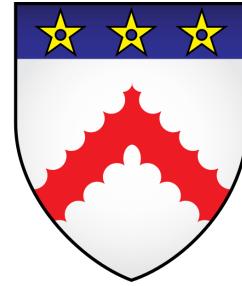
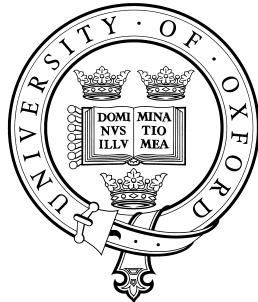


# A finite element method for linear and non-linear poroelasticity: theory and application to lung modelling

Lorenz Berger

Keble College



Department of Computer Science

University of Oxford

Michaelmas Term 2014

This thesis is submitted to the Department of Computer Science,  
University of Oxford, for the degree of Doctor of Philosophy. This thesis is  
entirely my own work, and, except where otherwise indicated, describes my  
own research.

Lorenz Berger  
Keble College

Doctor of Philosophy  
Michaelmas Term 2014

**A finite element method for linear and  
non-linear poroelasticity: theory and  
application to lung modelling**

**Abstract**

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Thesis motivation . . . . .	1
1.2	Thesis goals . . . . .	4
1.3	Thesis structure and contributions . . . . .	5
<b>2</b>	<b>Background</b>	<b>8</b>
2.1	Lung physiology . . . . .	8
2.2	Computational lung models . . . . .	12
2.3	Poroelastic models for lung parenchyma and other biological tissue . . . . .	14
2.4	Finite element methods for poroelasticity . . . . .	15
2.4.1	Linear three-field discretisations . . . . .	16
2.4.2	Methods valid in large deformations . . . . .	18
<b>3</b>	<b>Poroelasticity theory</b>	<b>20</b>
3.1	Kinematics . . . . .	21
3.1.1	Volume fractions . . . . .	25

3.1.2	Balance of mass . . . . .	27
3.1.3	Balance of momentum . . . . .	28
3.1.4	Constitutive relations . . . . .	29
3.1.5	Momentum balance of the mixture . . . . .	31
3.1.6	Momentum balance of the fluid . . . . .	31
3.1.7	Summary of the general poroelasticity model . . . . .	31
3.1.8	Simplification and reformulation of the model . . . . .	34
3.2	Linear poroelasticity . . . . .	35
<b>4</b>	<b>Finite element method</b>	<b>38</b>
4.1	Introduction . . . . .	38
4.2	Functional analysis . . . . .	39
4.3	Model problem . . . . .	41
4.3.1	Weak formulation . . . . .	42
4.3.2	Time discretisation . . . . .	43
4.3.3	Finite element discretisation . . . . .	44
4.4	Mixed methods . . . . .	47
<b>5</b>	<b>Stabilized low-order finite element approximation for linear three-field poroelasticity</b>	<b>52</b>
5.1	Weak formulation . . . . .	52
5.1.1	Continuous weak formulation . . . . .	54
5.2	Fully-discrete model . . . . .	56
5.2.1	Auxiliary results . . . . .	58

5.2.2	Existence and uniqueness of the fully-discrete model at a time step . . . . .	61
5.2.3	Energy estimate of the fully-discrete model . . . . .	67
5.3	A-priori error analysis . . . . .	74
5.3.1	Galerkin orthogonality . . . . .	76
5.3.2	Auxiliary error estimates . . . . .	77
5.4	Numerical Results . . . . .	83
5.4.1	Implementation . . . . .	84
5.4.2	2D convergence study . . . . .	85
5.4.3	3D convergence study . . . . .	88
5.4.4	2D cantilever bracket problem . . . . .	90
5.4.5	3D unconfined compression stress relaxation . . . . .	93
5.5	Conclusion . . . . .	95
<b>6</b>	<b>A stabilized nonlinear finite element method for three-field incompressible poroelasticity valid in large deformations</b>	<b>98</b>
6.0.1	Proposed method . . . . .	98
6.0.2	Overview of the paper . . . . .	100
6.1	The stabilized nonlinear finite element method . . . . .	100
6.1.1	Weak formulation . . . . .	101
6.1.2	Newton's Method . . . . .	102
6.1.3	Discretization of the Newton step . . . . .	105
6.1.4	Newton algorithm . . . . .	108

6.2	Numerical results . . . . .	112
6.2.1	3D unconfined compression stress relaxation . . . . .	113
6.2.2	Terzaghi's problem . . . . .	115
6.2.3	Swelling test . . . . .	119
6.3	Conclusion . . . . .	122
<b>7</b>	<b>A poroelastic model coupled to a fluid network with applications in lung modelling</b>	<b>124</b>
7.1	Model assumptions . . . . .	124
7.1.1	Approximating lung parenchyma using a poroelastic medium . . . . .	125
7.1.2	Approximating the airways using a fluid network model	130
7.2	A mathematical lung model . . . . .	130
7.2.1	A poroelastic model for lung parenchyma . . . . .	131
7.2.2	A network flow model for the airway tree . . . . .	136
7.2.3	Coupling the fluid network to the poroelastic model .	137
7.2.4	Summary of the complete lung model . . . . .	139
7.2.5	Implementation . . . . .	140
7.3	Model generation . . . . .	142
7.3.1	Mesh generation . . . . .	142
7.3.2	Reference state, boundary conditions and initial conditions . . . . .	144
7.3.3	Simulation parameters . . . . .	145

