by

$$\mathbf{E}_l^m := \tfrac{1}{2} \left[(\mathbf{F}^m)^T \mathbf{V} + (\mathbf{V})^T \mathbf{F}^m \right]. \tag{5.14}$$

In this notation, the Green-Lagrange strain tensor at time t^m reads $\mathbf{E}^m = \frac{1}{2}(\mathbf{C}^m - \mathbf{I})$. This weak form is the one used in [Bon08].

5.3.1 Time integration of $\mathbf{F}_a$

If we consider the problem (5.10) where we have a time evolution of the inhomogeneity $\mathbf{F}_a$, we have also to discretize (2.85.3) in time. If we use a θ-method this equation can be written as

$$\frac{\mathbf{F}_a^m - \mathbf{F}_a^{m-1}}{\tau} = \left(\theta \mathbf{L}_a^m \mathbf{F}_a^m + (1-\theta)\mathbf{L}_a^{m-1}\mathbf{F}_a^{m-1}\right).$$

We notice that if $\theta = 1$ we obtain the implicit Euler scheme, if $\theta = 0$ we obtain the explicit Euler scheme, and if $\theta = 0.5$ we obtain the Crank-Nickolson method. Collecting $\mathbf{F}_a^m$ at the left-hand side and $\mathbf{F}_a^{m-1}$ at the right-hand side, a step of the method can be written as

$$(\mathbf{I} - \tau\theta \mathbf{L}_a^m)\mathbf{F}_a^m = (\mathbf{I} + \tau(1-\theta)\mathbf{L}_a^{m-1})\mathbf{F}_a^{m-1}. \tag{5.15}$$

For each material point, eq. (5.15) has be solved indipendently. The iteration matrix of this scheme is

$$\mathbf{B} = (\mathbf{I} - \tau\theta \mathbf{L}_a^m)^{-1}(\mathbf{I} + \tau(1-\theta)\mathbf{L}_a^{m-1}). \tag{5.16}$$

Let us consider the case in which $\mathbf{L}_a$ is constant over the time step. In section 2.2.1, we showed that if $\mathrm{tr}\,(\mathbf{L}_a) = 0$ the evolution of $\mathbf{F}_a$ preserves the determinant J_a. In order to get this property also at the discrete level, the matrix $\mathbf{B}$ should have the unitary determinant.

A particular case is obtained if $\mathbf{L}_a$ has eigenvalues $\{-\lambda, 0, \lambda\}$ with $\lambda \in \mathbb{R}$. Here, the choice $\theta = 0.5$ gives the condition $\det(\mathbf{B}) = 1$.

For general forms $\mathbf{L}_a$ with $\mathrm{tr}\,\mathbf{L}_a = 0$, a possible solution would be to choose at each point a different θ such that $\det(\mathbf{B}) = 1$. This would lead to different schemes at different points of the domain.

5.4 Newton's method for poroelasticity system

Formally, the semi-discrete problem (5.12) can be rewritten in compact form as

$$\mathfrak{F}_{\boldsymbol{u}}(\boldsymbol{u}^m, \pi^m, \mathbf{v}) = 0, \tag{5.17}$$

$$\mathfrak{F}_{\pi}(\boldsymbol{u}^m, \pi^m, q) = 0, \tag{5.18}$$

where $\mathfrak{F}_{\boldsymbol{u}}$ and $\mathfrak{F}_{\pi}$ are functionals of displacement and pressure, and of their virtual counterparts, i.e. the test functions $\mathbf{v}$ and q. In the sequel, we drop the dependence of these functionals on $\mathbf{v}$ and q for the sake of a simpler notation.

The non-linear system of equations (5.17) and (5.18) is solved by Newton's method [Bon08, QSS00]. At the k-th iteration, the linearized system reads

$$\begin{aligned}
\mathfrak{F}_{\boldsymbol{u}}(\boldsymbol{u}^{m,k-1},\pi^{m,k-1}) &+ \mathfrak{D}_{\mathbf{u}}\mathfrak{F}_{\boldsymbol{u}}(\boldsymbol{u}^{m,k-1},\pi^{m,k-1})[\mathbf{h}^{m,k-1}] \\
&+ \mathfrak{D}_{\pi}\mathfrak{F}_{\boldsymbol{u}}(\boldsymbol{u}^{m,k-1},\pi^{m,k-1})[\theta^{m,k}] = 0, \qquad (5.19)\\
\mathfrak{F}_{\pi}(\boldsymbol{u}^{m,k-1},\pi^{m,k-1}) &+ \mathfrak{D}_{\mathbf{u}}\mathfrak{F}_{\pi}(\boldsymbol{u}^{m,k-1},\pi^{m,k-1})[\mathbf{h}^{m,k-1}] \\
&+ \mathfrak{D}_{\pi}\mathfrak{F}_{\pi}(\boldsymbol{u}^{m,k-1},\pi^{m,k-1})[\theta^{m,k}] = 0. \qquad (5.20)
\end{aligned}$$

where $\mathfrak{D}_{\mathbf{u}}\mathfrak{F}_{\alpha}(\mathbf{u},\pi)[\mathbf{h}]$ and $\mathfrak{D}_{\pi}\mathfrak{F}_{\alpha}(\mathbf{u},\pi)[\theta]$ denote the Gateaux differentials with respect to $\mathbf{u}$ and with respect to π of a functional $\mathfrak{F}_{\alpha}$ evaluated at $(\mathbf{u},\pi)$. The functions $\mathbf{h}$ and θ are the directions along which the differential is computed. Given a pair $(\boldsymbol{u}^{m,k-1},\pi^{m,k-1})$, in the neighborhood of which the functionals (5.17) and (5.18) are linearized, the increments $\boldsymbol{h}^{m,k}$ and $\theta^{m,k}$ are computed by solving the following linear problem whose unknown is the couple $(\boldsymbol{h}^{m,k-1},\theta^{m,k-1})$

$$\begin{aligned}
&\mathfrak{D}_{\mathbf{u}}\mathfrak{F}_{\boldsymbol{u}}(\boldsymbol{u}^{m,k-1},\pi^{m,k-1})[\mathbf{h}^{m,k-1}] + \mathfrak{D}_{\pi}\mathfrak{F}_{\boldsymbol{u}}(\boldsymbol{u}^{m,k-1},\pi^{m,k-1})[\theta^{m,k}] = -\mathfrak{F}_{\boldsymbol{u}}(\boldsymbol{u}^{m,k-1},\pi^{m,k-1}),\\
&\mathfrak{D}_{\mathbf{u}}\mathfrak{F}_{\pi}(\boldsymbol{u}^{m,k-1},\pi^{m,k-1})[\mathbf{h}^{m,k-1}] + \mathfrak{D}_{\pi}\mathfrak{F}_{\pi}(\boldsymbol{u}^{m,k-1},\pi^{m,k-1})[\theta^{m,k}] = -\mathfrak{F}_{\pi}(\boldsymbol{u}^{m,k-1},\pi^{m,k-1}).
\end{aligned}$$

The updated pair of fields $(\boldsymbol{u}^{m,k},\pi^{m,k})$ is then determined as follows

$$\boldsymbol{u}^{m,k} = \boldsymbol{u}^{m,k-1} + \boldsymbol{h}^{m,k}, \qquad (5.21)$$

$$\pi^{m,k} = \pi^{m,k-1} + \theta^{m,k}. \qquad (5.22)$$

Hence, in the continuation of this section we present the linearization of the terms appearing in (5.17) and (5.18).

5.4.1 Linearization of momentum equation

We now proceed in the linearization of the momentum equation (5.17). In this equation, we distinguish three terms:

- the constitutive part
$$\int_{\mathscr{B}_r} \mathbf{P} : \mathbf{V}\mathrm{d}\mathbf{X},$$
- the constraint arising from incompressibility
$$\int_{\mathscr{B}_r} -pJ\mathbf{F}^{-T} : \mathbf{V}\mathrm{d}\mathbf{X},$$
- the boundary forces.

We don't linearize the boundary integral because, if a Finite Element Method is used, it would require to know the connectivity from boundary elements and the connected volume element, but this information is not usually available in the standard finite element softwares. Hence we have to proceed in the linearization of the tensor $\mathbf{P}$ and of the constraint $-pJ\mathbf{F}^{-T}$.

Linearization of the first Piola-Kirchhoff tensor

Since the integral in (5.17) is a linear operator, we can linearize its argument. As a constitutive equation, we use the neo-Hookean material law (2.34), and hence the argument of the integral is

$$\mathbf{P}:\mathbf{V} = J_a\left(\mu_n\left(\mathbf{F}\mathbf{C}_a^{-1} - \mathbf{F}^{-T}\right) + \lambda_n \log\left(\frac{J}{J_a}\right)\mathbf{F}^{-T}\right):\mathbf{V}.$$

We write $\mathbf{P}$ as the sum of three terms

$$\mathbf{P} = \mathbf{P}_1 + \mathbf{P}_2 + \mathbf{P}_3,$$

with

$$\mathbf{P}_1 = J_a\mu_n\mathbf{F}\mathbf{C}_a^{-1}, \quad \mathbf{P}_2 = -J_a\mu_n\mathbf{F}^{-T}, \text{and} \quad \mathbf{P}_3 = \lambda_n \log\left(\frac{J}{J_a}\right)\mathbf{F}^{-T}.$$

The Gâteaux differentials of these three terms are

$$\mathfrak{D}_{\mathbf{u}}(\mathbf{P}_1:\mathbf{V})[\mathbf{h}] = J_a\mu_n\mathbf{H}\mathbf{C}_a^{-1}:\mathbf{V} = J_a\mu_n\mathbf{C}_a^{-1}:(\mathbf{H}^T\mathbf{V}),$$

$$\mathfrak{D}_{\mathbf{u}}(\mathbf{P}_2:\mathbf{V})[\mathbf{h}] = J_a\mu_n\mathbf{F}^{-T}\mathbf{H}^T\mathbf{F}^{-T}:\mathbf{V} = J_a\mu_n(\mathbf{H}\mathbf{F}^{-1})^T:(\mathbf{V}\mathbf{F}^{-1}),$$

$$\mathfrak{D}_{\mathbf{u}}(\mathbf{P}_3:\mathbf{V})[\mathbf{h}] = -\lambda_n \log\left(\frac{J}{J_a}\right)(\mathbf{H}\mathbf{F}^{-1})^T:(\mathbf{V}\mathbf{F}^{-1}) + \lambda_n(\mathbf{F}^{-T}:\mathbf{H})(\mathbf{F}^{-T}:\mathbf{V}).$$

Hence, we can write the linearization of the constitutive part of (5.17) at the point $(\mathbf{u}^{m,k}, \pi^{m,k})$ as

$$\mathfrak{D}_{\mathbf{u}}\left(\int_{\mathscr{B}_t}\mathbf{P}^{m,k}:\mathbf{V}\mathrm{d}\mathbf{X}\right)[\mathbf{h}^{m,k}] = a_1^{m,k}(\mathbf{h}^{m,k},\mathbf{v}) + a_2^{m,k}(\mathbf{h}^{m,k},\mathbf{v}) + a_3^{m,k}(\mathbf{h}^{m,k},\mathbf{v}),$$

where

$$a_1^{m,k}(\mathbf{h},\mathbf{v}) = \int_{\mathscr{B}_r} J_a\mu_n\mathbf{C}_a^{-1}:(\mathbf{H}^T\mathbf{V})\mathrm{d}\mathbf{X},$$

$$a_2^{m,k}(\mathbf{h},\mathbf{v}) = \int_{\mathscr{B}_r} J_a \left[\mu_n - \lambda_n \log\left(\frac{J^{m,k}}{J_a}\right)\right] (\mathbf{H}(\mathbf{F}^{m,k})^{-1})^T : (\mathbf{V}(\mathbf{F}^{m,k})^{-1})^T \mathrm{d}\mathbf{X},$$

$$a_3^{m,k}(\mathbf{h},\mathbf{v}) = \int_{\mathscr{B}_r} J_a \lambda_n ((\mathbf{F}^{m,k})^{-T} : \mathbf{H})((\mathbf{F}^{m,k})^{-T} : \mathbf{V}) \mathrm{d}\mathbf{X}.$$

Linearization of the incompressibility constraint

Now, we compute the directional derivative of

$$\int_{\mathscr{B}_r} -J\pi\mathbf{F}^{-T} : \mathbf{V}\mathrm{d}\mathbf{X}.$$

The derivative with respect to $\mathbf{u}$ is

$$\mathfrak{D}_{\mathbf{u}}\left(-J\pi\mathbf{F}^{-T} : \mathbf{V}\right)[\mathbf{h}] = J\pi(\mathbf{H}\mathbf{F}^{-1})^T : (\mathbf{V}\mathbf{F}^{-1}) - \pi J(\mathbf{F}^{-T} : \mathbf{H})(\mathbf{F}^{-T} : \mathbf{V}),$$

and the derivative with respect to π is

$$\mathfrak{D}_{\pi}\left(-J\pi\mathbf{F}^{-T} : \mathbf{V}\right)[\theta] = -J\theta\mathbf{F}^{-T} : \mathbf{V}.$$

Hence, we can write the linearization with respect to $\mathbf{u}$ of the incompressibility constraint part of (5.17) at the point $(\mathbf{u}^{m,k}, \pi^{m,k})$ as

$$\mathfrak{D}_{\mathbf{u}}\left(\int_{\mathscr{B}_r} -J^{m,k}\pi^{m,k}(\mathbf{F}^{m,k})^{-T} : \mathbf{V}\mathrm{d}\mathbf{X}\right)[\mathbf{h}^{m,k}] = a_4^{m,k}(\mathbf{h}^{m,k},\mathbf{v}) + a_5^{m,k}(\mathbf{h}^{m,k},\mathbf{v}),$$

where

$$a_4^{m,k}(\mathbf{h},\mathbf{v}) = \int_{\mathscr{B}_r} J^{m,k}\pi^{m,k}(\mathbf{H}(\mathbf{F}^{m,k})^{-1})^T : (\mathbf{V}(\mathbf{F}^{m,k})^{-1})\mathrm{d}\mathbf{X},$$

$$a_5^{m,k}(\mathbf{h},\mathbf{v}) = \int_{\mathscr{B}_r} -\pi^{m,k}J^{m,k}((\mathbf{F}^{m,k})^{-T} : \mathbf{H})((\mathbf{F}^{m,k})^{-T} : \mathbf{V})\mathrm{d}\mathbf{X}.$$

The linearization with respect to π of the incompressibility constraint part of (5.17) at the point $(\mathbf{u}^{m,k}, \pi^{m,k})$ is given by

$$\mathfrak{D}_{\pi}\left(\int_{\mathscr{B}_r} -J^{m,k}\pi^{m,k}(\mathbf{F}^{m,k})^{-T} : \mathbf{V}\mathrm{d}\mathbf{X}\right)[\mathbf{h}^{m,k}] = b^{m,k}(\mathbf{v},\theta^{m,k}).$$

5.4.2 Linearization of mass balance

We now show the linearization of the total mass balance equation (5.18). Also for this equation, we use a quasi-Newton method and we neglect the differentiation of the boundary integral and of the permeability tensor $\mathbf{K}$.

Linearization of determinant

The non-linear form is the integral of $J^{m,k}q$ and thus we have only to differentiate the determinant with respect to $\mathbf{u}$. The derivative with respect to π is null. The Gâteaux differential of $J^{m,k}q$ reads

$$\mathfrak{D}_{\mathbf{u}}\left(\int_{\mathscr{B}_r} J^{m,k}q\mathrm{d}\mathbf{X}\right) = b^{k,m}(\mathbf{h},q)[\mathbf{h}].$$

Linearization of non-linear Poisson term

The linearization of the non-linear Laplacian invokes both derivatives in $\mathbf{u}$ and π. The derivative with respect to $\mathbf{u}$ is

$$\mathfrak{D}_{\mathbf{u}}\left(\int_{\mathscr{B}} \mathbf{K}_r^m \nabla\pi^m \cdot \nabla q\mathrm{d}\mathbf{X}\right)[\mathbf{h}] = c_1^{m,k}(\mathbf{h},q) + c_2^{m,k}(\mathbf{h},q) + c_3^{m,k}(\mathbf{h},q),$$

where

$$c_1^{m,k}(\mathbf{h},q) = \int_{\mathscr{B}_r} J^{m,k}(\mathbf{F}^{m,k})^{-T} : \mathbf{H}((\mathbf{F}^{m,k})^{-1}\mathbf{K}(\mathbf{F}^{m,k})^{-T}\nabla\pi^{m,k}) \cdot \nabla q\mathrm{d}\mathbf{X},$$

$$c_2^{m,k}(\mathbf{h},q) = \int_{\mathscr{B}_r} J^{m,k}((\mathbf{F}^{m,k})^{-1}\mathbf{H}(\mathbf{F}^{m,k})^{-1}\mathbf{K}(\mathbf{F}^{m,k})^{-T}\nabla\pi^{m,k}) \cdot \nabla q\mathrm{d}\mathbf{X},$$

and

$$c_3^{m,k}(\mathbf{h},q) = \int_{\mathscr{B}_r} J^{m,k}((\mathbf{F}^{m,k})^{-1}\mathbf{K}(\mathbf{F}^{m,k})^{-T}\mathbf{H}^T(\mathbf{F}^{m,k})^{-T}\nabla\pi^{m,k}) \cdot \nabla q\mathrm{d}\mathbf{X}.$$

Since d_{NL} is linear in the pressure, its differential coincides with itself and hence

$$\mathfrak{D}_{\pi}\left(\int_{\mathscr{B}} \mathbf{K}_r^m \nabla\pi^m \cdot \nabla q\mathrm{d}\mathbf{X}\right)(\pi^{m,k},q;\mathbf{u}^{m,k})[\theta] = d_{NL}(\theta,q;\mathbf{u}^{m,k}) =: d^{m,k}(\theta,q).$$

5.4.3 Linearized problem

After the computation of the Gâteaux differentials of equations (5.17) and (5.18), we now write the weak problem that has to be solved at each step k of Newton's method and at each time step m.