7.4	Model exploration . . . . .	146
7.4.1	Normal breathing . . . . .	147
7.4.2	Breathing with airway constriction . . . . .	153
7.4.3	Breathing with tissue weakening . . . . .	157
7.5	Discussion . . . . .	157
7.5.1	Contributors of airway resistance and tissue mechanics to lung function . . . . .	158
7.5.2	Limitations and future work . . . . .	159
7.5.3	Summary . . . . .	161
7.6	Appendix . . . . .	162
7.6.1	Finite element matrices . . . . .	162
7.6.2	Matrix voigt notation . . . . .	164
7.6.3	Neo-Hookean example . . . . .	165
<b>8</b>	<b>Conclusion</b>	<b>167</b>
<b>A</b>		<b>168</b>
A.1	Tangent modulus . . . . .	168
A.2	Matrix voigt notation . . . . .	169
A.2.1	Neo-Hookean example . . . . .	170

## **Acknowledgements**

Cheers.

# Chapter 1

## Introduction

### 1.1 Thesis motivation

The main function of the lungs is to exchange gas between air and blood, supplying oxygen during inspiration and removing carbon dioxide by subsequent expiration. Gas exchange is optimised by ensuring efficient matching between ventilation and blood flow, the distributions of which are largely governed by tissue deformation, gravity and branching structure of the airway and vascular trees. In this work, we focus on the link between tissue deformation and ventilation. Previous work has typically focused on modelling either ventilation or tissue deformation in isolation. However evaluation of each component (i.e. tissue deformation and ventilation) separately does not necessarily give accurate ventilation predictions or provide a good indication of how the integrated organ works, this is because both components are inter-

dependent. To gain a better understanding of the biomechanics in the lung it is therefore necessary to fully couple the tissue deformation with the ventilation. To achieve this tight coupling between the tissue deformation and the ventilation we propose a novel multiscale model that approximates the lung parenchyma by a biphasic (tissue and air, ignoring blood) poroelastic model, that is then coupled to an airway fluid network model.

An integrated model of ventilation and tissue mechanics will be particularly important for understanding respiratory diseases since nearly all pulmonary diseases lead to some abnormality of lung tissue mechanics Suki and Bates [2011]. For example, chronic obstructive pulmonary disease (COPD) encompasses emphysema (destruction of alveolar tissue) and chronic bronchitis which can cause severe, airway remodeling, bronchoconstriction and air trapping, all of which can significantly alter tissue properties. If the tissue mechanics are affected so too will the ventilation and vice versa, again emphasising the importance of a model that fully couples the ventilation and tissue mechanics in the lung. The impact of alterations during disease, such as airway narrowing or changes in tissue properties, on regional ventilation and tissue stresses are not well understood. For example, one hypothesis is that airway disease may precede emphysema Galbán et al. [2012]. The computational lung model could be applied to investigate the impact of airway narrowing and tissue stiffness during obstructive lung diseases on tissue stresses, alveoli pressure and ventilation.

Developing such a fully coupled model has to our knowledge not yet

been achieved. There are many difficulties involved in creating a model that is physiologically accurate and can be solved numerically. We will need to develop methodology techniques to solve the poroelastic equations, and develop solution techniques to couple the poroelastic model to the airway fluid network model.

In particular, in the diseased lung, abrupt changes in tissue properties and heterogeneous airway narrowing are possible. This can result in a patchy ventilation and pressure distribution [Venegas et al., 2005]. In this situation existing methods that solve the poroelastic equations using a continuous pressure approximation would struggle to capture the steep gradients in pressure, and result in localized oscillations in the pressure. By developing a method that utilises a discontinuous approximation for the pressure we will be able to approximate these steep pressure gradients reliably, and avoid localized oscillations in the pressure.

The proposed methodology could also be adapted to model other biological tissues where blood vessels flow through and interact with a deforming tissue. For example, when modelling perfusion of blood flow in the beating myocardium [Chapelle et al., 2010, Cookson et al., 2012], modelling brain oedema [Li et al., 2010] or hydrocephalus [Wirth and Sobey, 2006], or microcirculation of blood and interstitial fluid in the liver lobule [Leungchavaphongse, 2013]. In addition to this poroelasticity theory has also been used in various geomechanical applications ranging from reservoir engineering [Phillips and Wheeler, 2007a] to modelling earthquake fault zones

[White and Borja, 2008]. The theory developed in this thesis could be applied in these fields.