Find $\mathbf{h}^{m,k} \in V$ and $\theta^{m,k} \in Q$ such that:

$$\begin{cases} a^{m,k-1}(\mathbf{h}^{m,k}, \mathbf{v}) & -b^{m,k-1}(\theta^{m,k}, \mathbf{v}) & = & -\mathfrak{F}_{\boldsymbol{u}}(\boldsymbol{u}^{m-1}, \pi^{m-1}, \mathbf{v}) = & \forall \mathbf{v} \in V \\ -c^{m,k-1}(\mathbf{h}^{m}, q) & -\tau d^{m,k-1}(\theta^{m,k-1}, q) & = & -\mathfrak{F}_{\pi}(\boldsymbol{u}^{m-1}, \pi^{m-1}, q) = 0 & \forall q \in Q. \end{cases} \tag{5.23}$$

where

$$a^{m,k-1}(\mathbf{h}, \mathbf{v}) = \sum_{i=1}^{5} a^{m,k-1}(\mathbf{h}, \mathbf{v}),$$

and

$$c^{m,k-1}(\mathbf{h}, q) = b^{m,k-1}(\mathbf{h}, q) + \tau \sum_{i=1}^{3} c_i^{m,k-1}(\mathbf{h}, q).$$

After the solution of this two-field problem, the new displacement and pressure can be computed by

$$\mathbf{u}^k = \mathbf{u}^{k-1} + \mathbf{h}^k$$

and

$$\pi^k = \pi^{k-1} + \theta^k.$$

The problem (5.23) is a generalized SP problem according to the definition of [BGL05]. In fact, the weak form a is in general not positive definite and the global problem is not symmetric, i.e $b \neq c$. Hence, it cannot be written as an inf-sup problem but the formal setting is the one presented in Sec. 3.5.3. Existence may be proved by means of Proposition 4 and Proposition 5 of Sec. 3.5.3. The weak forms a, b, c, and d are all continuous. The kernel of b and c is empty, hence a is coercive on the kernel. The weak form d is positive definite and symmetric. The weak form a is symmetric but it is not in general positive. Indeed, this property depends on the used material law and on the pressure distribution at the previous Newton iteration. Hence, even the use of convex material laws does not ensure positiveness of a. Thus, existence and uniqueness strictly depends on the pressure.

5.5 Finite element formulation

In order to write the finite element formulation of the problem, we introduce a conforming shape-regular mesh $\mathcal{T}_h$ consisting of N_h non-overlapping elements $\{K_i\}_{i=1}^{N_h}$. Moreover, we define the following finite dimensional spaces

$$\begin{aligned} V_h^l &= \left\{ \mathbf{v}_h \in V : \ \mathbf{v}_h|_{K_i} \in (\mathbb{P}_l)^3, \text{ for } i = 1, \dots N_h \right\}, \\ Q_h^\sigma &= \left\{ \pi_h \in Q : \ \pi_h|_{K_i} \in \mathbb{P}_\sigma, \text{ for } i = 1, \dots N_h \right\}, \end{aligned} \tag{5.24}$$

where $(\mathbb{P}_l)^3$ and $\mathbb{P}_\sigma$ are spaces of polynomials of degree l and σ, respectively. In particular, the notation $(\mathbb{P}_l)^3$ means that all components of the three-dimensional vector $\boldsymbol{v}_h|_{K_i}$ are polynomials of degree l. Denoting the Lagrangian basis functions of V_h^l and Q_h^σ by $\{\boldsymbol{\varphi}_q\}_{q=1}^M$ and $\{\phi_s\}_{s=1}^N$, with $M = \dim(V_h^l)$ and $N = \dim(Q_h^\sigma)$, respectively, the discrete solutions can be written as

$$\mathbf{h}_h^{m,k} = \sum_{q=1}^{M} h_q^{m,k} \boldsymbol{\varphi}_q \,, \qquad \theta_h^{m,k} = \sum_{s=1}^{N} \theta_s^{m,k} \phi_s \,. \tag{5.25}$$

Employing these definitions, the finite element formulation of the problem reads

Find $\boldsymbol{h}_h^{m,k} \in V_h^l$ and $\theta_h^{m,k} \in Q_h^\sigma$, such that:

$$\left\{ \begin{array}{rcrcl} a(\boldsymbol{h}_h^{m,k}, \boldsymbol{\varphi}_p) & - & b(\boldsymbol{\varphi}_p, \theta_h^{m,k}) & = & f(\boldsymbol{\varphi}_p) \quad , \quad p = 1 \dots M, \\ -c(\boldsymbol{h}_h^{m,k}, \phi_r) & - & d(\theta_h^{m,k}, \phi_r) & = & g(\phi_r) \quad , \quad r = 1 \dots N, \end{array} \right. \tag{5.26}$$

or in an algebraic form:

$$\begin{pmatrix} \mathbf{A} & -\mathbf{B}^T \\ -\mathbf{C} & -\mathbf{D} \end{pmatrix} \begin{pmatrix} \underline{h} \\ \underline{\theta} \end{pmatrix} = \begin{pmatrix} \underline{f} \\ \underline{g} \end{pmatrix}. \tag{5.27}$$

The matrices $\mathbf{A} \in \mathbb{R}^{M \times M}$, $\mathbf{B}$, $\mathbf{C} \in \mathbb{R}^{M \times N}$, and $\mathbf{D} \in \mathbb{R}^{N \times N}$ are the algebraic representations of the weak forms $a(\cdot,\cdot)$, $b(\cdot,\cdot)$, $c(\cdot,\cdot)$, and $d(\cdot,\cdot)$, respectively, and their entries are given by

$$\begin{aligned} \mathbf{A}_{ij} &= a(\boldsymbol{\varphi}_j, \boldsymbol{\varphi}_i) & \qquad \mathbf{B}_{lj} &= b(\boldsymbol{\varphi}_l, \phi_j) \\ \mathbf{C}_{lj} &= c(\boldsymbol{\varphi}_j, \phi_l) & \qquad \mathbf{D}_{ln} &= d(\phi_n, \phi_l). \end{aligned} \tag{5.28}$$

The entries of vectors $\underline{f}$ and $\underline{g}$ are given by

$$\underline{f}_i = f(\boldsymbol{\varphi}_i), \qquad \underline{g}_l = g(\phi_l), \tag{5.29}$$

while the vectors $\underline{h}$ and $\underline{\theta}$ are defined such that $\underline{h}_q = h_q$ and $\underline{\theta}_s = \theta_s$.

Analysis of the problem (5.27) is the same as for (5.23). Hence, the solution is unique only if the matrix **A** is positive semi-definite. In this case it is possible to prove stability of the solution and convergence of the discrete solution $(\mathbf{h}_h^{m,k}, \pi_h^{m,k})$ to the continuous one $(\mathbf{h}^m, k, \pi^m, k)$. Our choice of $\mathbb{P}_1 - \mathbb{P}_1$ elements is not compliant with the LBB condition. This results in matrices **B** and **C** with a non-trivial kernel. Still, since **D** is non-zero, spurious pressure modes are avoided. This effect of the Laplacian is the same as obtained by the Brezzi-Pitktäranta stabilization for Stokes and Navier-Stokes problems [BJ84].

Let us point out that in the case of very small permeability the stabilizing effect of the diffusion matrix **D** becomes less effective, which can be the source of numerical instabilities. In this case, the choice of discrete spaces that satisfy the LBB condition may reduce wiggles. Otherwise, stabilization like the FPL presented in Chap. 4 has to be employed.

In our simulations, no occurrence of volume locking was observed. However, despite the formulation of the problem in a saddle-point form, volume locking can occur in the case of large bulk moduli. To overcome this issue, different variational formulations involving a mixed problem [SR90, Bra09] or second order finite elements can be applied.

5.5.1 Spatial discretization of the inhomogeneities

Since we choose FE spaces of linear order both for **u** and π, the gradient **F** is constant on the element. For each element it is belongs to $\mathbb{R}^{3\times 3}$. Hence, we choose the tensor $\mathbf{F}_a$ in the same space. Evolution of $\mathbf{F}_a$ is governed by Eq. (5.15) on each element of $\mathscr{T}_h$

5.6 Solution of a benchmark problem

We apply the model presented in Sec. 2.3.5 to describe a confined compression test under given loading conditions. We consider the case in which the biphasic material (a specimen of soft biological tissue) is positioned inside a rigid cylinder and left free to grow. The cylindrical sample is then compressed between two plates: an external compressive force is applied at the upper plate, parallel to the symmetry axis of the specimen. The cylindric wall and the lower plate are impermeable, whereas the upper plate allows for fluid exudation, so that the liquid embedded in the material can escape from the specimen due to compression. The experimental apparatus is schematically shown in Fig. 5.1.

The formulation of the confined compression is based on the assumption that the matrix representation of the deformation gradient is given in a very simple diagonal form. Indeed, since the cylindrical walls of the parallel-plate-apparatus are supposed to be rigid and impervious, it is reasonable to assume deformations and velocities of all constituents to be along the Z-axis.

Therefore, using a cylindrical coordinate system, material point can be written as $\boldsymbol{X} = (R, \Theta, Z)$. Since we consider deformations generated by an uniaxial force applied along the Z-axis, the problem can be considered a monodimensional one with the following unknowns

$$\chi^z(t, \mathbf{X}) = \chi^z(t, Z) = z, \tag{5.30}$$

and

$$\pi^z(t, \mathbf{X}) = \pi^z(t, Z). \tag{5.31}$$

The deformation gradient is hence given by

$$\mathsf{F} = \operatorname{diag}\left\{1, 1, \partial_Z \chi^z\right\}. \tag{5.32}$$

We remark that, due to the particular form of F, the identity $J = \partial_Z \chi^z$ holds. Then, let consider $\mathbf{F}_a$, i.e. the tensor of anelastic deformation that maps the tangent space of the reference configuration, $T\mathscr{B}_r$, onto the tangent space of the natural configuration, $T\mathscr{B}_n$. We assume that $\mathbf{F}_a$ has the diagonal form

$$\mathsf{F}_a = \operatorname{diag}\left\{g_1, g_1, g_3\right\}. \tag{5.33}$$

This case of non-spherical growth is used to study the influence of a non-isotropic growth on the distribution of load and pressure throughout compression. From (5.32) and (5.33), it follows that

$$\mathsf{F}_e = \operatorname{diag}\left\{\frac{1}{g_1}, \frac{1}{g_1}, \frac{J}{g_3}\right\}. \tag{5.34}$$

System (2.85) in cylindrical coordinates can be found in [GGF+12] and [Giv03].

5.6.1 Physical parameters

We restrict our analysis to the case in which the external force increases linearly in time until $t_{max} = 30$ s when the maximum force, $F_{max} = -0.2 \cdot 9.81$ N, is reached.

In (5.33), we consider

$$\mathsf{F}_a = \operatorname{diag}\left\{g_1, g_1, g_3\right\} = \operatorname{diag}\left\{g, g, g + \epsilon\right\}, \tag{5.35}$$

where ϵ measures the deviation of $\mathbf{F}_a$ from a spherical anelastic deformation. We recall that, for the considered problem, the only possible value of g is unity. This implies that the deviations of $\mathbf{F}_a$ from a referential spherical tensor are actually the deviations from the identity tensor.

All the parameters are listed in Tab. 5.1.

The value of P^{zZ} at the upper boundary is given by $P^{zZ}(t, L) = \left(F_{appl}(t, L)\right)/S$, where $F_{appl}(t, L) = F_{max}[t/t_{max}]$. The area of the surface over which the applied load is distributed, $S = \pi R^2$, coincides with the cross section of the specimen in the reference configuration.

Parameter	Description	Value
L	Height of the specimen	10 mm
$2R$	Diameter of the specimen	5 mm
F_{max}	Maximum applied force	$-0.2 \cdot 9.81$ N
t_{max}	Time of load application	30 s
k_0	Hydraulic conductivity	$3.6454 \cdot 10^{-12}$ m^4/(N $\cdot$ s)
m_0	Material parameter	0.0848
m_1	Material parameter	4.638
ϕ_{s0}	Referential value of solidity	0.6
ϕ_{sn}	Solidity in the relaxed configuration	0.6
λ_n	Lamé's first modulus	0.3137 MPa
μ_n	Shear modulus	0.3566 MPa

Table 5.1. Parameters of the benchmark problem

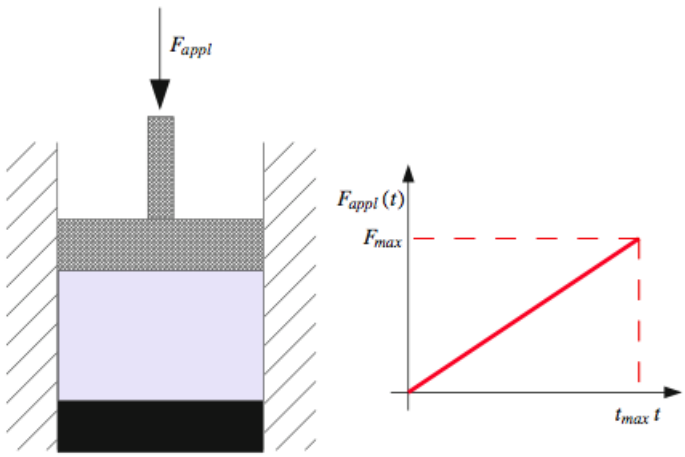


Figure 5.1. Schematic representation of the parallel plate apparatus used in the benchmark problem: the lower plate is impermeable, whereas liquid can flow through the upper plate. The applied load is linearly increasing in time.

5.6.2 Boundary and initial conditions

In order to solve (2.85) in the one-dimensional case, we have to supply boundary conditions (BCs) and an initial condition (IC). In particular, the mass balance (2.85.2) requires two BCs and one IC, whereas (2.85.1) requires one BC only. Boundary conditions have to be provided at the boundary faces at $Z = 0$ and $Z = L$, which identify the lower and upper boundary of the specimen, respectively.

The boundary conditions have to be consistent with the following requirements:

- the axial stress at the upper boundary of the specimen has to be equal to the applied load, $P_{appl}(t)$;
- the velocity of the fluid and of solid phase has to be zero at the bottom because the lower plate is impermeable and fixed;
- the pressure π has to be zero at $Z = L$ since the liquid is in equilibrium with the atmosphere.

These observations are translated in the following set of boundary conditions

$$\chi^z(t,0) = 0, \tag{5.36}$$

$$\frac{\partial \pi}{\partial Z}(t,0) = 0, \tag{5.37}$$

$$\mathbf{N}^T\mathbf{P}\mathbf{N} = P^{zZ} = P_{appl}(t), \tag{5.38}$$

$$\pi(t,L) = 0. \tag{5.39}$$

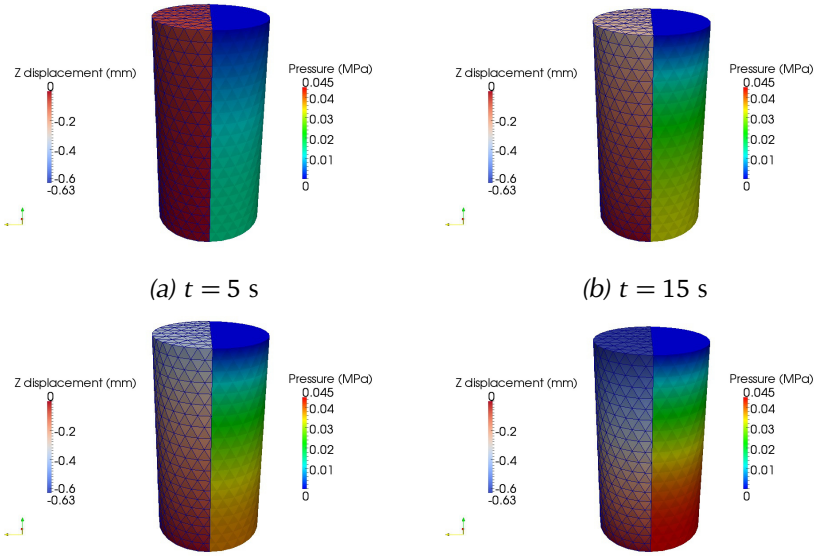
As initial condition, we take $\mathbf{u} = (0, 0, g_3 Z)$. Indeed, at the initial time, there is no elastic deformation, although the anelastic deformation has already occurred. We remark that, for consistency, the condition $\mathbf{F}_e(\mathbf{X}, 0) = \mathbf{I}$ entails that $g_1 = 1$.

5.6.3 Results

The discretization method described above has been implemented in the software toolbox UG/Obslib++ [BBJ+97, GK08]. For the validation of our discretization strategy we consider a grid a small number ($10k$) of degrees of freedom. The solution of the linear system (5.27) is done with UMFPACK [Dav04]. For Newton's method we used a tolerance of 10^{-8}.

The solution procedure described in Sec. 5.3 and 5.4 has been validated in [GGF+12]. Because of the cylindrical symmetry of this setting, we could compare our results to those obtained for the one-dimensional case with a different solution strategy.

We ran a set of simulations with $\epsilon = \{0, 0.01, 0.1, 0.2\}$ in order to highlight the influence of this parameter on the response of the material (e.g., vertical displacement and deformation inside the specimen).