## 1.2 Thesis goals

The main goal of this is to rigorously develop a finite element method for solving the poroelastic equations, and then use this methodology to simulate the lung breathing on a realistic geometry. More specific targets are:

1. Develop a low-order finite element method for solving the linear poroelastic equations using a discontinuous pressure approximation. Prove theoretical results about the discretisation, including existence and uniqueness, an energy estimate and an optimal a-priori error estimate.
2. Extend the method to a non-linear finite element method to solve the poroelastic equations valid in large deformations.
3. Rigorously test the method using numerous test problems to verify theoretical stability and convergence results, and its ability to reliably capture steep pressure gradients.
4. Derive a poroelastic model for lung parenchyma coupled to an airway fluid network model, and develop a stable method to numerically solve the coupled model.
5. Solve the computational lung model on a realistic geometry, with realistic boundary conditions extracted from imaging data, to simulate

breathing. Evaluate the effect of tissue weakening and airway narrowing on lung function.

### 1.3 Thesis structure and contributions

The contributions of each chapter to the thesis are as follows:

**Chapter 2:** We give a brief overview of lung physiology, review the literature on ventilation models and existing poroelastic models, and discuss numerical methods currently available to solve the poroelastic equations.

**Chapter 3:** We introduce the general theory of poroelasticity valid in large deformations, and derive the linear poroelastic equations, valid in small deformations.

**Chapter 4:** We outline the basic concepts of the standard continuous Galerkin finite element method. We then discuss mixed problems and their stability requirement.

**Chapter 5:** A stabilized conforming finite element method for the linear three-field (displacement, fluid flux and pressure) poroelasticity problem is presented. By applying a local pressure jump stabilization term to the mass conservation equation we avoid pressure oscillations. For the fully discretized

problem we prove existence and uniqueness, an energy estimate and an optimal a-priori error estimate. Numerical experiments in 2D and 3D illustrate the convergence of the method, show the effectiveness of the method to overcome spurious pressure oscillations, and evaluate the added mass effect of the stabilization term.

**Chapter 6:** We apply the method developed in Chapter 5 to solve the three-field nonlinear quasi-static incompressible poroelasticity problem valid in large deformations. We present the linearization and discretisation of the equations, and give a detailed account of the implementation. Numerical experiments in 3D verify the method and illustrate its ability to reliably capture steep pressure gradients.

**Chapter 7:** We present the model assumptions required for the proposed poroelastic lung model and outline its mathematical formulation and coupling to the airway fluid network. A finite element method is presented to discretize the equations in a monolithic way to ensure convergence of the nonlinear problem. Finally, numerical simulations on a realistic lung geometry that illustrate the coupling between the poroelastic medium and the network flow model are presented. Numerical simulations of tidal breathing are shown to reproduce global physiological realistic measurements. We also investigate the effect of airway constriction and tissue weakening on the ventilation, stress and alveolar pressure distribution.

**Chapter 8:** We summarise the main results and propose future lines of research.

# Chapter 2

## Background

### 2.1 Lung physiology

We will now give a basic review of lung physiology. A more complete introduction can be found in West [2008].

During inspiration, the volume of the thoracic cavity increases and air is drawn into the lung by creating a sub-atmospheric pressure. The increase in volume is brought about mainly by contraction of the diaphragm, which causes it to descend, and partly by the action of the intercostal muscles, which raise the ribs. The lung is elastic and returns passively to its preinspiratory volume during resting breathing. During expiration the intra-alveolar pressure becomes slightly higher than atmospheric pressure and gas flows out of the lungs West [2008]. The pressure required to move gas through the airways in a healthy lung is very small. During normal inspiration, an air

flow rate of 1 liter per second requires a pressure drop along the airways of less than 200 Pa. Compare this to a smokers pipe, which needs a pressure of about 50,000 Pa for the same flow rate West [2008].

The airway tree is divided into a conducting zone and a respiratory zone. The trachea divides into right and left main bronchi, which in turn divide into lobar and then segmental bronchi. This process continues down to the terminal bronchioles, which are the smallest airways without alveoli. All of these bronchi make up the conducting airways. The terminal bronchioles, which appear at around generation 15-16, then continue to divide into respiratory bronchioles, which have occasional alveoli budding from their walls. Finally, we get to the alveolar ducts, which are completely lined with alveoli, see Figure 2.1b. This alveolated region of the lung where the gas exchange occurs is known as the respiratory zone West [2008]. Table 2.1 documents the different flow characteristics found in the airway tree during slow and rapid breathing.

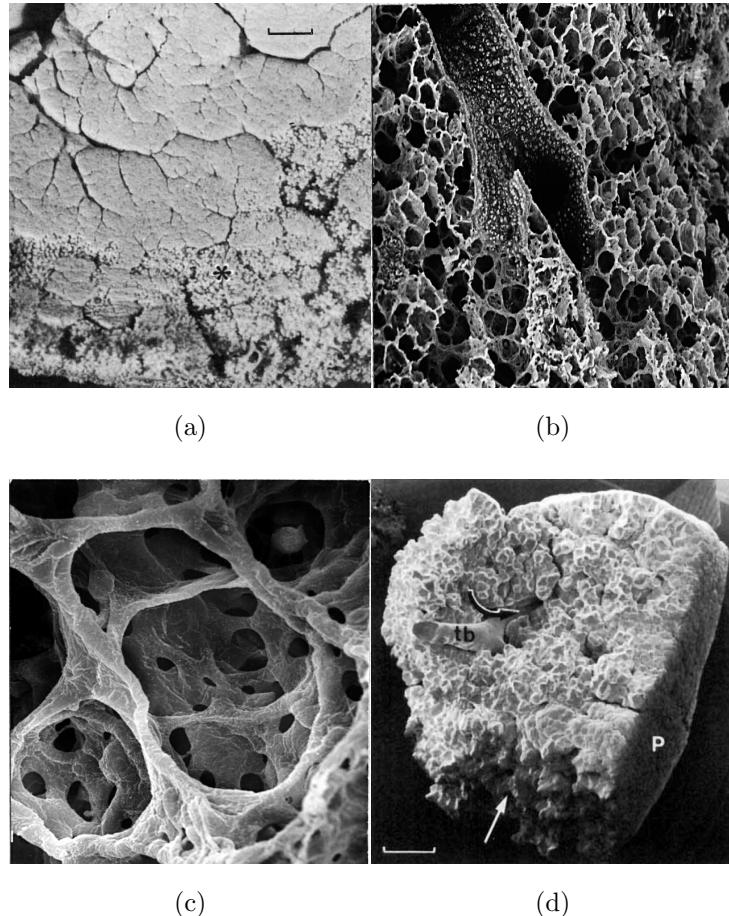


Figure 2.1: (a) Portions of silicone rubber casts of upper lobes of human lungs; asterisk marks incompletely filled regions. The outline of individual unfilled acinar units can also be seen. Scale marker, 5 mm. (b) Transition from terminal bronchiole to alveolar duct, from conducting airway to oxygen transfer area, diameter of terminal bronchiole is 0.5 mm. (c) A few alveoli in an alveolar duct. The dark round openings are pores between alveoli. The alveolar wall is quite thin and contains a network of capillaries. The average diameter of one alveoli is 0.2 mm. (d) Scanning electron micrograph of complete acinus with transitional bronchiole (tb) and surface abutting on pleura (P). Note the irregular surface where alveolar sacs of adjacent acini interdigitate (straight arrow). Scale marker, 1 mm. Images are reproduced from Lawrence Berkeley National Laboratory [1995].

Generation	Diameter cm	Length cm	Flow rate 10L/min Velocity (m/s)	Re	Flow rate 100L/min Velocity (m/s)	Re
Trachea	1.80	12.0	65.8	775	658	7750
1	1.22	4.76	71.6	573	716	5730
5	0.35	1.07	53.6	123	536	1230
10	0.13	0.46	12.55	10.6	125	106
15	0.066	0.20	1.48	0.63	14.8	6.30
20	0.045	0.083	0.10	0.031	1.00	0.31

Table 2.1: Shows dimensions, velocity and the corresponding Reynolds number for different sections of the airway tree during slow and rapid breathing. These values have been taken from Pedley et al. [1970].

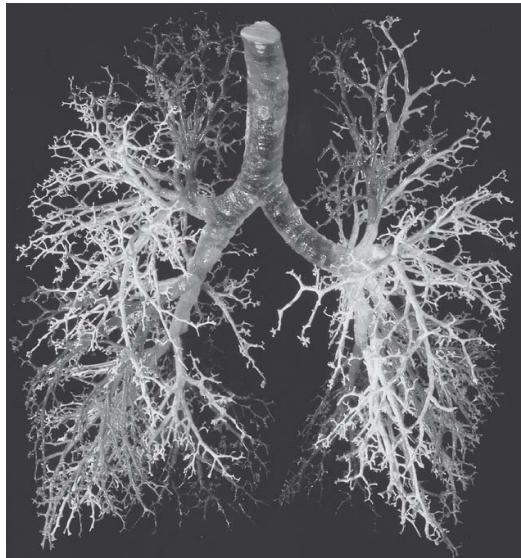


Figure 2.2: A rubber cast of the conducting of a human lung. The image is reproduced from West [2008].