(a) $t = 5$ s *(b)* $t = 15$ s

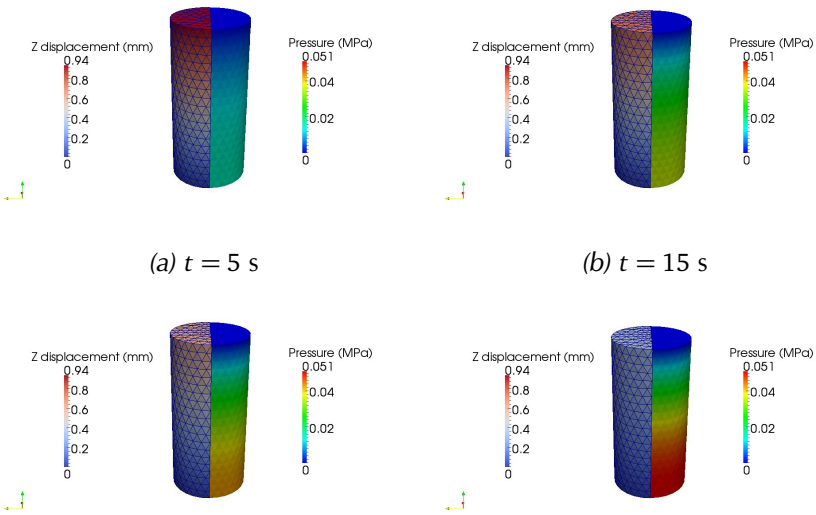
(c) $t = 20$ s *(d)* $t = 30$ s

Figure 5.2. Time evolution of the displacement and the pressure without growth ($\epsilon = 0$).

Results of the three-dimensional simulations for $\epsilon = 0.0$ and $\epsilon = 0.1$ are reported in Figs. 5.2 and 5.3, respectively.

In Fig. 5.4 we report the evolution of vertical component of the first Piola-Kirchhoff tensor and vertical relative displacement, $[z(t,Z) - z(0,Z)]/L_{\text{in}}$, for $\epsilon = 0.1$. Here, L_{in} denotes the length of the specimen at time $t = 0$s, which is defined by $L_{\text{in}} := \int_0^L \partial_Z \chi^z(0,Z)\mathrm{d}Z = J_a L$. The amplitude of the displacement increases in time with the applied load (cf. Fig. 5.4b). This behaviour is qualitatively the same also for the other considered values of ϵ.

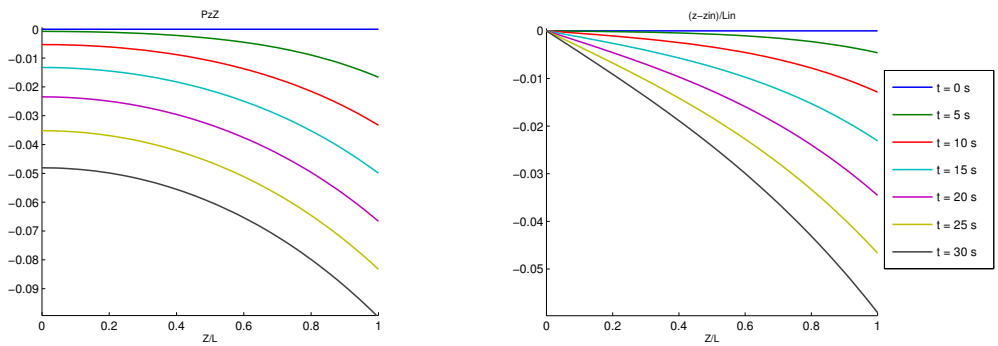
Pressure and vertical relative displacement for different ϵ are plotted in Fig. 5.5 at time $t = t_{max}$. The value of the pressure at the lower boundary rises as ϵ increases, and the pressure distribution tends to become more inhomogeneous for larger deviations of $\mathbf{F}_a$ from sphericity. For the considered load, the normalized final displacement, $\left[z(t_{max},Z) - z(0,Z)\right]/L_{\text{in}}$, which is zero at the bottom of the specimen, diminishes with increasing ϵ.



(a) $t = 5$ s *(b)* $t = 15$ s

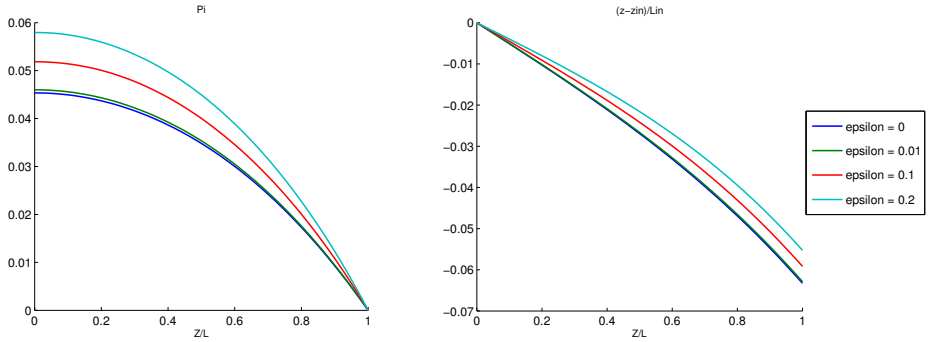
(c) $t = 20$ s *(d)* $t = 30$ s

Figure 5.3. Time evolution of the displacement and the pressure with $\epsilon = 0.1$.



(a) *(b)*

Figure 5.4. Evolution in time of the distribution of P^{zZ} (a) and relative displacement (b). Results are plotted every 5 s up to 30 s, with $\epsilon = 0.1$.



(a) *(b)*

Figure 5.5. Distribution of the pressure over space (a) and relative displacement (b) at $t_{max} = 30$ s, for different values of ϵ.

5.7 Multigrid method for the poroelasticity system

Multigrid (MG) methods allow for an efficient solution of linear systems arising from the discretization of PDEs. The idea of MG is based on two observations:

- although classical iterative solvers have poor convergence properties, they can smooth in few iterations the highest frequencies representable on a mesh; because of this reason, iterative solvers are called *smoothers*.
- the smooth component of the error can be approximated on a coarse mesh.

For this reason, a hierarchy of grids is constructed and on each level a different method is used to smooth the error.

We explain the idea for a two-grid algorithm. It is realized by repeatedly applying the following steps.

1. The correction is smoothed applying few steps (3 to 5) of a Jacobi or Gauss-Seidel solver.
2. The new residual is computed and it is transferred to the coarse level. It is done applying a *resctriction* operator.
3. On the coarse level the correction is computed exactly or “almost-exactly”.
4. The correction is *interpolated* to the fine grid using the *prolongation* operator.
5. Again, some steps of a Jacobi or Gauss-Seidel solver are applied to smooth the error.

The MG cycle is realized applying a two-grid algorithm for the solution of point 3. It gives rise to the so-called *V-cycle* MG. Hence, the ingredients of a MG solver are the smoothers, restriction and prolongation operators, and the coarse grid operators.

MG methods have been initially developed for scalar elliptic problems. For this class of problems, it was shown that MG methods have linear complexity with respect the number of unknowns. The straightforward application of MG to the generalized SP problem (5.27) or to SP problems in general is not usually so efficient. In particular, when the coupling between the unknowns is not strong the smoothing is not so effective.

A possible remedy is to use the so-called *point-block* smoothers. The idea of these smoothers is to consider the stiffness matrix as a block matrix where each block contains the degrees of freedom of the different variables associated to the same node. Then, the smoothing steps are applied to this matrix in block form. This can be considered the simplest remedy to the low smoothing factors of standard iterative solvers. This strategy is the one we will use in our numerical experiments but in [TOS01] it is pointed out how the obtained performances are not always satisfactory.

An extension of these smoothers are the box relaxation smoothers. They consist in the subdivision of the the mesh into subdomains and the linear system of each subdomain is exactly solved. A different class is composed by the *distributive* smoothers. These smoothers transform the system by pre-conditioning or post-conditioning the global matrix of (5.27). In this way, two decoupled subsystems are obtained and point-wise smoothing can be done separately on each of them. The application of these smoothers requires a good approximation of the pressure Schur complement of the system. Two possible preconditioners are presented in [LWW02, FKJ12].

Finally, we want to point out that we are using the MG solver inside a non-linear solution strategy. We are in the framework of the so-called *Multigrid Newton Method* (MGNM). An alternative would be using the non-linear multigrid method. This would require the development of non-linear smoothing strategies for the poroelasticity system. The MGNM does not require the exact solution of each step of the Newton's method but few MG cycles are enough [Bra09]. This implies that quadratic convergence of Newton's method is lost. Another ingredient that has to be specified for MGNM is how to assemble the operators on the coarser levels. Indeed the coarse operators depend on the current value of displacement and pressure. In our solution strategy, we project them on the coarser levels and then we assemble the operator on every level. An alternative is to use the Galerkin assembling strategy.

5.7.1 Efficiency of the solver

The MGNM has been implemented using the MG solver of the toolbox UG/obslib++. In this test, we do not fix the number of MG cycles but we solve each Newton iteration with a tolerance of 10^{-12}. For the coarse level solution, we have implemented an interface to the library PETSc. We tested two different coarse solvers: the parallel direct solver MUMPS and the iterative solver GMRES. For this latter we used a tolerance of 10^{-12}.

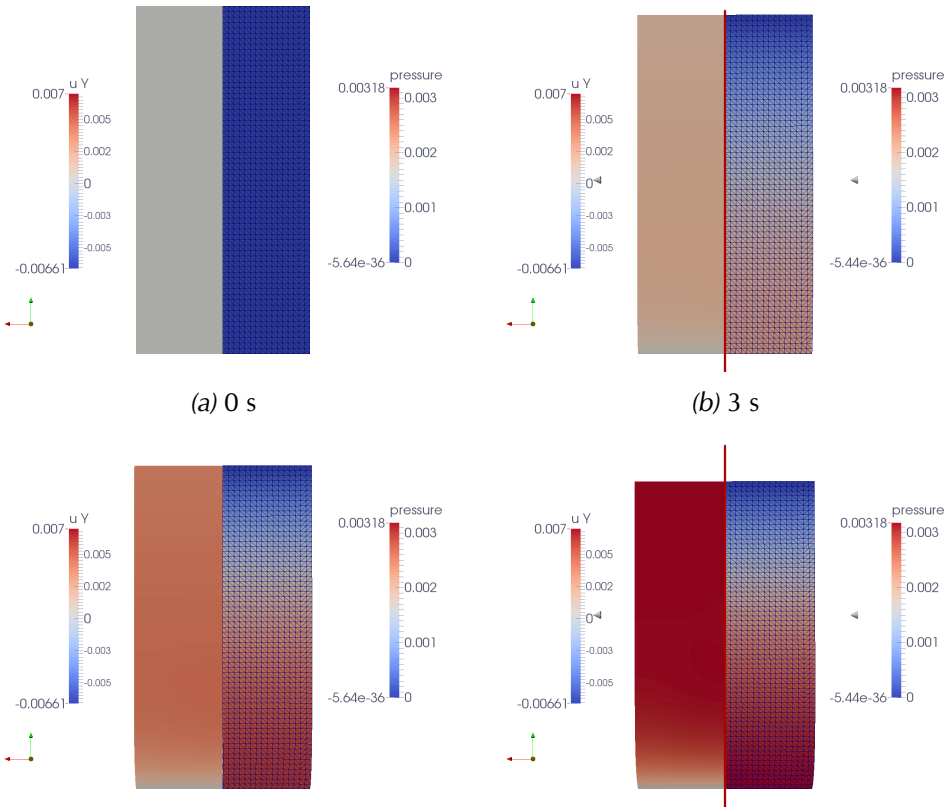
As smoother we used 5 iterations of a point-block Gauss-Seidel. Since we are testing our solver in a parallel framework, the Gauss-Seidel smoother is applied only locally on each subdomain.

We consider an unconfined test of a brick of size $1 \times 1 \times 2$. The unconfined compression test cannot be reduced to a one-dimensional case. The physical parameters are the same as presented in Tab.5.1. The force is applied on an area of 1×1. Results of this test are reported in Fig. 6.9.

The original geometry has been divided into 8192 cubic elements and then each of them has been divided into six tetrahedra. Then we apply up to 5 uniform refinements of the original grid. Hence, we test our MG implementation from 2 up to 6 levels. The number of processors and nodes is reported in Tab. 5.2.

In order to consider a non-linear case, we compute the convergence rate and the scaling at the second Newton iteration of the second time step.

The convergence rates with GMRES and MUMPS as coarse solvers are reported in Figs. 5.7a and 5.8a, respectively. In both cases the convergence rate drops with the



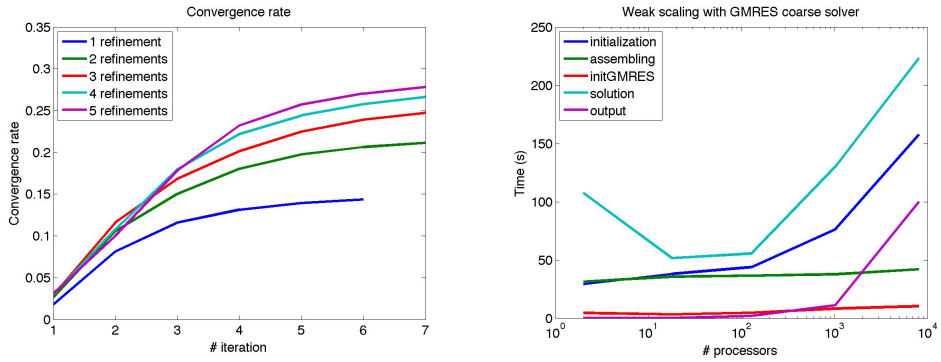
(a) 0 s *(b)* 3 s

(c) 8 s *(d)* 13 s

Figure 5.6. Evolution of unconfined compression test on a brick. Mesh with 70k nodes.

refinements	0	1	2	3	4	5
mesh nodes	9k	70k	545k	4.2M	33.8M	269.7M
# cores	1	2	16	128	1024	8192

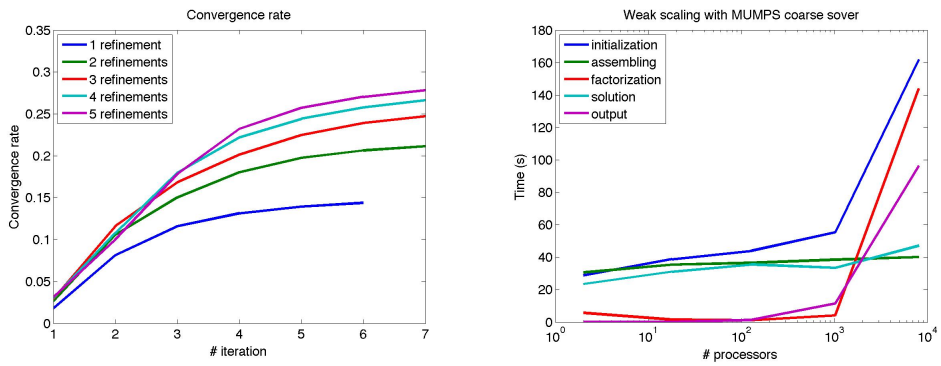
Table 5.2. Number of nodes for each mesh of the ideal scaling test.



(a) *(b)*

Figure 5.7. Convergence rate and time with a GMRES solver.

iterations but it reaches an asymptotic limit. The larger the number of levels is, the larger the convergence rate is.



(a) *(b)*

Figure 5.8. Convergence rate and time with a MUMPS solver.

The scaling results of the components of the solution process are reported in Fig. 5.7b for GMRES and Fig. 5.8b for MUMPS. The assembling time is the same and it scales with the number of processors. The solution time with GMRES increases with the number of processors. The solution time with MUMPS is constant but it requires the factorization of the coarse level operator. The time of this phase is reported in the red line of Fig. 5.8b. This can be seen as a strong scaling test for MUMPS since the coarse grid is the same for

all the tests. The factorization time is constant up to 1024 processors.

Chapter 6

Efficient solution and validation of a coupled elastic-poroelastic model for a tooth-PDL system

In this chapter we present an efficient solution method to simulate a tooth-periodontal ligament system. The proposed solution strategy is based on a MG solver. It has to be adapted for the tooth-PDL system because of the SP nature of the problem and because of the large jump in the coefficients between the two materials. Finally, we validate our model against physical simulation realized on a porcine tooth.

6.1 Introduction

The biomechanics of tooth and the surrounding structures is a fundamental and challenging task in dentistry.

The main goal of dental biomechanics is to study the tooth displacement under physiological loadings (chewing, swallowing, and clenching) and under application of orthodontic treatments. Moreover, it is important to understand the mechanical parameters of tooth and their changes due to aging and to parodontal diseases, such as gingivitis and periodontitis. The former is a reversible inflammatory disease of the gingival tissue. The latter is responsible for the destruction of the supportive tissues of tooth and it might even lead to its loss.

Teeth can only virtually be thought as rigidly linked to *dental alveoli* in the jaw bone. Actually, they are connected to their socket by the *periodontium* (from Greek, peri-, meaning “around”, and -odons, meaning “tooth”). This is a group of specialized tissues that support the teeth and maintain them in the maxillary and mandibular bones.

The periodontium absorbs the mechanical stresses exerted on the tooth, for example generated by chewing, allowing micro-displacements. The forces acting on a tooth can also be large: for example, a tight bite can correspond to a force of 80 N.

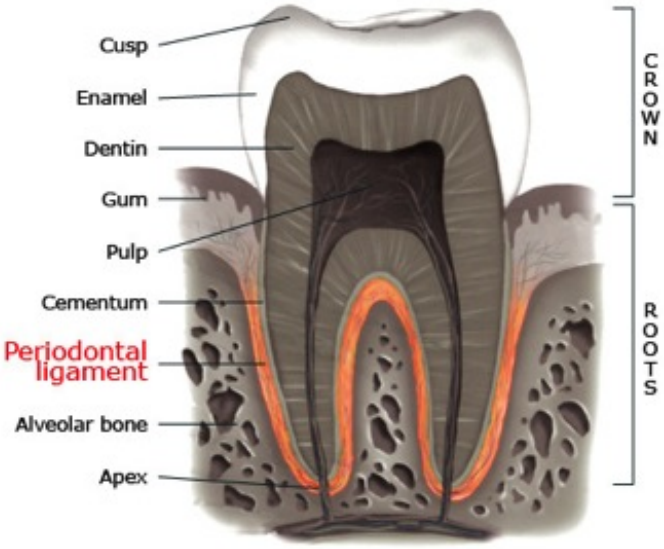


Figure 6.1. The different tissues of tooth and periodontium. Image taken from [stu].

The periodontium consists of four principle tissues: the *alveolar bone*, the *cementum*, the *periodontal ligament* (PDL), and the *gingiva*. They have been sketched in Fig. 6.1. Each of these components has its own location, tissue architecture, and bio-chemical composition.

- The alveolar bone, or *alveolar process*, is a dense ridge of bone that covers the dental alveoli.
- The cementum is a calcified substance that surrounds the surface of the root of the teeth.
- The PDL is a soft connective tissue that joins the cementum to the inner wall of the alveolar socket.
- The gingiva, or *gum*, is the most superficial tissue. It is a mucous membrane that surrounds the teeth and provides a seal to the periodontium.

Among the structures of periodontium, the PDL is the most important but also the most complicated, as it does not have only a tooth-supportive function. It will be described more in details in the following subsection.

6.1.1 The periodontal ligament

The PDL has a width going from 0.15 mm to 0.38 mm [BHM09]. It is larger at the coronal and apical ends of the alveolus, and thinner in the middle third of the root.

Moreover, the periodontal ligament is thicker for functioning teeth subjected to larger loadings than for teeth out of function or in light function. At the level of the neck of the tooth, the PDL has a circular shape. It creates the so-called *dent-gingival seal*. It protects the deeper tissues from the infiltration by saliva, liquids, food residuals, and, above all, bacteria.

The PDL has inherently a biphasic structure, composed by a solid phase and a fluid phase.

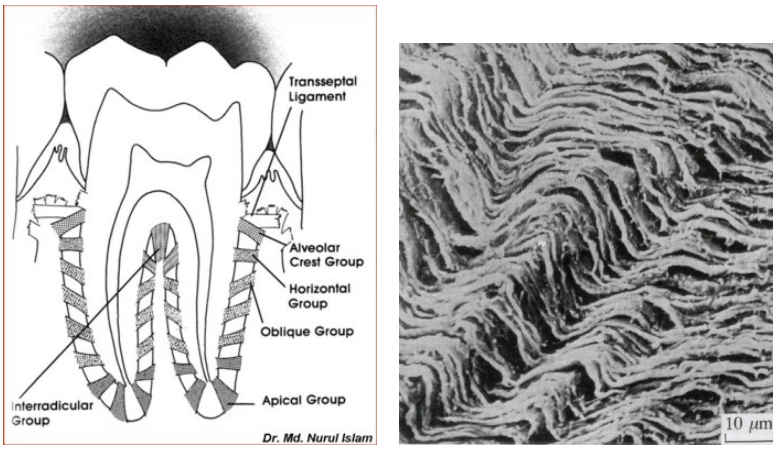
The solid phase is the largest one: it forms from 53% to 74% of the PDL depending on the zone. The fluid phase (or *interstitial fluid*) is composed by a 70% of water. There is also a blood component inside the vessels that run through the PDL.

The solid phase has a cellular component that, together with nervous fibers and blood vessels, is plunged in an extra-cellular matrix.

Extra-cellular component

The extra-cellular matrix is mainly formed by bundles of connective fibers (or elastic collagen fibers), that have a different orientation depending on the zone where they are located. The collagen fibers are divided into two principal subgroups: the fibers of the *alveodental ligament* and the *transseptal fibers*.

The fibers of the alveodental ligament originate from the alveolar process and from the cementum and they meet inside the PDL. The ones connected to the alveolar process are larger, while the others are smaller but they are several. These fibers, that are also called Sharpey fibers, are classified in *alveolar crest*, *horizontal*, *oblique*, *apical*, and



(a) *(b)*

Figure 6.2. Classification of the fibers inside the PDL of a multi-rooted tooth (a). Image taken from web page of Prof. Nurul Islam [Isl]. Micro-CT scan of the periodontal fibers (b).

interradicular fibers, with respect to the zone where they originate. The interradicular fibers are only present in multi-rooted teeth.

The transseptal fibers are located near the alveolar bone crest and they are connected to the cementum of teeth. Since they do not have an attachment to the alveolar bone, they are considered part of the gingival tissue. They form interdental ligaments that keep all the teeth aligned.

The role of the collagen fibers is to mediate the forces acting on the teeth. Intra-oral forces range from below 1 N to more than 100 N with time intervals between fractions of a second to hours or weeks [DKH+11]. The fibers withstand both compressive forces and traction forces. In case of traction forces, the pressure on the tooth elongates the bundles of the PDL and it is transmitted as a traction on the cementum and the bone.

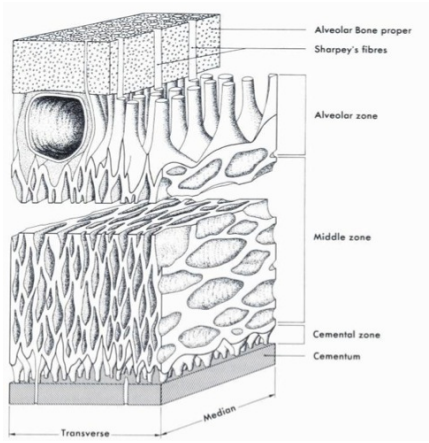


Figure 6.3. Sketch of the periodontal structure. Sharpey fibers at the two ends of the ligaments are connected to the cementum and the alveolar bone. Image from [BHM92].

Cellular component

The cellular component of the periodontal ligament consists of:

- fibroblasts, that are predominant as they are responsible for collagen production; they are disposed parallel to the fibers of the connective matrix;
- osteoblasts and cementoblasts, that form the alveolar bone and the cementum (in areas of tension);
- precursor cells that can differentiate into the specialized cells (fibroblasts, osteoblasts, and cementoblasts, respectively);

- osteoclasts and cementoclasts, that are present only during re-adsorbing phenomena (along lines of compression);
- macrophages;
- epithelial cell rests of Malassez.

These cellular components provide a high remodeling function of PDL. This self-renewal capability of PDL is necessary to ensure its physiological maintenance and repair against the effects of the rapidly changing physical forces due to chewing, speech and orthodontic movements.

Vascular component

The periodontal ligament is well-vascularized by blood vessels that reach it through the fundus of the alveoli. This vascular component provides a nutritive function for the various cells of PDL and also contribute to soften the pressure due to chewing.

Neural component

The periodontal ligament is richly supplied also with neural fibers that provide a sensory function. In particular, the nerve endings are primarily receptors for pain, pressure and contact.

6.1.2 State-of-the-art of the modeling of PDL

Due to its structure, the biomechanics of the PDL is highly elaborated to describe. The elastic nature of collagen fibers and the damping behaviour of the fluid phase are responsible for the non-linear and time-dependent response of the PDL.

PDL may be considered a biphasic material: the collagen fibers are the solid matrix, while the interstitial fluid, that is basically water, is the fluid phase. A further motivation for using a poroelastic model is the PDL force response to relaxation experiments. In fact, as we will show in the next Sec. 6.12d, the response to an instantaneous displacement is a rapid increase of force followed by a relaxation.

Even if this biphasic nature is well known, very few models investigated this aspect. The most classical studies consider the PDL as an isotropic and linear elastic material [TS83, WE84, AMPM91, SJ06]. In order to validate the model also for large displacements, a bilinear model has been tested [PBJ02, KBZ$^+$04, KBJ03]. The idea of these works is to introduce a Young modulus that can assume two different values depending on strain state.

In [NCP$^+$07, Pro00, QCK01], linear and non-linear anisotropic material laws were investigated. Anisotropy was included to take into account fiber directions. In [PCB$^+$02], the PDL was considered as an incompressible elastic material. In [TDLE02, Jus04], in

order to take into account the PDL time-dependent behaviour, a disco-elastic material law was considered. In these models the influence of the interstitial liquid is implicitly given by means of abstract material parameters which can be adjusted in order to improve the approximation strength of the numerical simulations to the outcome of the in-vitro measurements.

In [NPSS02], a poroelastic model has been presented assuming small displacements. The non-linear behaviour was included assuming a tangential constitutive tensor the depends on the current stress and strain. Moreover a non-linear permeability was assumed. In this work, a FE discretization was presented. Solution was done using a implicit scheme for time integration and the Newton-Raphson method for the solution of the non-linearities. Because of the small displacement assumptions, the model was validated only in small ranges of forces. The obtained results presented a good fitting in this range but the model was not validated on realistic geometries and no mesh dependence was studied.

A detailed work is the PhD thesis [Ber08]. Here, both in-vitro and in-silico experiments were realized on five specimina of bovine teeth. Thanks to the in-vitro experiments, the PDL was studied both under filled and drained conditions. The flux of fluid at the boundary was included.

In [FGD+11], a biphasic model was applied to study the force response to indentation experiments on realistic geometries of porcine tooth, obtained by micro-CT scan. Here, a small displacement model was studied, but non-linear permeability was included. A detailed coupling of the tooth-PDL models was presented, with FE discretization, time advancing algorithm and a predictor-corrector scheme to treat the non-linearity. A sensitivity analysis of permeability parameters was performed even with large displacements, while most of the previously published papers focused on a small loading regime.

Another discussed problem in the field on dentistry is the identification of the correct material parameters. Indeed, the values used in the different articles about the modeling were highly inconsistent both for the tooth and for the PDL. A detailed review of the different measured mechanical parameters for the tooth is reported in [KMM03]. It was observed that the range of the Young modulus for the dentine was 1-30 GPa. In elastic problems, it was usually higher than 20-25 GPa, while in the considered viscoelastic models it was about 12 GPa. The anisotropy of the measurements was also observed and the material was subject to fatigue.

Since the early work on 1935 [DS35], the two phases were not taken into account and the PDL was considered as an incompressible material. This hypothesis was assumed also in [RJ97], where usually the Poisson ratio was taken as 0.49. The same work collects also a large range of parameters observed for the Young modulus of PDL. Measured values go from 0.07 MPa to 1750 MPa. In [Ber08], the properties of the drained solid phase were experimentally measured. Here, in contrast with the assumptions of the previous works, the mechanical response of the PDL results as the one of a

porous, compressible, and hyperelastic material.

Since the biphasic model was not highly investigated in the literature, almost no value of the permeability under loading is available. In [Ber08], from physical measurements on bovine specimina the value of 50 $mm^4 \cdot N^{-1} \cdot s^{-1}$ has been found. In [NPSS02], a range of values from 50 to 200 $mm^4 \cdot N^{-1} \cdot s^{-1}$ was identified.

An important effect that has been studied in [PCB+02] is the sensitivity of the results with respect to the mesh resolution. It was showed that coarse meshes overestimate the force response but a convergence of the force response was observed for $h \to 0$.

6.1.3 Goal of this work

The work presented in this chapter is part of a larger project, namely the Clinical Research Unit KFO208 of University of Bonn titled "Aetiology and Sequelae of Periodontal Diseases. Genetic, Cell Biological and Biomechanical Aspects", partially supported by German Research Foundation (DFG). This project proposed an interdisciplinary research on periodontal diseases, since they are showing a significant increase in population. The major goal of the Research Unit was to improve the prevention, diagnostics and therapy of periodontitis, through a better understanding of its causes and effects as well as of regenerative processes in the involved tissues.

In particular, the biomechanical aspects of the periodontal ligament have been investigated in collaboration between the work group of Prof. Bourauel (Universität Bonn) and the work group of Prof. Krause (Università della Svizzera italiana). These groups worked together on the "Project #5", titled "Development of Mathematical Models and Efficient Algorithms to Simulate the Loading Conditions of the Periodontium in Dental Biomechanics". Specifically, the aims of our research group were:

- the development of a closed mathematical model of the biphasic and time-dependent constitutive law of PDL;
- the development of robust and efficient numeric solution methods for the arising non-linear systems;
- the application of these methods to the PDL;
- the validation of the developed models with experimental biomechanics results and clinical data, collected by the research group of prof. Bourauel.

6.2 Setup of the in-vitro experiments

In [DKH+11], our co-workers presented a novel intra-oral device to determine the biomechanical characteristics of the periodontal ligament. The measurement set-up is reported in Fig. 6.4. It allows to determine the force-deflection characteristics of a tooth both in-vivo and in-vitro. A piezoelectric actuator is used to apply displacements

on the tooth's crown. The point of displacement application is chosen in the centre of the labial side of the tooth's crown. Resulting forces are simultaneously recorded by an ultra miniature force sensor. Two cylindrical magnets are attached on the top of the tooth's crown. The movement of these magnets is detected using an array of eight Hall effect sensors for each magnet, and the movement of the loaded tooth is derived from the movement of the magnets.

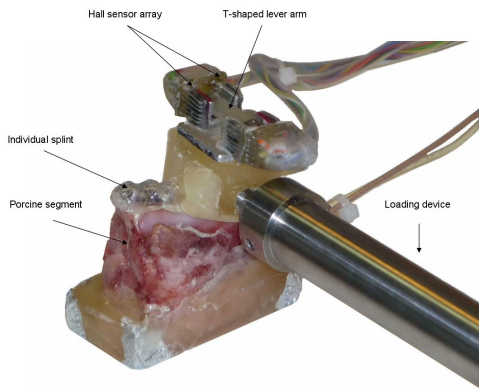


Figure 6.4. Measurement set-up used for the in-vitro measurements.

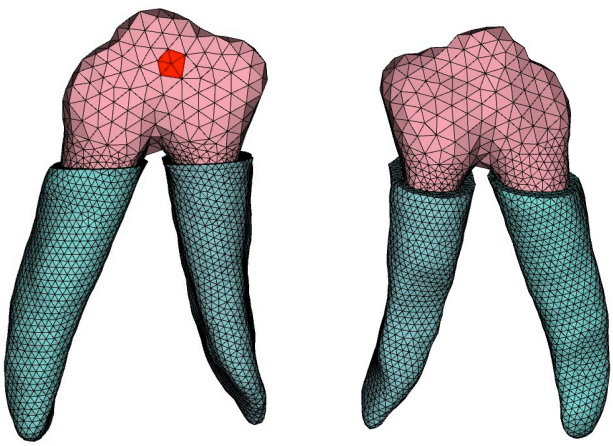
The tooth employed for the calibration of the device and for the validation of the numerical experiments is a premolar from a pig jaw bone segment. It was used for this investigation as the omnivore diet of pigs as well as the size (here, 21.5 mm height) of the pig's two-rooted premolars allow an easy transfer of the results to the human masticatory system. After experimentation the jaw bone segment was scanned in an in-house μ-CT scanner (SkyScan 1174, Belgium) and a FE mesh of the relevant geometries (tooth and PDL) was created using the software ADOR3D [RKB+05]. Figure 6.5 shows the resulting mesh which was used to validate our biphasic material model of the PDL by comparing the numerical results with the measured force-deflection behaviour of the same specimen, as described in the next sections.

6.2.1 Physical experiments

Two physical experiments are considered in this work:

- increasingly linear loading;
- relaxation experiments.

In both these experiments a maximum displacement of 0.2 mm is applied on the tooth's crown.



(a) (b)

Figure 6.5. Geometry used for the simulations of the tooth-PDL system: labial side (a) and buccal side (b). In the labial side, the contact area between the tooth and the bolt is highlighted. The contact surface is centred at 19 mm away from the root.

Linear loading

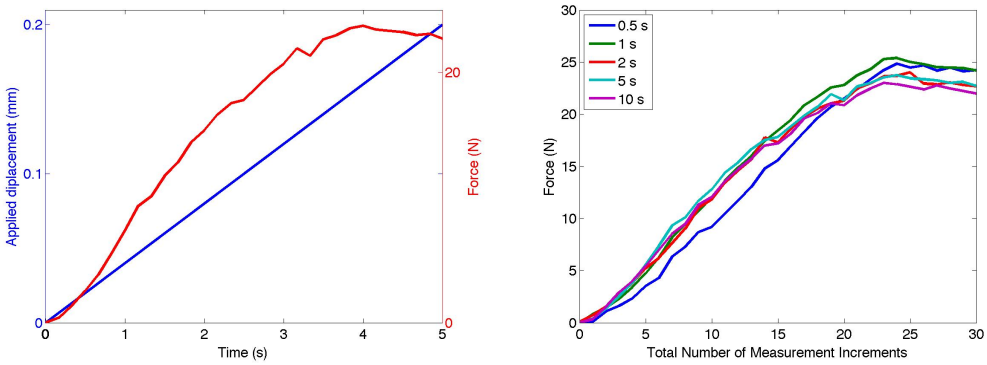
In this experiment a linear increasing in-time displacement is applied by the indenter with different loading times $T_{fin} = 0.5, 1.0, 2.0, 5.0$ s. The force response with different loading velocities is reported in Fig. 6.6a.

Fig. 6.6b shows the force response when the maximum displacement is reached in 5 s. The force response gives a non-linear curve divided into three phases: a stiffening at low loading (up to 0.05 mm), a softening behaviour (up to 0.15 mm), and a plateau for larger displacements. From the stiffening to the softening phase, a change of curvature is observed. In some specimen, the first convex response is more pronounced.