Lung parenchyma refers to the portion of the lung made up of the small air chambers (alveoli) participating in gas exchange. The alveoli are made up of collagen, elastin fibers and membranous structures containing the capillary network, see Figure 2.1c. Alveoli are arranged in sponge like structures

and fill the entire volume of the lungs surrounding the conducting passages. Figure 2.1a shows a rubber cast of lung parenchyma, the dark lines outline the branching structure of the airways. The right and left lung are partitioned into three and two lobes, respectively. Lung segments of conic shape are then the first subdivision of these lobes. These structures are bounded by connective tissue such that surgical separation is often possible. In the right lung, there are usually ten segments whereas only nine can be found in the left lung. Within the segments, the bronchi branch about six to twelve times. The terminal bronchioles which appear after roughly 15–16 branching generations then finally feed into approximately 30,000 so-called acini, see Figure 2.1d. These acini represent the largest lung units of which all airways are alveolated and thus participate in gas exchange Weichert [2011].

## 2.2 Computational lung models

There exist a large number of computational ventilation and deformation models for the lung. Some models are designed to model particular phenomena whilst others are more general. They also range in spatial complexity from 0D compartment type models to 3D models which are able to incorporate ‘patient-specific’ geometries extracted from CT images. In this review, we will focus on models that couple ventilation with tissue deformation and can be used as patient-specific models. A review of popular lung models can be found in Bates [2009]. One study that couples ventilation and tis-

sue deformation using a one way coupling approach and then applies the model to a full 3D geometry is described in Tawhai and Lin [2010]. Here a mechanics model for elastic deformation of compressible lung tissue is used to provide flow and pressure boundary conditions for an embedded airway model which makes the resultant ventilation distribution dependent on the tissue deformation due to gravity. In Swan et al. [2012], air flow is simulated in patient specific conducting airways which are coupled to geometrically simplified terminal acinar units with varying volume dependent compliances. The fluid flow in the airways is approximated by Poiseuille flow with an added correction term for airway bifurcations. The end terminal acinar units are able to expand but are assumed to be independent of neighbouring acinar units. This does not allow for feedback from neighbouring acini that are infact tightly connected by a matrix of fibers, collagen and capillaries. Other sophisticated flow models of the whole airway tree, as previously mentioned, also exist [Ismail et al., 2013, Yin et al., 2013]. These models solve the full 3D Navier-Stokes equations in the upper airways, segmented from CT images, to capture high Reynolds number effects. The 3D fluid model is then coupled to a 0D laminar flow model of the lower airways.

## **2.3 Poroelastic models for lung parenchyma and other biological tissue**

Some early work on a mechanical model of lung parenchyma as a poroelastic medium has already been proposed in Kowalczyk [1993]. This work developed a similar poroelastic model to the one we propose, however it has only been applied to a very simple 2D geometry. Also in Owen and Lewis [2001] homogenisation theory has been used to derive macroscopic poroelastic equations for average air flows and tissue displacements in lung parenchyma during high frequency ventilation. The resulting model is a one dimensional system of equations that is used to investigate the effect of high-frequency ventilation on strain in the parenchymal tissue. The use of a poroelastic model has also been applied to modelling other biological tissues. For example modelling protein based hydrogels embedded with cells Galie et al. [2011], perfusion of blood flow in the beating myocardium Chapelle et al. [2010], Cookson et al. [2012], the modelling of brain oedema (swelling) Li et al. [2010] and hydrocephalus Wirth and Sobey [2006]. Another application is the modelling of interstitial fluid and tissue in articular cartilage and intervertebral discs Mow et al. [1980], Holmes and Mow [1990], Galbusera et al. [2011].

## 2.4 Finite element methods for poroelasticity

The method that we use for spatially discretising our equations in this work is the finite element method (FEM) for obtaining approximations to the solution of partial differential equations.

After many decades of research there remain numerous challenges associated with the numerical solution of the poroelastic equations. When using the finite element method the main challenge is to ensure stability and convergence of the method and prevent numerical instabilities that often manifest themselves in the form of spurious oscillations in the pressure. It has been suggested that this problem is caused by the saddle point structure in the coupled equations resulting in a violation of the famous Ladyzhenskaya-Babuska-Brezzi (LBB) condition [Haga et al., 2012], highlighting the need for a stable combination of mixed finite elements. Another numerical challenge in practical 3D applications is the algebraic system arising from the finite element discretisation. This can lead to a very large matrix system that has many unknowns and is severely ill-conditioned, making it difficult to solve using standard iterative solvers. Therefore low-order finite element methods that allow for efficient preconditioning are preferred [Ferronato et al., 2010, White and Borja, 2011].