Relaxation experiments

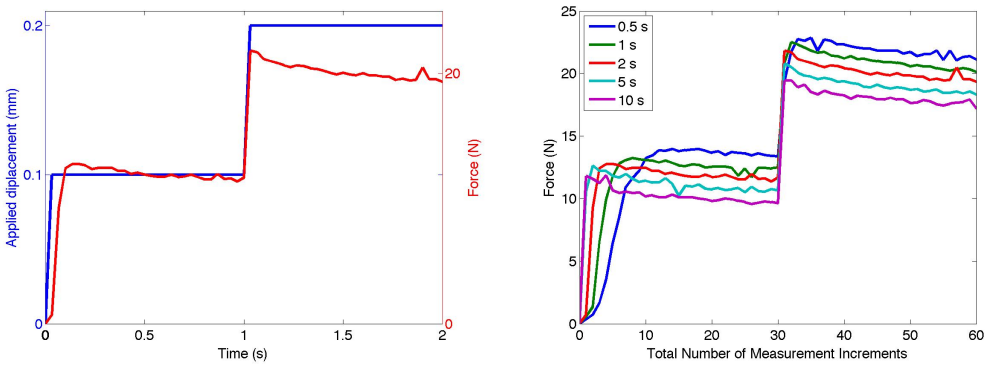
In the relaxation experiments, the maximum displacement is applied in two steps of 0.1 mm. This is realized for several loading times $T_{fin} = 0.5, 1.0, 2.0, 5.0$ s. After the first step the system reaches an equilibrium state. The load is kept for a time $T_{fin}/2$ and then the second displacement is applied.

In Fig. 6.7a, the result for an experiment with $T_{fin} = 2$ s is reported. This short holding time resulted in maximum forces up to 14 and 23 N in the peak point of the first and the second holding period, respectively.



(a) *(b)*

Figure 6.6. Applied displacement (a) and force response (b) for a linear loading experiment.



(a) *(b)*

Figure 6.7. Applied displacement (a) and force response (b) for a relaxation experiment.

6.3 Modeling

For the simulation of the in-vivo experiments realized by our co-workers, we create a coupled model. Since we are reproducing only short-time experiments, we can focus on the tooth and the PDL: the bone is assumed to be rigid, remodeling takes long time to occur and gums are not considered. Hence, we present a coupled elastic-poroelastic model for the simulation of the tooth-PDL system. In particular we discuss the coupling between these two models. For the tooth we use the elastic model presented in Sec. 1.1 and for the PDL the biphasic model presented in Sec. 1.2. Since we have to distinguish between quantities defined for the tooth and for the PDL, we will use the superscript $Tooth$ referring to the tooth and the superscript PDL referring to the PDL.

6.3.1 Kinematics and coupling conditions

To formally derive the coupled system, we introduce the body manifolds $\mathscr{B}^{Tooth}$, $\mathscr{B}_s^{PDL}$, and $\mathscr{B}_\ell^{PDL}$. The points of these manifolds are denoted by X^{Tooth}, X_s^{PDL}, and X_ℓ^{PDL}, respectively. $\mathbf{x}^{Tooth}$ and $\mathbf{x}^{PDL}$ indicate the points of the current configurations $\mathscr{B}_r^{Tooth}$ and $\mathscr{B}_r^{PDL}$. Their motions are denoted by

- $\mathbf{x}^{Tooth} = \boldsymbol{\chi}^{Tooth}(\mathsf{X}^{Tooth}, t)$,
- $\mathbf{x}^{PDL} = \boldsymbol{\chi}_s^{PDL}(\mathsf{X}_s^{PDL}, t)$,
- $\mathbf{x}^{PDL} = \boldsymbol{\chi}_\ell^{PDL}(\mathsf{X}_\ell^{PDL}, t)$.

Points of the two solid body manifolds are mapped into their reference configurations $\mathscr{B}_r^{Tooth}$ and $\mathscr{B}_r^{PDL}$ by two maps $\mathbf{X}^{Tooth} = \boldsymbol{\kappa}^{Tooth}(\mathsf{X}^{Tooth})$ and $\mathbf{X}_s^{PDL} = \boldsymbol{\kappa}_s^{PDL}(\mathsf{X}_s^{PDL})$. The motion of the solid phase can be parametrized with respect to these two reference configurations. From here on, we assume that the reference configurations of the two solid phases coincide with the initial configurations, i.e.

$$\mathbf{X}^{Tooth} = \boldsymbol{\chi}^{Tooth}(\mathbf{X}^{Tooth}, 0) \quad \text{and} \quad \mathbf{X}^{PDL} = \boldsymbol{\chi}_s^{PDL}(\mathbf{X}^{PDL}, 0). \tag{6.1}$$

The two displacements are denoted by $\mathbf{u}^t$ and $\mathbf{u}^{PDL}$. The displacements and the pressure π are the unknowns of the differential problem. For them we have to set initial and boundary conditions that are physically meaningful. The initial condition for the displacements is homogeneous because of (6.1).

The two current configurations represent the space occupied by the tooth and the PDL. The boundary of $\mathscr{B}_r^{Tooth}$ is hence divided into three subsets Γ_I, Γ_D^{Tooth}, and Γ_N^{Tooth}. The boundary of $\mathscr{B}_r^{PDL}$ is instead divided into three subsets Γ_I, Γ_D^{PDL}, and Γ_N^{PDL}. Γ_I represents the boundaries of the two manifolds where they are in contact. The following compatibility condition is imposed

$$\boldsymbol{u}^{Tooth} = \boldsymbol{u}^{PDL} \quad \text{on } \Gamma_I. \tag{6.2}$$

The subset Γ_D^{PDL} denotes the part of the boundary that is in contact with the bone. Here we assume that no displacement occurs in our experiments, and hence

$$\boldsymbol{u}^{PDL} = \mathbf{0} \quad \text{on } \Gamma_D^{PDL}.$$

In our model, the contact between the indenter and the tooth is described by means of time-dependent Dirichlet boundary condition. This condition is imposed only on the horizontal component u of the displacement. We consider two different kinds of conditions on Γ_D^{Tooth} corresponding to the linear loading and the relaxation experiments. The former is described by

$$u(\mathbf{X}, t) = \frac{t}{T_{fin}} u_{\max} \quad \mathbf{X} \in \Gamma_D^{Tooth}, \tag{6.3}$$

and the latter is described by

$$u(\mathbf{X}, t) = u_{\max} \quad \mathbf{X} \in \Gamma_D^{Tooth}. \tag{6.4}$$

The value of $u_{\max}$ depends on the considered experiment.

On the other two directions, homogeneous Neumann conditions are imposed. The same is done on Γ_N^{Tooth} and Γ_N^{DPL}.

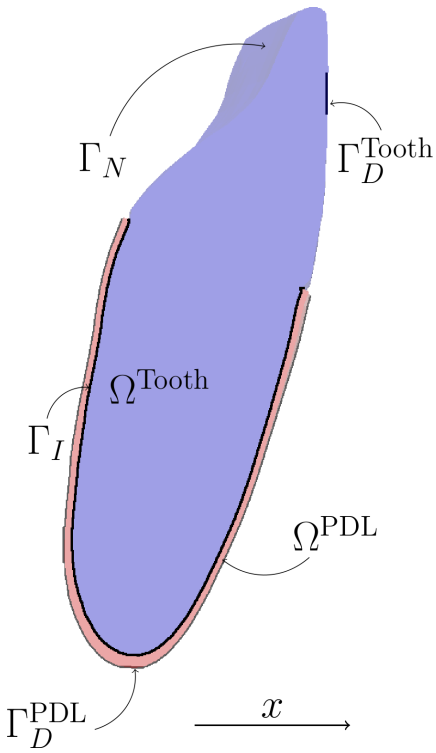


Figure 6.8. The different domains of tooth and periodontium.

The spatial velocities are denoted by $\mathbf{v}^{Tooth}$, $\mathbf{v}_s^{PDL}$ and $\mathbf{v}_\ell^{PDL}$ with obvious meaning of the symbols involved. The boundary conditions for the pore pressure π are imposed by means of the seepage velocity $\mathbf{w} = \mathbf{v}_\ell^{PDL} - \mathbf{v}_s^{PDL}$. Neglecting volume forces, pore pressure and seepage velocity are related by the Darcy law

$$\mathbf{w} = -k \operatorname{grad} \pi.$$

The hydraulic conductivity k can be a function of the displacement. A particular form has been discussed in Sec. 2.3.5. On Γ_I, interstitial fluid cannot flow through dentine boundary. Hence, on Γ^I fluid cannot flow in or out the tooth. Hence, velocities of the fluid and of the solid are the same in normal direction. This results in the following boundary Neumann condition

$$-k \operatorname{grad} \pi \cdot \mathbf{n} = 0. \tag{6.5}$$

Same argument applies to Γ_N^{Tooth} and Γ_N^{PDL}.

At the contact between the PDL and the alveolar bone, small channels allow the fluid to escape the PDL volume. We investigate their effect, by imposing for the pressure

- homogeneous Dirichlet boundary condition, i.e. the fluid is free to escape;
- homogeneous Neumann boundary condition, i.e. the fluid is constrained in the PDL.

6.3.2 Continuity equations

Our final aim is the formulation of a coupled system, which allows for describing the mechanical properties of the PDL-tooth system. In our model, we assume that the tooth is a very stiff elastic body and the PDL is a poroelastic material. Hence, to describe the tooth we use the quasi-static elasticity model (1.46) and to describe the PDL we employ the poroelastic model (1.83). The two balances of force that have to be solved on each domain are:

$$-\operatorname{Div} \mathbf{P}^{Tooth} = \mathbf{0} \tag{6.6}$$

and

$$-\operatorname{Div}\left[\mathbf{P}_s^{PDL} + \mathbf{P}_\ell^{PDL}\right] \quad = \quad \mathbf{0}. \tag{6.7}$$

Between the two domains a transmission condition for the continuity of stress has to be imposed. In material coordinates it reads

$$\mathbf{P}^{Tooth}\mathbf{N}^{Tooth} = -(\mathbf{P}_s^{PDL} + \mathbf{P}_\ell^{PDL})\mathbf{N}^{PDL}. \tag{6.8}$$

Assuming that the fluid is inviscid and the components of the solid phase are incompressible, the total stress tensor in (6.7) takes the form

$$\mathbf{P}^{PDL} - J^{PDL}\pi(\mathbf{F}^{PDL})^{-T}. \tag{6.9}$$

The tensors $\mathbf{P}^{Tooth}$ and $\mathbf{P}^{PDL}$ describe the material properties of the tooth and of the solid phase of the PDL, respectively. For $\mathbf{P}^{Tooth}$ we employ the neo-Hookean material law (1.49). For $\mathbf{P}^{PDL}$ we use the Fung material law (1.50).

Because of the incompressibility hypothesis, on $\mathscr{B}_0^{PDL}$ also the following continuity equation has to be solved

$$\dot{J}^{PDL} = \operatorname{Div}\left[(J^{PDL} - \Phi_{s0})(\mathbf{F}^{PDL})^{-1}\mathbf{K}(\mathbf{F}^{PDL})^{-T}\operatorname{Grad}(\pi)\right]. \tag{6.10}$$

We remind that Φ_{s0} is the solid volume fraction in the reference configuration, i.e. the local percentage of solid in the unstressed configuration. This equation is mathematically well-defined as long as $J \geq \Phi_{s0}$. If equality is reached, it is said that the material has reached its *compaction point*. It means that, in a reference volumetric element (RVE) of material, only solid component remains. When compaction is reached, because of the

incompressibility hypothesis, material should behave as an incompressible solid. This constraint is not enforced in (6.7) and (6.9) and the model stops being valid when this point is reached. We will use this condition as exit condition for our simulations.

The force response of the system is measured on the surface Γ_D^{Tooth} at the position of the bolt. Thus, the resulting force is evaluated as the residual boundary forces as

$$\int_{\Gamma_D^{\text{Tooth}}} \mathbf{P}^{Tooth}\mathbf{N}_x \cdot \mathbf{N}_x.$$

6.3.3 Parameters of the model

The presented model depends on six physical parameters, i.e. the two elastic parameters for the tooth, the two elastic parameter of the PDL, the permeability k, and the stiffening parameter a of the Fung potential used for the PDL. The chosen mechanical parameters of the tooth-PDL system are reported in Tab. 6.1. The parameters of the tooth are similar to the ones usually measured for teeth [Ber08, KMM03]. The elastic parameters employed for the PDL are similar to the ones presented in [NPSS02]. In [Ber08], a value of 0.02 MPa was used for the PDL.

As presented in the work of Berkowitz [BHM09], the solid-fluid ratio is not uniform in the PDL. Solid volume fraction ranges from 53% to 74%. Since it is not possible to know the local effective distribution of the phases in the PDL, we employ a constant value $\Phi_{s0} = 0.60$. This value is larger from the others used in literature. In [NPSS02] and [Ber08], the employed values were 0.4 and 0.37, respectively.

Parameter	E^{PDL}	ν^{PDL}	E^{Tooth}	ν^t	Φ_{s0}	$\Phi_{\ell 0}$
value	0.2×10^6	0.3	20×10^9	0.31	0.6	0.4
units	$\text{N}\cdot\text{m}^{-2}$	#	$\text{N}\cdot\text{m}^{-2}$	#	#	#

Table 6.1. Physical parameters used in the numerical experiments.

In the next sections, we investigate the role of the stiffening parameter a and the permeability k on the force response.

6.4 Multigrid Newton's method for the coupled system

Conservation Eqs. (6.6) and (6.7) can be written as a unique continuity equation

$$-\text{Div}\,\mathbf{P} = \mathbf{0} \quad \text{on } \mathscr{B}_0,$$

where $\mathscr{B}_0 = \overline{\overline{\mathscr{B}_0^{Tooth}} \dot{\cup} \overline{\mathscr{B}_0^{PDL}}}$. The Piola-Kirchhoff tensor $\mathbf{P}_0$ is defined as follows

$$\mathbf{P} = \begin{cases} \mathbf{P}^{PDL} - J^{PDL}\pi(\mathbf{F}^{PDL})^{-1} & \text{in } \mathscr{B}_0^{PDL} \\ \mathbf{P}^{Tooth} & \text{in } \mathscr{B}_0^{Tooth} \end{cases} \tag{6.11}$$

and, hence, it takes different forms depending on the subdomain. Because of the compatibility condition (6.2), $\mathbf{P}$ can be seen as a function of the displacement $\mathbf{u}$, that is

$$\mathbf{u} = \begin{cases} \mathbf{u}^{PDL} & \text{in } \mathscr{B}_0^{PDL} \\ \mathbf{u}^{Tooth} & \text{in } \mathscr{B}_0^{Tooth}. \end{cases} \tag{6.12}$$

$\mathbf{P}$ depends on π only in $\mathscr{B}_0^{PDL}$. We define the interface between the two domains as $\Gamma_I = \overline{\mathscr{B}_0^{Tooth}} \cap \overline{\mathscr{B}_0^{PDL}}$.

For the discretization of the problem, on the domain $\mathscr{B}_0$ a regular triangulation $\mathscr{T}_h$ of tetrahedral elements is introduced. We assume that Γ_I does not cut any element. Hence, we can see this as two triangulations of the subdomains that are compatible on Γ_I.

We consider a discrete formulation with linear FEs for both displacement and pressure. The vectors of the FE unknowns are denoted with different subscripts. We use the subscript T to denote the FE vectors of the nodes that belong to $\mathscr{B}_0^{Tooth} \backslash \Gamma_I$, e.g. $\underline{u}_T$. We use the subscript P to denote the FE vectors of the nodes that belong to $\mathscr{B}_0^{PDL} \backslash \Gamma_I$, e.g. $\underline{u}_P$. We use the subscript Γ to denote the FE vectors of the nodes that belong to Γ_I, e.g. $\underline{u}_\Gamma$.

At each Newton step, a linear system of the following form has to be solved:

$$\begin{pmatrix} \mathbf{A}_{TT} & \mathbf{A}_{T\Gamma} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_{\Gamma T} & \mathbf{A}_{\Gamma\Gamma} & \mathbf{A}_{\Gamma P} & -\mathbf{B}_{\Gamma\Gamma} & -\mathbf{B}_{\Gamma P} \\ \mathbf{0} & \mathbf{A}_{P\Gamma} & \mathbf{A}_{PP} & -\mathbf{B}_{P\Gamma} & -B_{PP} \\ \mathbf{0} & -\mathbf{C}_{\Gamma\Gamma} & -\mathbf{C}_{\Gamma P} & -\mathbf{D}_{\Gamma\Gamma} & -\mathbf{D}_{\Gamma P} \\ \mathbf{0} & -\mathbf{C}_{P\Gamma} & -\mathbf{C}_{PP} & -\mathbf{D}_{P\Gamma} & -\mathbf{D}_{PP} \end{pmatrix} \begin{pmatrix} \underline{h}_T \\ \underline{h}_\Gamma \\ \underline{h}_P \\ \underline{\theta}_\Gamma \\ \underline{\theta}_P \end{pmatrix} = \begin{pmatrix} \underline{f}_T \\ \underline{f}_\Gamma \\ \underline{f}_P \\ \underline{g}_\Gamma \\ \underline{g}_P \end{pmatrix}. \tag{6.13}$$

The presence of a high jump in the coefficients in the coupled problem introduces two further issues in the application of MG to (6.13). It heavily reduces the convergence rate of the smoothers and it may prevent the convergence of iterative solvers on the coarse level. Because of the former, we use a MG Newton's Method (MGNM) with a fixed number (15) of W multigrid cycles. The latter effect may be overcome by the use of direct solvers. Indeed, in the performed simulations we observed a non-convergence of the GMRES on the coarse level.

A further issue that affects the application of the MG to this problem is related to the stability condition presented in Chap. 4. In fact, the efficiency of MG is based on the principle that the coarse grid operators have to be a good representation of the fine problem. It is not sufficient that the stability condition is met on the finest levels because

oscillations in the correction obtained from unstable levels may lead to a divergence of MG. To this purpose, we extend the stabilization strategy presented in [GLR10]. In this work, the authors propose to introduce a Finite Pressure Laplacian stabilization on all the levels.

Our idea is to check the stability during the assembling routine. Thanks to the stability limit derived in Chap. 4, we check if the local stiffness matrix is stable and, if it is not stable, we introduce a numerical diffusion k_h that is the smallest that renders the local matrix stable.

We assemble the local stiffness matrix $\mathbf{d}_K$ of the non-linear diffusion operator on the element K. We assemble the reduced mass matrix that in this case reads

$$\mathbf{m}_K^R = \frac{1}{32}\frac{J}{(\mu+G)}\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{vmatrix}. \tag{6.14}$$

Hence, we study the M-matrix property of $\mathbf{s}_K = (1+\beta)\mathbf{d}_K + \mathbf{m}_K^R$ as a function of β. We find the smallest positive β that renders the matrix $\mathbf{s}_K$ an M-matrix.

Hence, the permeability is modified as follows:

$$\begin{cases} k_h = k & \text{if } \beta = 0 \\ k_h = k(1+\beta) & \text{if } \beta > 0. \end{cases} \tag{6.15}$$

Before proceeding with the numerical tests, we remind the discretization and solution scheme that we employ:

1. implict Euler scheme for time integration;

2. linear FEs for the spatial discretization of $\mathbf{u}$ and π with element-wise FPL stabilization;

3. MGNM for the efficient solution of the linear system:

   - projection of the current displacement and pressure on the coarse levels to assemble the coarse operators;
   - W-cycle;
   - 15 MG cycles at each Newton step;
   - 5 steps of Gauss-Seidel smoother;
   - direct solver (MUMPS) for the solution of the coarse operator.

refinements	0	1	2	3	4
mesh nodes	10k	84k	645k	5M	42M
# cores	1	2	16	128	1024

Table 6.2. Number of nodes.

6.4.1 Efficiency of the solver in a realistic case

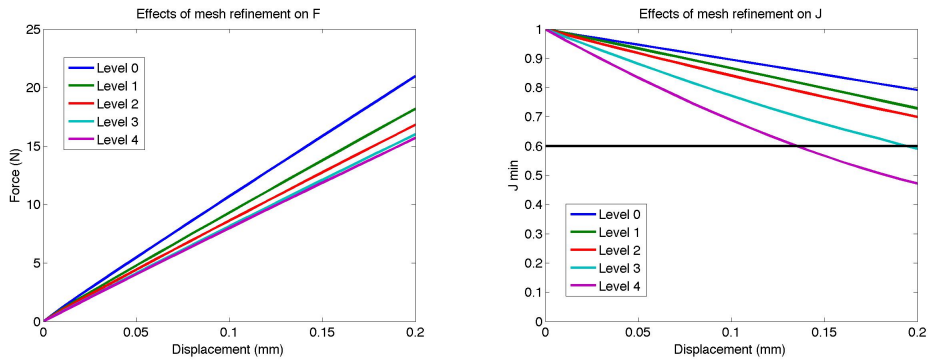
We test the efficiency of our solution method in the simulation of a realistic test. We reproduce an indentation experiment on the mesh reported in Fig. 6.5. Moreover, this test allows to study the effect of mesh refinement on the force response. In fact, as reported also in [PCB+02], coarse meshes of the PDL are responsible for an overestimation of the simulated force response. In order to have physically meaningful results, a very high resolution of the domain is necessary.

The initial mesh has 10k vertices and it is refined up to 4 times. Hence, we test the MGNM with a number of levels from 2 to 5. The number of refinements, mesh nodes, and used cores is reported in Tab. 6.2.

In Fig. 6.9a, the time evolution of the force response for different numbers of refinements is reported. The force response between the experiments with 1 and 4 refinements is stiffer of about 25%. A convergence of the force is observed for increasing number of mesh refinements.

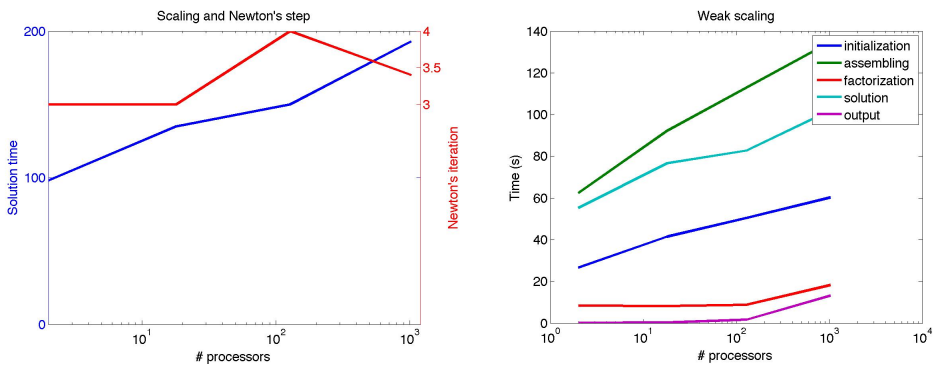
We also studied the effect of the mesh refinement on the minimum of the deformation determinant. The study of this variable is necessary to guarantee the physical validity of our model: values of J small than 0.6 give a negative value of the permeability. The results are reported in Fig. 6.9b. Differently from the force, J does not show a convergence with respect to the mesh refinement. Further simulations with different geometries, models and parameters showed that this effect depends on the geometrical settings of our problem. Indeed, the same behavior has been observed in a toy problem where the tooth was substituted by a hexahedron and the PDL by a layer partially surrounding the hexahedron. Similar simulations have also been performed using an elastic material law and different material parameters, always showing a non-convergence of J. Finally, we want to state that no convergence prove exists for the deformation determinant.

In Fig. 6.10a we notice that, even if we are using a MGNM with a fixed number of MG cycles, the number of Newton iterations remains almost constant with the number of refinements but an increase of the total computation time is observed. This can be explained studying the times of the different parts of the solution strategy, see Fig. 6.10b. Differently from the test case in Sec. 5.7.1, an increase in time of the assembling and the solution is observed. This difference can be explained by a non-optimal load balancing. In fact, even if a subdomain has just one element more then the others on the coarsest level, this difference is magnified up to 8^4 elements when four refinements are applied. This affects the scaling of a parallel solver.



(a) *(b)*

Figure 6.9. Force response (a) and minimum of J (b) for different levels for an indentation experiment.



(a) *(b)*

Figure 6.10. Total time and time of the different parts of the code.

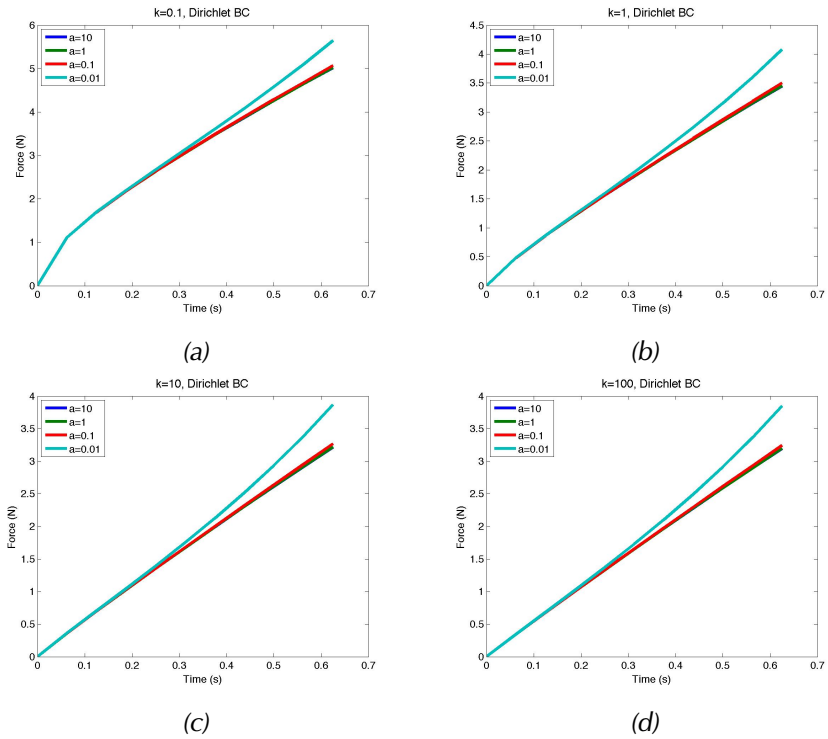


Figure 6.11. Effect of permeability on the force response in a linear loading experiment.

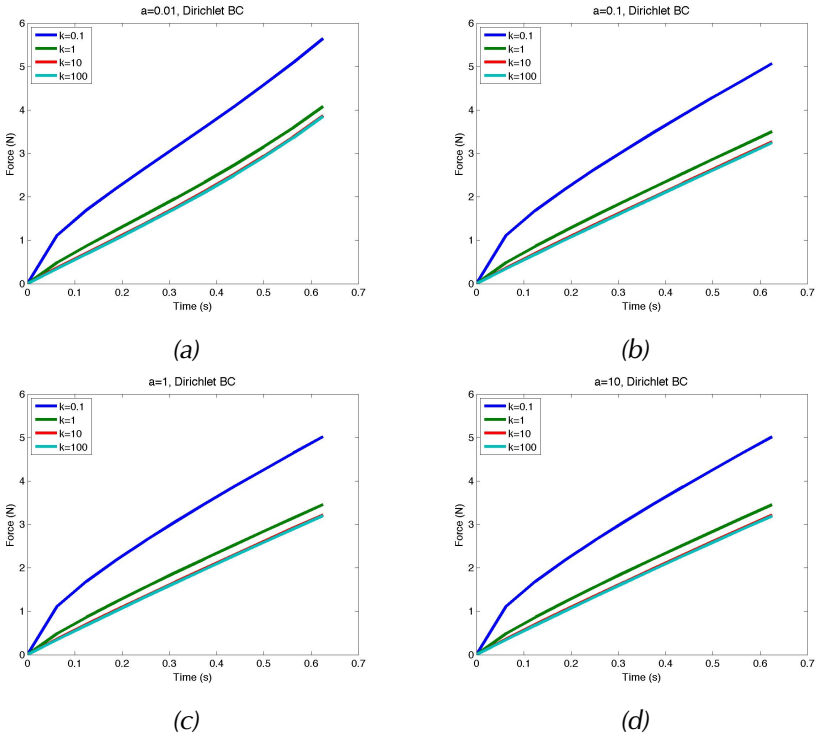


Figure 6.12. Effect of the stiffening parameter a on the force response in a linear loading experiment.

6.5 Validation and sensitivity analysis

In order to validate the model in a realistic case, we simulate a linear loading and a relaxation test. The geometry employed is the one showed in Fig. 6.5. The experiments have been realized with three refinements of the original mesh.

In our experiment we study the effect of the two "free" parameters, i.e. the stiffening parameter a of the Fung material law and the permeability k.

In Fig. 6.11, we report the force responce for different permeabilities. The permeability influences the force in the early stages of the experiment. In particular, smaller permeabilities give a concave and stiffer force response. Instead, the parameter a affects the force response for large displacement.

The optimal fit has been found for $a = 0.0065$ and $k = 100$. In Fig. 6.13, we compared the simulated force obtained with the optimal parameters and the measured one in an indentation experiment for a porcine and a human tooth. For both the specimina, we stopped our simulation when the compaction point was reached. In Fig. 6.13a, this happens where the change of curvature in the experimental response is observed. Hence, this may be explained by the compaction of the material. For a linear material law (green line), the compaction point is reached for larger displacements.

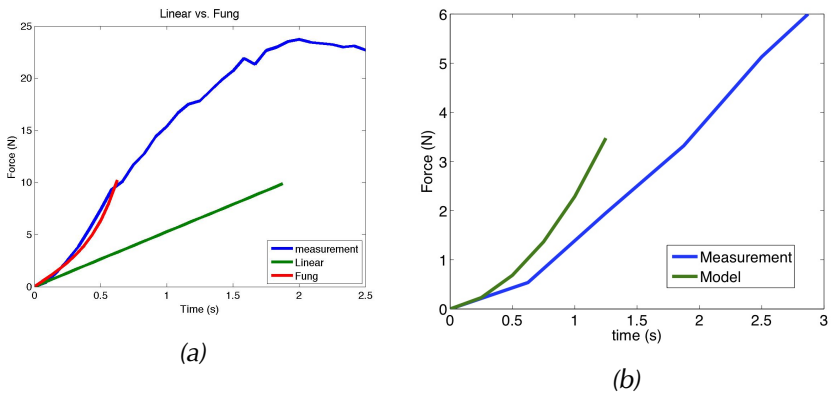


Figure 6.13. Comparison between simulated and measured force, for a porcine tooth (a) and a human tooth (b).

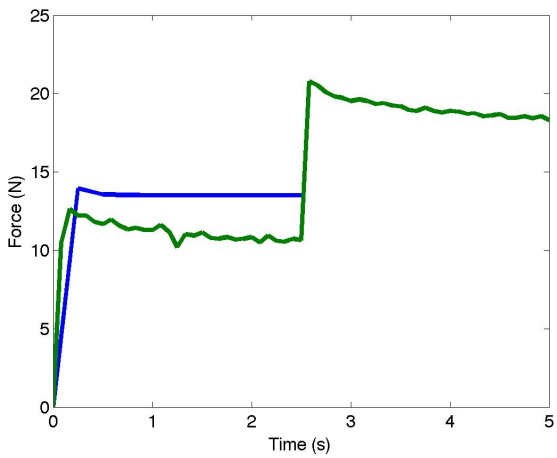


Figure 6.14. Comparison between simulated and measured force of a relaxation experiment.

Finally, we used the obtained parameters for the simulation of a relaxation experiment. The results can be found in Fig. 6.14. The simulated and measured forces have a comparable behaviour. In both of them an instantaneous peak and a following relax-

ation phase are observed. After the first displacement step, the simulated value is about 15% larger than the measured one. In the simulation, the relaxation time is shorter than the measured one. Our simulation stops after the first displacement step because the compaction point would be reached with the second step.

In numerical tests with different boundary conditions, we observed that the force response was larger when the flux at the boundary is not considered.

6.6 Hyperelastic material law with unilateral constraint of compaction

In a poroelastic material, even if the two phases are incompressible, local volume changes may happen. The fluid can flow from one point to others and this is described by changes of the spatial volume fractions. If all the fluid has moved out of a RVE , the solid is at compaction and the solid phase should behave as an incompressible solid. Hence, if this phenomenon has to be included in the material law, a term that diverges for $J \to \Phi_{S0}$ has to be present in the equations of the poroelastic model.

A possible formulation for the elastic constitutive equations that includes this limit has be presented in [FG12b]. Before briefly discussing its numerical implementation, we want to point out that it cannot be realized by changing the elastic energy or by using a decoupled approach of shear and volumetric parts. Indeed, a potential that diverges at Φ_{S0} would not be compatible with linear elasticity and the decoupled approach would be appropriate only for isochoric material laws.

The idea presented in [FG12b] is to add to the elastic energy an additional term $U(J)$ such that

- it is active only in compression;
- it respects the zero-zero condition, i.e. it should not contribute to the stiffness at zero strain.

The proposed potential is:

$$U(J) = H(J_{crit} - J)(J - J_{crit})^{2q}(J - \Phi_{S0})^{-r}, \tag{6.16}$$

where H is the Heaviside function and J_{crit} is the point where that potential becomes active. The value of J_{crit} has to belong to the interval $(\Phi_{S0}, 1)$. The parameters $q \geq 2$ and $r \in (0, 1)$ are related to the material properties of the matrix.

The first Piola-Kirchhoff tensor obtained deriving this potential is

$$\mathbf{P}_U = H(J_{crit} - J)(2q(J - J_{crit})^{2q-1}(J - \Phi_{S0})^{-r}\mathbf{F}^{-T} - r(J - J_{crit})^{2q}(J - \Phi_{S0})^{-r-1}\mathbf{F}^{-T}).$$

Adding this tensor to the material law enforces the constraint of incompressibility for $J \to \Phi_{s0}$. Implementation of this constraint to the material law is responsible for

volumetric locking. This phenomenon usually happens for incompressible materials when they are discretized with low order FEs. Two approaches are possible to overcome this problem: the use of high order FEs or the introduction of a SP formulation. Since we focus on low order FEs, we propose a possible SP formulation for the second strategy.

Roughly speaking, we consider (6.16) as a sort of volumetric term whose bulk modulus is given by $(J - \Phi_{s0})$. Indeed, it is the term that diverges for $J \to \Phi_{s0}$. As done in (3.25) for almost incompressible linear elasticity, we define the function

$$p = (J - \Phi_{s0})^r (J - J_{crit})^q. \tag{6.17}$$

Hence, the potential would take the form

$$U(J, p) = pH(J_{crit} - J)(J - J_{crit})^q. \tag{6.18}$$

Derivative of (6.18) and definition (6.17) give rise to the system

$$\begin{cases} -\mathrm{Div}\,(\mathbf{P} & + & pq(J - J_{crit})^{q-1} & = & \mathbf{0} \\ (J - J_{crit})^q & - & \frac{p}{(J-\Phi_{s0})} & = & 0 \end{cases}. \tag{6.19}$$

This system has to be coupled with the evolution of the deformation gradient (6.10). Hence, the poroelasticity system with the unilateral constraint (6.16) has been written as SP problem with a primal variable $\mathbf{u}$ and two Lagrange multipliers π and p. Finally, this system can be discretized also with linear FEs for the three unknowns.

6.7 Summary

The poroelastic model presented in this chapter allowed for a good representation of the mechanical force response of the tooth. In particular:

- We derived two optimal values for a and k.
- We were able to reproduce the early stages of linear loading experiment up to 0.05 mm and 10 N.
- The linear loading experiments for this displacement reach the compaction point. This corresponds with the change of curvature in the measured force.
- In the relaxation experiments, the force was overestimated and the relaxation time was a bit shorter than the observed one.
- A good approximation has also been obtained on human tooth.