### 2.4.1 Linear three-field discretisations

The poroelastic equations are often solved in a reduced displacement and pressure formulation, from which the fluid flux can then be recovered [Murad and Loula, 1994, White and Borja, 2008]. Murad and Loula [1994] have analysed the stability and convergence of this reduced displacement pressure ( $\mathbf{u}/p$ ) formulation and were able to show error bounds for inf-sup stable combinations of finite element spaces (e.g. Taylor-Hood elements). In this paper we will keep the fluid flux variable resulting in a three-field, displacement, fluid flux, and pressure formulation. Keeping the fluid flux as a primary variable has the following advantages:

- i It avoids the calculation of the fluid flux in post-processing.
- ii Physically meaningful boundary conditions can be applied at the interface when modelling the interaction between a fluid and a poroelastic structure [Badia et al., 2009].
- iii It allows for greater accuracy in the fluid velocity field. This can be of interest whenever a consolidation model is coupled with an advection diffusion equation, e.g. to account for thermal effects, contaminant transport or the transport of nutrients or drugs within a porous tissue [Khaled and Vafai, 2003].
- iv It allows for an easy extension of the fluid model from a Darcy to a Brinkman flow model, for which there are numerous applications in

modelling biological tissues [Khaled and Vafai, 2003].

- v It reduces the order of the spatial derivative of the pressure, allowing for a discontinuous pressure approximation without any additional penalty terms.

Phillips and Wheeler [2007a,b], have proven error estimates when solving the three-field formulation problem using continuous piecewise linear approximations for displacements and mixed low-order Raviart Thomas elements for the fluid flux and pressure variables. However this method was found to be susceptible to spurious pressure oscillations [see, Phillips and Wheeler, 2009]. To overcome these pressure oscillations, Li and Li [2012] analysed a discontinuous three-field method, and Yi [2013] analysed a nonconforming three-field method. In addition to these monolithic approaches there has been considerable work on operating splitting (iterative) approaches for solving the poroelastic equations [Feng and He, 2010, Kim et al., 2011, Wheeler and Gai, 2007]. Although these methods are often able to take advantage of existing elasticity and fluid finite element software, and result in solving a smaller system of equations, these schemes are often only conditionally stable. To ensure that the method is unconditionally stable, monolithic approaches are often preferred. The method proposed in this work is monolithic, and will therefore retain the advantage of being unconditionally stable .

### 2.4.2 Methods valid in large deformations

We will now give a brief overview of different approaches for solving the poroelastic equations valid in large deformations. There has been some work on operating splitting (iterative) approaches where the poroelastic equations are separated into a fluid problem and deformation problem [Chapelle et al., 2010]. Again, this approach is only conditionally stable. Some notable quasi-static incompressible large deformation monolithic approaches include a mixed-penalty formulation, and a mixed solid velocity-pressure formulation, both outlined in [Almeida and Spilker, 1998], the solid velocity-pressure formulation is similar to the commonly used reduced ( $\mathbf{u}/p$ ) formulation [Ateshian et al., 2010]. These two-field ( $\mathbf{u}/p$ ) formulations require a stable mixed element pair such as the popular Taylor-Hood element to satisfy the LBB inf-sup stability requirement. To reduce the number of unknowns, and allow for an equal-order, piecewise linear approximation, a stabilized reduced ( $\mathbf{u}/p$ ) formulation has been proposed in [White and Borja, 2008]. This method introduces a stabilization term to the mass conservation equation to overcome the spurious pressure oscillations. The key difficulty, however, that this stabilized element cannot escape is that jumps in material coefficients may introduce large solution gradients across the interface, requiring severe mesh refinement. This is because a continuous pressure element is used, which is unable to reliably capture jumps in the pressure solution [White and Borja, 2008]. In [Levenston et al., 1998] a three-field (displacement, fluid flux, pressure) formulation has been outlined, however this method uses

a low-order mixed finite element approximation without any stabilization and therefore is not inf-sup stable.

# Chapter 3

## Poroelasticity theory

We will now develop the theory required for modelling of the complete porous medium made up of a solid and fluid phase. There exist two main approaches for modelling a deformable porous medium. There is the mixture theory approach also known as the Theory of Porous Media (TPM) [bow, Boer, 2005, Bowen, 1980], which has its roots in the classical theories of gas mixtures, and makes use of a volume fraction concept, where the porous medium is represented by spatially superposed interacting media. The other approach is purely macroscopic and is mainly associated with the work of Biot, a detailed description can be found in the book by Coussy Coussy [2004]. For a good comparison between the two theories see Coussy et al. [1998]. In this work we will use the mixture theory approach, outlined in the book by R. Boer Boer [2005], to derive the equations since this is the common approach taken for applications of poroelasticity in biology.

### 3.1 Kinematics

Within the theory of continuum mixture theory, a poroelastic medium is treated as the superposition of two interacting continua simultaneously occupying the same physical space. The superscript  $\alpha \in s, f$  denotes a quantity related to the solid or fluid. Before presenting the mixture theory, we give a review of solid mechanics. This will form the basis of the description of the solid skeleton. The following review of continuum mechanics closely follows chapter 4 in Gonzalez and Stuart [2008], and the standard Poromechanics book by Coussy [2004]. Most of the kinematic quantities described here will be associated with the solid skeleton, since this also describes the motion of the fluid domain.