Conclusions and future perspectives

We presented the formal derivation of a poroelastic model including inhomogeneities due to volumetric growth. The study of the residual dissipation allowed us to formally derive Darcy law for the flow of fluid in a porous medium.

We proposed a discretization strategy based on the implicit Euler scheme, Newton's method and a mixed finite element method. Our solver was coupled with a time-integration strategy for the evolution of growth under the hypothesis that the tensor of inelastic rate is given. This solution strategy has been validated in a simulation of a confined compression test against a different solution method.

We showed that, in the linear poroelastic problem, linear finite element spaces for both displacement and pressure are stable even if they are not LBB compliant.

We derived a Peclet number for the poroelasticity problem in two and three dimensions and showed how it depends on the physical and discretization parameters. In particular, we found that this Peclet number has to be smaller than one to prevent instabilities. The Peclet number allowed to find an optimal parameter for the fluid pressure Laplacian stabilization method to regularize unstable problems.

We presented a coupled model for a tooth-PDL system. The PDL was modeled as a poroelastic material and the tooth as an elastic one. We developed an efficient solution strategy based on the multigrid Newton's method with a fixed number of multigrid cycles. We introduced some tunings to improve convergence: direct solver on the coarsest level, W-cycle, stabilization of coarser levels if the grid does not respect the stability condition, and point-block Gauss-Seidel smoother. With these choices the inexact solver converged with a fixed number of Newton iterations and demonstrated good weak scalability.

With this coupled model, we simulated indentation experiments on porcine and human teeth. We were able to reproduce the force response of linear loading experiments up to 10 N and displacements up to 0.05 mm. In the relaxation test, we obtained comparable force response but the relaxation times were underestimated.

We contributed to the project "Aetiology and Sequelae of Periodontal Diseases. Genetic, Cell Biological and Biomechanical Aspects", with the realization of an efficient simulation tool for the reproduction of loading experiments on teeth. Our software will allow the workgroup to understand the change of mechanical parameters due to diseases, aging, and application of dental braces.

Finally, we want to point out other topics that should be the subject of future research.

- In our solution method, we considered the rate of inelastic deformation as a data of the problem. In realistic cases, it may also be dependent on the current deformation or other variables of the problem. For this reason, our solution strategy has to be extended. A possible solution would be to adapt the return mapping algorithm used in plasticity to the field of growth.

- The stability limit that we derived was obtained only by means of numerical simulations. An analytical proof is missing but simulations confirm the validity of our result.

- In the multigrid solver, we employed a point-block Gauss-Seidel smoother. Other more efficient smoothers used for the linear poroelastic model have to be investigated also in the non-linear case.

- In the simulations performed on the tooth, we were able to reproduce large deformations of the PDL up to 25% of the domain-size. At this point, compaction of the material was reached. Material laws that include the volumetric constraint have also to be considered in order to simulate the complete indentation experiments.

- The simulation of relaxation experiments showed that the relaxation time was shorter than the measured one. Since it heavily depends on the permeability, more appropriate models for it have to be taken into account.

Appendix A

Linearization of kinematical descriptors

In this thesis, we used the Newton-Raphson method for the solution of nonlinear poroelasticity problems. It requires the linearization of the conservation equations with respect to the displacement $\mathbf{u}$. The nonlinearities are found in the deformation determinant, the inverse of the deformation gradient, and the inverse of the Cauchy-Green tensor. They are dependent on $\mathbf{u}$ as

$$J = \det(\mathbf{I} + \mathbf{U}),$$

$$\mathbf{F}^{-1} = (\mathbf{I} + \mathbf{U})^{-1},$$

and

$$\mathbf{C}^{-1} = \mathbf{F}^{-1}\mathbf{F}^{-T} = (\mathbf{I} + \mathbf{U})^{-1}(\mathbf{I} + \mathbf{U})^{-T},$$

with $\mathbf{U} = \nabla\mathbf{u}$. Since $\mathbf{F}$ depends on $\mathbf{u}$ through the affine transformation $\mathbf{I} + \nabla(\cdot)$, the directional derivatives can be computed with respect to the deformation gradient. In the following we report the formulas for the computation of the linearization of these basic quantities by means of the application of the definition of directional derivative.

A.1 Linearization of determinant of tensor

In order to compute the linearization of the deformation determinant, we first consider the incremental ratio

$$\frac{\det(\mathbf{F} + \epsilon\mathbf{H}) - \det(\mathbf{F})}{\epsilon}.$$

Collecting J, he previous expression can be written as

$$J\frac{\det(\mathbf{F})^{-1}\det(\mathbf{F}+\epsilon\mathbf{H})-1}{\epsilon}=J\frac{\det(\mathbf{I}+\epsilon\mathbf{F}^{-1}\mathbf{H})-1}{\epsilon}.$$

Defining $\tilde{\mathbf{H}}=\mathbf{F}^{-1}\mathbf{H}$, the determinant of $\mathbf{I}+\epsilon\tilde{\mathbf{H}}$ can be written as $1+\epsilon\mathrm{tr}\,(\tilde{\mathbf{H}})+\mathscr{O}(\epsilon^2)$. Hence the following expression holds true

$$J\frac{\det(\mathbf{I}+\epsilon\mathbf{F}^{-1}\mathbf{H})-\det(\mathbf{I})}{\epsilon}=J\frac{\epsilon\mathrm{tr}\,(\tilde{\mathbf{H}})+\mathscr{O}(\epsilon^2)}{\epsilon}=J\mathrm{tr}\,(\tilde{\mathbf{H}})+\mathscr{O}(\epsilon).$$

Now for $\epsilon\to 0$ and using the property of trace, we obtain

$$\mathfrak{D}_{\mathbf{F}}J\,[\mathbf{H}]=\mathfrak{D}_{\mathbf{U}}J\,[\mathbf{H}]=J\mathbf{F}^{-T}:\mathbf{H}.$$

A.2 Linearization of the inverse of a tensor

As in the previous section, we start from the incremental ratio for the computation of the Gâteaux differential:

$$\frac{(\mathbf{F}+\epsilon\mathbf{H})^{-1}-\mathbf{F}^{-1}}{\epsilon}.$$

Multiplying from the right by $\mathbf{F}\mathbf{F}^{-1}$, the previous expression can be written as

$$\frac{(\mathbf{F}+\epsilon\mathbf{H})^{-1}\mathbf{F}-\mathbf{F}^{-1}\mathbf{F}}{\epsilon}\mathbf{F}^{-1}=\frac{(\mathbf{F}+\epsilon\mathbf{H})^{-1}(\mathbf{F}^{-1})^{-1}-\mathbf{I}}{\epsilon}\mathbf{F}^{-1}=\frac{(\mathbf{I}+\epsilon\mathbf{F}^{-1}\mathbf{H})^{-1}-\mathbf{I}}{\epsilon}\mathbf{F}^{-1}.$$

If coincides with the computation of the inverse of the identity matrix by setting the increment $\tilde{\mathbf{H}}=\mathbf{F}^{-1}\mathbf{H}$. Hence we consider the following incremental ratio

$$\frac{(\mathbf{I}+\epsilon\tilde{\mathbf{H}})^{-1}-\mathbf{I}}{\epsilon}=\frac{(\mathbf{I}-\epsilon\tilde{\tilde{\mathbf{H}}})^{-1}-\mathbf{I}}{\epsilon}.$$

with $\tilde{\tilde{\mathbf{H}}}=-\tilde{\mathbf{H}}$. Because we are interested in computing the limit for $\epsilon\to 0$, we can use the formula for the geometric series of matrices, and the last expression becomes

$$\frac{(\mathbf{I}-\epsilon\tilde{\tilde{\mathbf{H}}})^{-1}-\mathbf{I}}{\epsilon}=\frac{\sum_{n=0}^{\infty}(\epsilon\tilde{\tilde{\mathbf{H}}})^n-\mathbf{I}}{\epsilon}=\frac{\sum_{n=1}^{\infty}(\epsilon\tilde{\tilde{\mathbf{H}}})^n}{\epsilon}=\tilde{\tilde{\mathbf{H}}}+\sum_{n=2}^{\infty}(\epsilon^{n-1}\tilde{\tilde{\mathbf{H}}}^n).$$

Taking the limit in the last expression and using the definition $\tilde{\tilde{\mathbf{H}}}$, we obtain that

$$\mathfrak{D}_{\mathbf{F}}\mathbf{F}^{-1}[\mathbf{H}]=\mathfrak{D}_{\mathbf{U}}\mathbf{F}^{-1}[\mathbf{H}]=-\mathbf{F}^{-1}\mathbf{H}\mathbf{F}^{-1}. \tag{A.1}$$

Since the transpose operator is linear, we directly obtain

$$\mathfrak{D}_{\mathbf{F}}\mathbf{F}^{-T}[\mathbf{H}]=-\mathbf{F}^{-T}\mathbf{H}^T\mathbf{F}^{-T}. \tag{A.2}$$

A.3 Linearization of the inverse of the Cauchy-Green tensor

The computation of the differential of $\mathbf{C}^{-1} = \mathbf{F}^{-1}\mathbf{F}^{-T}$ with respect to the deformation gradient can be done either applying the product rule or alternatively the chain rule. Using the first strategy we obtain

$$\mathfrak{D}_{\mathbf{F}}\mathbf{C}^{-1}[\mathbf{H}] = \mathfrak{D}_{\mathbf{F}}(\mathbf{F}^{-1}\mathbf{F}^{-T})[\mathbf{H}] = -\mathbf{F}^{-1}\mathbf{H}\mathbf{F}^{-1}\mathbf{F}^{-T} - \mathbf{F}^{-1}\mathbf{F}^{-T}\mathbf{H}^{T}\mathbf{F}^{-T}.$$

This last term can also be written as

$$\mathfrak{D}_{\mathbf{F}}\mathbf{C}^{-1}[\mathbf{H}] = -\mathbf{F}^{-1}\mathbf{H}\mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{H}^{T}\mathbf{F}^{-T}.$$

or

$$\mathfrak{D}_{\mathbf{F}}\mathbf{C}^{-1}[\mathbf{H}] = -\mathbf{C}^{-1}\mathbf{F}^{T}\mathbf{H}\mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{H}^{T}\mathbf{F}\mathbf{C}^{-1} = -\mathbf{C}^{-1}(\mathbf{F}^{T}\mathbf{H} + \mathbf{H}^{T}\mathbf{F})\mathbf{C}^{-1}.$$

Notice that the last term in the previous equation equals

$$\mathfrak{D}_{\mathbf{C}}\mathbf{C}^{-1}[\mathfrak{D}_{\mathbf{F}}\mathbf{C}[\mathbf{H}]].$$

Bibliography

[AG07] D. Ambrosi and F. Guana. Stress-modulated growth. *Mathematics and Mechanics of Solids*, 12:319–342, 2007.

[AGDM08] D. Ambrosi, A. Guillou, and E. Di Martino. Stress-modulated remodeling of a non-homogeneous body. *Biomechan. Model. Mechanobiol.*, 7:63–76, 2008.

[AGLR08] G. Aguilar, F. Gaspar, F. Lisbona, and C. Rodrigo. Numerical stabilization of biot's consolidation model by a perturbation on the flow equation. *International Journal for Numerical Methods in Engineering*, 75(11):1282–1300, 2008.

[AMPM91] K.L. Andersen, H.T. Mortensen, E.H. Pedersen, and B. Melsen. Determination of stress levels and profiles in the periodontal ligament by means of an improved three-dimensional finite element model for various types of orthodontic and natural force systems. *Journal of Biomedical Engineering*, 13(4):293 – 303, 1991.

[APV10] D. Ambrosi, L. Preziosi, and G. Vitale. The insight of mixtures theory for growth and remodeling. *Z. Angew. Math. Phys.*, 61:177–191, 2010.

[AW10] G. A. Ateshian and J. A. Weiss. Anistropic hydraulic permeability under finite deformation. *Journal of Biomechanical Engineering*, 132:111004–1, 2010.

[Bal76] JohnM. Ball. Convexity conditions and existence theorems in nonlinear elasticity. *Archive for Rational Mechanics and Analysis*, 63(4):337–403, 1976.

[BBF08] Daniele Boffi, Franco Brezzi, and Michel Fortin. Finite elements for the Stokes problem. In *Mixed finite elements, compatibility conditions, and applications. Lectures given at the C.I.M.E. summer school, Cetraro, Italy, June 26–July 1, 2006*, pages 45–100. Berlin: Springer; Florenz: Fondazione CIME Roberto Conti, 2008.

[BBJ+97] P. Bastian, K. Birken, K. Johannsen, S.Lang, N. Neuß, H. Rentz-Reichert, and C.Wieners. UG – a flexible software toolbox for solving partial differential equations. *Computing and Visualization in Science*, 1:27–40, 1997.

[BCP95] K. E. Brenan, S. L. Campbell, and L. R. Petzold. *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*. Society for Industrial and Applied Mathematics, 1995.

[Bea72] J. Bear. *Dynamics of fluid in Porous Media*. Dover Publications, INC. New York, 1972.

[Ber08] M. Bergomi. *Experimental and Numerical Investigations on the Fluid Contribution to the Tensile-compressive Mechanical Behavior of the Bovine Periodontal Ligament*. 2008.

[BF91] F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element Methods*. Springer-Verlag, 1991.

[BGL05] M. Benzi, G. H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Numerica*, 14:1–137, 2005.

[BHM92] B. K. B. Berkovitz, G. R. Holland, and B. J. Moxham. *A Colour Atlas and Textbook of Oral Anatomy. Histology and Embryology*. Wolfe Medical Publications, London, 2nd edition, 1992.

[BHM09] B. K. B. Berkowitz, G. R. Holland, and B. J. Moxham. *Oral anatomy, Histology and Embryology*. Mosby, 4th edition, 2009.

[Bio40] M. A. Biot. General theory of three-dimensional consolidation. *Journal of applied physics*, 12(2):155–164, 1940.

[BJ84] F. Brezzi and J.Pitkäranta. On the stabilization of finite element approximations of the Stokes equations. In W. Hackbusch, editor, *Efficient Solutions of Elliptic Systems*, pages 11–19. Vieweg, 1984.

[BMC00] L. S. Bennethum, M. A. Murad, and J. H. Cushman. Macroscale Thermodynamics and the Chemical Potential for Swelling Porous Media. *Transport in Porous Media*, 39:187–225, 2000.

[Bon08] D. Bonet, J.and Wood. *Nonlinear continuum mechanics for finite element analysis*. Cambridge, 2 edition, 2008.

[Bra09] Dietrich Braess. *Finite Elements : Theory, Fast Solvers, and Applications in Solid Mechanics*. Cambridge University Press, 2009.

[CFS01] P. Cermelli, E. Fried, and S. Sellers. Configurational stress, yield and flow in rate-independent plasticity. *Proc. R. Soc. Lond. A*, 457:1447–1467, 2001.

[CR73] P.G Ciarlet and P.-A Raviart. Maximum principle and uniform convergence for the finite element method. *Computer Methods in Applied Mechanics and Engineering*, 2(1):17 – 31, 1973.

[Dav04] Timothy A. Davis. Algorithm 832: Umfpack v4.3—an unsymmetric-pattern multifrontal method. *ACM Trans. Math. Softw.*, 30:196–199, June 2004.

[DKH+11] M. Drolshagen, L. Keilig, I. Hasan, S. Reimann, J. Deschner, K.T. Brinkmann, R. Krause, M. Favino, and C. Bourauel. Development of a novel intraoral measurement device to determine the biomechanical characteristics of the human periodontal ligament. *Journal of Biomechanics*, 44(11):2136 – 2143, 2011.

[DQ02] A. DiCarlo and S. Quiligotti. Growth and balance. *Mechanics Research Communications*, 29, 449–456 2002.

[DS35] M. L. Dyment and J. L. Synge. The elasticity of the periodontal membrane. *Oral Health*, (25):105–109, 1935.

[EM00] M. Epstein and G. A. Maugin. Thermomechanics of volumetric growth in uniform bodies. *International Journal of Plasticity*, 16:951–978, 2000.

[EM01] W. Ehlers and B. Markert. A linear viscoelastic biphasic model for soft tissues based on the theory of porous media. *Journal of Biomechanical Engineering*, 123:418–424, 2001.

[EM02] W. Ehlers and B. Markert. A viscoelastic two-phase model for cartilage tissues. In Wolfgang Ehlers, editor, *IUTAM Symposium on Theoretical and Numerical Methods in Continuum Mechanics of Porous Materials*, volume 87 of *Solid Mechanics and Its Applications*, pages 87–92. Springer Netherlands, 2002.

[Eva97] L. C. Evans. *Partial differential equations in action*. American Mathematical Society, 1 edition, 1997.

[FFA11] L. Fusi, A. Farina, and D. Ambrosi. Mathematical modelling of solid liquid mixture with mass exchange between constituents. *Mathematics and Mechanics of Solids*, pages 575–595, 11.

[FG12a] S. Federico and A. Grillo. Elasticity and permeability of porous fibre-reinforced materials under large deformations. *Mechanics of Materials*, 44:58–71, 2012.

[FG12b] Salvatore Federico and Alfio Grillo. Elasticity and permeability of porous fibre-reinforced materials under large deformations. *Mechanics of Materials*, 44(0):58 – 71, 2012. Microstructures and Anisotropies.

[FGD+11] M. Favino, C. Gross, M. Drolshagen, L. Keilig, C. Bourauel, J.Deschner, and R. Krause. Validation of a heterogeneous elastic-biphasic model for the numerical simulation of the PDL. *Computer Methods in Biomechanics and Biomedical Engineering*, 2011.