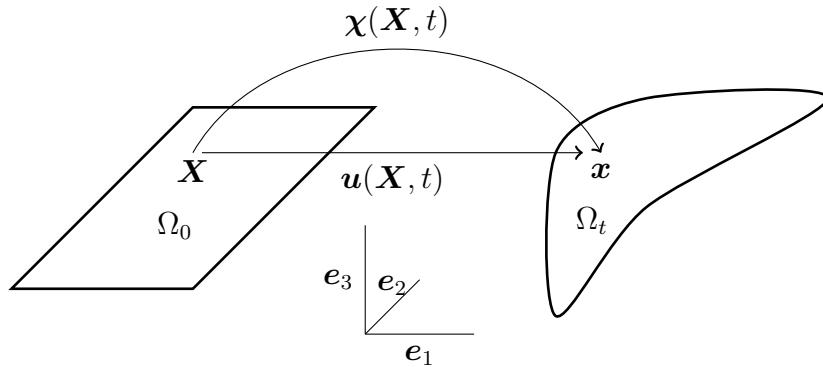


Figure 3.1: Illustration of the solid deformation.

Let the volume  $\Omega_0$  be the undeformed Lagrangian (material) reference configuration and let  $\mathbf{X} = \{X\mathbf{e}_1 + Y\mathbf{e}_2 + Z\mathbf{e}_3\}$  indicate the position of a solid particle in  $\Omega_0$  at  $t = 0$ , where  $X, Y$  and  $Z$  are the components of the

position with respect to the standard orthonormal basis  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$  for  $\mathbb{R}^3$ . The position of a solid particle in the current Eulerian (spatial) configuration  $\Omega_t$  is given by  $\mathbf{x} = \{x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3\}$ , with  $\mathbf{x} = \chi(\mathbf{X}, t)$ , shown in Figure 3.1. The deformation map,  $\chi(\mathbf{X}, t)$ , is a continuously differentiable, invertible mapping from  $\Omega_0$  to  $\Omega_t$ . Thus the inverse of the deformation map,  $\chi^{-1}(\mathbf{x}, t)$ , is such that  $\mathbf{X} = \chi^{-1}(\mathbf{x}, t)$ . The displacement field is given by

$$\mathbf{u}(\mathbf{X}, t) = \chi(\mathbf{X}, t) - \mathbf{X}. \quad (3.1)$$

The deformation gradient tensor,

$$\mathbf{F} = \frac{\partial \chi(\mathbf{X}, t)}{\partial \mathbf{X}} = \mathbf{I} + \frac{\partial \mathbf{u}(\mathbf{X}, t)}{\partial \mathbf{X}}, \quad (3.2)$$

maps a material line element in the reference configuration  $d\mathbf{X}$ , to a line element  $d\mathbf{x}$  in the current configuration, i.e.  $d\mathbf{x} = \mathbf{F}d\mathbf{X}$ . The symmetric right Cauchy-Green deformation tensor is given by

$$\mathbf{C} = \mathbf{F}^T \mathbf{F}. \quad (3.3)$$

The jacobian is defined as

$$J = \det(\mathbf{F}), \quad (3.4)$$

and represents the change in an infinitesimal small volume from the reference to the current configuration. Also note that  $J > 0$ , to avoid self penetration

of the body. We also have that  $J$  represents the change in an infinitesimal small volume from a reference volume element  $d\Omega_0$  to a current configuration volume element  $d\Omega_t$

$$d\Omega_t = J d\Omega_0. \quad (3.5)$$

Also,  $\mathbf{F}$  is invertible, and it is easy to see that the inverse of the deformation gradient is the deformation gradient of the inverse map

$$\mathbf{F}^{-1} = \frac{\partial \boldsymbol{\chi}^{-1}(\mathbf{x}, t)}{\partial \mathbf{x}} = \frac{\partial \mathbf{X}}{\partial \mathbf{x}}. \quad (3.6)$$

We denote by  $\mathbf{V}(\mathbf{X}, t)$  the velocity at time  $t$  of the material (fixed) solid particle  $\mathbf{X}$ . By definition we have

$$\mathbf{V}(\mathbf{X}, t) = \frac{\partial}{\partial t} \boldsymbol{\chi}(\mathbf{X}, t). \quad (3.7)$$

Similarly, we denote by  $\mathbf{A}(\mathbf{X}, t)$  the acceleration, given by

$$\mathbf{A}(\mathbf{X}, t) = \frac{\partial^2}{\partial t^2} \boldsymbol{\chi}(\mathbf{X}, t). \quad (3.8)$$

We see that the velocity and acceleration of material particles are material fields. Also note that  $\frac{\partial}{\partial t} \mathbf{u}(\mathbf{X}, t) = \frac{\partial}{\partial t} \boldsymbol{\chi}(\mathbf{X}, t)$ . We will also require a spatial description of these fields. We denote by  $\mathbf{v}^s(\mathbf{x}, t)$  the spatial description of

the material solid velocity field, such that

$$\mathbf{v}^s(\mathbf{x}, t) = \left[ \frac{\partial}{\partial t} \boldsymbol{\chi}(\mathbf{X}, t) \right]_{\mathbf{X}=\boldsymbol{\chi}^{-1}(\mathbf{x}, t)}. \quad (3.9)$$