[FGK] Marco Favino, Alfio Grillo, and Rolf Krause. *A Stability Condition for the Numerical Simulation of Poroelastic Systems*, chapter 110, pages 919–928.

[FGLR+05] S. Federico, A. Grillo, G. La Rosa, G. Giaquinta, and W. Herzog. A transversely isotropic, transversely homogeneous microstructural-statistical model of articular cartilage. *Journal of Biomechanics*, 38(10):2008 – 2018, 2005.

[FGT01] Massimiliano Ferronato, Giuseppe Gambolati, and Pietro Teatini. Ill-conditioning of finite element poroelasticity equations. *International Journal of Solids and Structures*, 38(34Ð35):5995 – 6014, 2001.

[FKJ12] Marco Favino, Rolf Krause, and Steiner Johannes. An efficient preconditioning strategy for schur complements arising from biphasic models. In *Biomedical Engineering / 765: Telehealth / 766: Assistive Technologies*, 2012.

[Fun90a] Y. C. Fung. *Biomechanics: Motion, Flow, Stress, and Growth*. Springer-Verlag, 1 edition, 1990.

[Fun90b] Y. C. Fung. *Biomechanics: Motion, Flow, Stress, and Growth*. Springer-Verlag Inc. New York, 1990.

[GAG+03] K. Garikipati, E. M. Arruda, K. Grosh, H. Narayanan, and S. Calve. A continuum treatment of growth in biological tissue: The coupling of mass transport and mechanics. *Journal of the Mechanics and Physics of Solids*, 2003.

[Gea84] C.W. Gear. Differential-algebraic equations. In EdwardJ. Haug, editor, *Computer Aided Analysis and Optimization of Mechanical System Dynamics*, volume 9 of *NATO ASI Series*, pages 323–334. Springer Berlin Heidelberg, 1984.

[GFA10] M. E. Gurtin, E. Fried, and L. Anand. *The Mechanics and Thermodynamics of Continu*. Cambridge University Press, Cambridge, New York, Melbourne, 2010.

[GFW+09] A. Grillo, S. Federico, G. Wittum, S. Imatani, G. Giaquinta, and M. V. Mićunović. Evolution of a fibre-reinforced growing mixture. *Il Nuovo Cimento*, 32 C(1):97–119, 2009.

[GFW11] A. Grillo, S. Federico, and G. Wittum. Growth, mass transfer, and remodeling in fiber-reinforced, multi-constituent materials. *International Journal of Non-Linear Mechanics*, 47(doi:10.1016/j.ijnonlinmec.2011.09.026), 2011.

[GFW12] A. Grillo, S. Federico, and G. Wittum. Growth, mass transfer, and remodeling in fiber-reinforced, multi-constituent materials. *International Journal of Non-Linear Mechanics*, 2012.

[GGF+12] Alfio Grillo, Chiara Giverso, Marco Favino, Rolf Krause, Michael Lampe, and Gabriel Wittum. Mass transport in porous media with variable mass. In J.M.P.Q. Delgado, Antonio Gilson Barbosa Lima, and Marta Vázquez Silva, editors, *Numerical Analysis of Heat and Mass Transfer in Porous Media*, volume 27 of *Advanced Structured Materials*, pages 27–61. Springer Berlin Heidelberg, 2012.

[Giv03] Chiara Giverso. *Cell and cell-aggregate mechanics: remodelling, growth and interaction with the extracellular environment.* PhD thesis, Politecnico di Torino, 2003.

[GK08] C. Groß and R. Krause. Import of geometries and extended informations into obslib++ using the exodus ii and exodus parameter file formats. Technical Report 712, Institute for Numerical Simulation, University of Bonn, Germany, January 2008.

[GL81] Philip M. Gresho and Robert L. Lee. Don't suppress the wiggles - they're telling you something! *Computers & Fluids*, 9(2):223 – 253, 1981.

[GLR10] F.J. Gaspar, F.J. Lisbona, and C. Rodrigo. Multigrid finite element method on semi-structured grids for the poroelasticity problem. In *Numerical Mathematics and Advanced Applications 2009*, pages 343–350. Springer Berlin Heidelberg, 2010.

[GLV03] F.J. Gaspar, F.J. Lisbona, and P.N. Vabishchevich. A finite difference analysis of biot's consolidation model. *Applied Numerical Mathematics*, 44(4):487 – 506, 2003.

[GWGM09] A. Grillo, G. Wittum, G. Giaquinta, and M. V. Mićunović. A multiscale analysis of growth and diffusion dynamics in biological materials. *International Journal of Engineering Science*, 47:261–283, 2009.

[GZBG07] A. Grillo, G. Zingali, D. Borrello, and G. Giaquinta. Transport phenomena in living systems and continuum physics. *La Rivista del Nuovo Cimento*, 31(11):485–562, 2007.

[Har04] Isaac Harari. Stability of semidiscrete formulations for parabolic problems at small time steps. *Computer Methods in Applied Mechanics and Engineering*, 193(15Ð16):1491 – 1516, 2004. Recent Advances in Stabilized and Multiscale Finite Element Methods.

[HG89] S. M. Hassanizadeh and W. G. Gray. Derivation of conditions describing transport across zones of reduced dynamics within multiphase systems. *Water Resources Research*, 25:529–539, 1989.

[HM90] M. H. Holmes and V. C. Mow. The nonlinear characteristics of soft gels and hydrated connective tissues in ultrafiltration. *J. Biomechanics*, 23(11):1145–1156, 1990.

[Hol01] G. A. Holzapfel. *Nonlinear solid mechanics*. Wiley, 1 edition, 2001.

[HOL12] Joachim Berdal Haga, Harald Osnes, and Hans Petter Langtangen. On the causes of pressure oscillations in low-permeable and low-compressible porous media. *International Journal for Numerical and Analytical Methods in Geomechanics*, 36(12):1507–1522, 2012.

[Hon] Shi Hongtao. http://www.phys-astro.sonoma.edu/people/faculty/shi/research.html.

[HRZ76] E. Hinton, T. Rock, and O. C. Zienkiewicz. A note on mass lumping and related processes in the finite element method. *Earthquake Engineering & Structural Dynamics*, 4(3):245–249, 1976.

[Hug09] T. J. R. Hughes. *The finite element method, linear static and dynamic finite element analysis*. Prentice Hall, 2009.

[HvCAH91] J. M. Huyghe, D. H. van Campen, T. Arts, and R. M. Heethaar. A two-phase finite element model of the diastolic left ventricle. *Journal of Biomechanics*, 24(7):527–538, 1991.

[Isl] Nurul Islam. http://www.kck.usm.my/ppsg/histology/pdl_6.htm.

[Jus04] J. Justiz. *With applications to the periodontal ligament a non-linear large strain viscoelastic law*. PhD thesis, École Polytechnique Fédérale de Lausanne, Switzerland, 2004.

[KBJ03] A. Kawarizadeh, C. Bourauel, and A. JŁger. Experimental and numerical determination of initial tooth mobility and material properties of the periodontal ligament in rat molar specimens. *The European Journal of Orthodontics*, 25(6):569–578, 2003.

[KBZ+04] Afshar Kawarizadeh, Christoph Bourauel, Dongliang Zhang, Werner GÃűtz, and Andreas JÃď'ger. Correlation of stress and strain profiles and the distribution of osteoclastic cells induced by orthodontic loading in rat. *European Journal of Oral Sciences*, 112(2):140–147, 2004.

[KMM03] J.H. Kinney, S.J. Marshall, and G.W. Marshall. The mechanical properties of human dentin: a critical review and re-evaluation of the dental literature. *Crit. Rev. Oral. Biol. Med.*, 2003.

[Krö59] E. Kröner. Die inneren Spannungen und die Inkompatibilitätstensor in der Elastizitätstheorie. *Zeitschrift für Angewandte Mathematik und Physik*, 7:249–257, 1959.

[Krö60] E. Kröner. Kontinuumstheorie der Versetzung und Eigenspannungen. *Archive for Rational Mechanics and Analysis*, 4:273–334, 1960.

[Lee69] E. H. Lee. Elastic-plastic deformation at finite strains. *ASME Transaction on Journal of Applied Mechanics*, 36:1–6, 1969.

[LP03] J. Lindenstrauss and D. Preiss. On Frèchet differentiability of Lipschitz maps between Banach spaces. *Annals of Mathematics*, 2003.

[LSo05] B. Loret and F. M. F. Simões. A framework for deformation, generalized diffusion, mass transfer and growth in multi-species multi-phase biological tissues. *European Journal of Mechanics A/Solids*, 24:757–781, 2005.

[LWW02] S. Lang, C. Wieners, and G. Wittum. *The application of adaptive parallel multigrid methods to problems in nonlinear solid mechanics*. J. Wiley, stein, e. (ed): adaptive methods in solid mechanics edition, 2002.

[MH94] J. E. Marsden and T. J. R. Hughes. *Mathematical Foundation of Elasticity*. Dover editions, 1 edition, 1994.

[MHE10] B. Markert, Y. Heider, and W. Ehlers. Comparison of monolithic and splitting solution schemes for dynamic porous media problems. *International Journal for Numerical Methods in Engineering*, 82(11):1341–1383, 2010.

[Mić09] M. V. Mićunović. *Thermodynamics of Viscoplasticity*, volume 20 of *Advances in Mathematics for Applied Sciences*. Springer Verlag, Heidelberg, 2009.

[ML92] Márcio A. Murad and Abimael F. D. Loula. Improved accuracy in finite element analysis of biot's consolidation problem. *Comput. Methods Appl. Mech. Eng.*, 95(3):359–382, March 1992.

[MNL11] Andrea Malandrino, JŐrŹme Noailly, and Damien Lacroix. The effect of sustained compression on oxygen metabolic transport in the intervertebral disc decreases with degenerative changes. *PLoS Computational Biology*, 7(8), 2011.

[MPK98] Michael I. Miga, Keith D. Paulsen, and Francis E. Kennedy. Von neumann stability analysis of biot's general two-dimensional theory of consolidation. *International Journal for Numerical Methods in Engineering*, 43(5):955–974, 1998.

[NCP+07] A.N. Natali, E.L. Carniel, P.G. Pavan, C. Bourauel, A. Ziegler, and L. Keilig. Experimental-numerical analysis of minipig's multi-rooted teeth. *Journal of Biomechanics*, 40(8):1701 – 1708, 2007.

[NPSS02] A.N. Natali, P.G. Pavan, B.A. Schrefler, and S. Secchi. A multi-phase media formulation for biomechanical analysis of periodontal ligament*. *Meccanica*, 37(4-5):407–418, 2002.

[NW06] J. Nocedal and S. Wright. *Numerical Optimization*. Springers, 2 edition, 2006.

[Ogd84] R. W. Ogden. *Non-Linear Elastic Deformations*. Dover Publications, INC. Mineola, New York, 1984.

[OK08] T. Olsson and A. Klarbring. Residual stresses in soft tissue as a consequence of growth and remodeling: application to an arterial geometry. *European Journal of Mechanics A/Solids*, 27:959–974, 2008.

[PBJ02] Mathias Poppe, Christoph Bourauel, and Andreas JŁger. Determination of the elasticity parameters of the human periodontal ligament and the location of the center of resistance of single-rooted teeth a study of autopsy specimens and their conversion into finite element models. *Journal of Orofacial Orthopedics / Fortschritte der KieferorthopŁdie*, 63(5):358–370, 2002.

[PCB+02] G. Pietrzak, A. Curnier, J. Botsis, S. Scherrer, A. Wiskott, and U. Belser. A nonlinear elastic model of the periodontal ligament and its numerical calibration for the study of tooth mobility. *Comput Methods Biomech Biomed Engin*, 2(5):91–100, Apr 2002.

[PDIS10] L. Placidi, F. Dell'Isola, N. Ianiro, and G. Sciarra. Variational formulation of pre-stressed solid-fluid mixture theory, with an application to wave phenomena. *European Journal of Mechanics*, 27, 2010.

[PP11] Matthias Preisig and Jean H. Prevost. Stabilization procedures in coupled poromechanics problems: A critical assessment. *International Journal for Numerical and Analytical Methods in Geomechanics*, 35(11):1207–1225, 2011.

[Pro00] Christopher G. Provatidis. A comparative fem-study of tooth mobility using isotropic and anisotropic models of the periodontal ligament. *Medical Engineering & Physics*, 22(5):359 – 370, 2000.

[PV11] L. Preziosi and G. Vitale. A multiphase model of tumor and tissue growth including cell adhesion and plastic reorganization. *Mathematical Models and Methods in Applied Sciences*, 21(9):1901–1932, 2011.

[QCK01] Haihong Qian, Jie Chen, and Thomas R. Katona. The influence of {PDL} principal fibers in a 3-dimensional analysis of orthodontic tooth movement. *American Journal of Orthodontics and Dentofacial Orthopedics*, 120(3):272 – 279, 2001.

[QSS00] A. Quarteroni, R. Sacco, and F. Saleri. *Numerical Mathematics*, volume 37 of *Texts in Applied Mathematics*. Springer-Verlag Inc. New York, 2000.

[Qua09] A. Quarteroni. *Numerical modeling for differential problems*. Springer Milan, 4 edition, 2009.

[Qui02] S. Quiligotti. On bulk growth mechanics of solid-fluid mixtures: kinematics and invariance requirements. *Theoretical and Applied Mechanics*, 28–29:277–288, 2002.

[RHD94] E. K. Rodriguez, A. Hoger, and MacCullogh A. D. Stress-dependent finite growth in soft elastic tissues. *Journal of Biomechanics*, 27:455–467, 1994.

[RJ97] J S Rees and P H Jacobsen. Elastic modulus of the periodontal ligament. *Biomaterials*, 18(14):995–9, 1997.

[RKB+05] Alireza Rahimi, Ludger Keilig, G. Bendels, Reinhard Klein, Thorsten M. Buzug, Iman Abdelgader, Marcus Abboud, and Christoph Bourauel. 3d reconstruction of dental specimens from 2d histological images and microct-scans. *Computer Methods in Biomechanics and Biomedical Engineering*, 8(3):167–176, 2005.

[RNL13] C. Ruiz, J. Noailly, and D. Lacroix. Material property discontinuities in intervertebral disc porohyperelastic finite element models generate numerical instabilities due to volumetric strain variations. *Journal of the Mechanical Behavior of Biomedical Materials*, 26(0):1 – 10, 2013.

[Roc70] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, NJ., 1970.

[RS04] K. R. Rajagopal and A. R. Srinivasa. On thermodynamical restrictions of continua. *Proc. R. Soc. Lond. A*, 460:631–651, 2004.

[Sal08] S. Salsa. *Partial differential equations in action*. Springer, 1 edition, 2008.

[Sch03] M. Schanz. On the equivalence of the linear biotÕs theory and the linear theory of porous media. *ASCE Engineering Mechanics Conference*, 2003.

[SH98] J. C. Simo and T. J. R. Hughes. *Compuational Inelasticity*. Springer Verlag, 1998.

[Sho00] R.E. Showalter. Diffusion in poro-elastic media. *Journal of Mathematical Analysis and Applications*, 251(1):310 – 340, 2000.

[SJ06] JC Maltha SH JŮnsdŮttir, EBW Giesen. Biomechanical behaviour of the periodontal ligament of the beagle dog during the first 5 hours of orthodontic force application, 2006.

[SR90] J. C. Simo and M. S. Rifai. A class of mixed assumed strain methods and the method of incompatible modes. *International Journal for Numerical Methods in Engineering*, 29(8):1595–1638, 1990.

[stu] www.studiodentaire.com/en/glossary/periodontal_ligament.php.

[Tab95] L. A. Taber. Biomechanics of growth, remodeling and morphogenesis. *ASME Appl. Mech. Rev.*, 48:487–545, 1995.

[Tal94] Patrick Le Tallec. Numerical methods for nonlinear three-dimensional elasticity. In P. G. Ciarlet and J. L. Lions, editors, *Handbook of Numerical Analysis*, volume 3, pages 465 – 622. Elsevier, 1994.

[TDLE02] R.S. Toms, G.J. Dakin, J.E. Lemons, and A.W. Eberhardt. Quasi-linear viscoelastic behavior of the human periodontal ligament. *Journal of Biomechanics*, 35:1411âĂŞ1415, 2002.

[TEST10] G. Psihoyos T. E. Simos and Ch. Tsitouras, editors. *Growth and Mass Transfer in Multi-Constituent Biological Materials*, volume 1, 2010.

[TN04] C. Truesdell and W. Noll. *The non-linear field theory of mechanics*. Springer, 3 edition, 2004.

[TOS01] Ulrich Trottenberg, C. W. Osterlee, and Anton Schuller. *Multigrid*. Academic Press, Inc., Orlando, FL, USA, 2001.

[TS83] K Tanne and M Sakuda. Initial stress induced in the periodontal tissue at the time of the application of various types of orthodontic force: three-dimensional analysis by means of the finite element method. *J Osaka Univ Dent Sch*, 23:143–71, 1983.

[VF06] M. Vianello and S. Forte. Meccanica dei continui (appunti incompleti). Dipartimento di Matematica, Politecnico di Milano, 2006.

[vT23] K. von Terzaghi. Die berechnung der durchlŁssigkeit des tones aus dem verlauf der hydromechanischen spannungserscheinungen. Sitzungsbericht der Akademie der Wissenschaften, 1923.

[VV81] P. A. Vermeer and A. Verruijt. An accuracy condition for consolidation by finite elements. *International Journal for Numerical and Analytical Methods in Geomechanics*, 5(1):1–14, 1981.

[Wan03] Jing Wan. *Stabilized finite element methods for coupled geomechanics and multiphase flow*. PhD thesis, Stanford University, 2003.

[WE84] K.R. Williams and J.T. Edmundson. Orthodontic tooth movement analysed by the finite element method. *Biomaterials*, 5(6):347 – 351, 1984.