Similarly, for the spatial description of the solid acceleration, we have

$$\mathbf{a}^s(\mathbf{x}, t) = \left[ \frac{\partial^2}{\partial t^2} \boldsymbol{\chi}(\mathbf{X}, t) \right]_{\mathbf{X}=\boldsymbol{\chi}^{-1}(\mathbf{x}, t)}. \quad (3.10)$$

Notice that  $\mathbf{v}^s(\mathbf{x}, t)$  and  $\mathbf{a}^s(\mathbf{x}, t)$  correspond to the velocity and acceleration of the solid material particle whose current coordinates are  $\mathbf{x}$  at time  $t$ . Also note that by definition of  $\mathbf{v}^s$  in (3.9) we have (see [Gonzalez and Stuart, 2008, page 132])

$$\mathbf{v}^s(\mathbf{x}, t)|_{\mathbf{x}=\boldsymbol{\chi}(\mathbf{X}, t)} = \frac{\partial}{\partial t} \boldsymbol{\chi}(\mathbf{X}, t). \quad (3.11)$$

The **particle derivative of a field**  $\mathcal{G}(\mathbf{x}, t)$  with respect to the particle  $\alpha$  ( $s$  or  $f$ ) is given by (see [Coussy, 2004, eqn. (1.43)])

$$\frac{d^\alpha}{dt} \mathcal{G} = \frac{\partial \mathcal{G}}{\partial t} + (\nabla \mathcal{G}) \mathbf{v}^\alpha, \quad (3.12)$$

where  $\nabla(\cdot) = \frac{\partial(\cdot)}{\partial \mathbf{x}}$  denotes the partial derivative with respect to  $\mathbf{x}$ . Since nearly all our workings will be performed in the current configuration we will keep the shorthand notation  $\nabla$  to denote the spatial gradient in the current configuration instead of explicitly writing  $\nabla_{\mathbf{x}}$ . The **particle derivative of a material volume** with respect to the  $\alpha$ -constituent is given by (see [Coussy,

2004, eqn. (1.42)])

$$\frac{d^\alpha}{dt}(d\Omega_t) = (\nabla \cdot \mathbf{v}^f) d\Omega_t. \quad (3.13)$$

The particle derivative also applies to a volume integral. Thus, for any quantity  $\mathcal{G}$ , associated with the  $\alpha$  constituent, we have (see [Coussy, 2004, eqn. (1.47)])

$$\frac{d^\alpha}{dt} \int_{\Omega_t} \mathcal{G} d\Omega_t = \int_{\Omega_t} \frac{d^\alpha}{dt}(\mathcal{G} d\Omega_t) = \int_{\Omega_t} \left( \frac{d^\alpha \mathcal{G}}{dt} + \mathcal{G} \nabla \cdot \mathbf{v}^\alpha \right) d\Omega_t = \int_{\Omega_t} \left( \frac{\partial \mathcal{G}}{\partial t} + \nabla \cdot \mathcal{G} \mathbf{v}^\alpha \right) d\Omega_t. \quad (3.14)$$

This is commonly known as the Reynolds transport theorem, in the last step we have used the identity  $\nabla \cdot (\psi \mathbf{s}) = \mathbf{s} \cdot \nabla \psi + \psi \nabla \cdot \mathbf{s}$  for some scalar  $\psi$  and vector  $\mathbf{s}$ .

### 3.1.1 Volume fractions

The porous medium is assumed to be saturated consisting of a solid part and a fluid part. The fluid part accounts for volume fractions  $\phi_0(\mathbf{X}, t = 0)$  and  $\phi(\mathbf{x}, t)$  of the total volume in the reference and the current (deformed) configuration. The fractions for the solid are thus given by  $1 - \phi_0$  and  $1 - \phi$  in the reference and the current configuration respectively,  $\phi$  is also often known as the porosity. In standard elasticity theory  $\rho$  denotes the density of the solid, and for an incompressible solid we have that  $\rho d\Omega_t = \rho_0 d\Omega_0$ . We now change the definition of  $\rho$  to be the density of the mixture in the current

configuration which is given by

$$\rho = \rho^s(1 - \phi) + \rho^f\phi \quad \text{in } \Omega_t, \quad (3.15)$$

where  $\rho^s$  and  $\rho^f$  are the densities of the fluid and solid, respectively. We will also assume that both the solid and the fluid are incompressible, such that  $\rho^s = \rho_0^s$  and  $\rho^f = \rho_0^f$ . For notational convenience we also define

$$\hat{\rho}^s = \rho^s(1 - \phi), \quad (3.16)$$

and

$$\hat{\rho}^f = \rho^f\phi. \quad (3.17)$$

Due to mass conservation and the incompressibility of both the solid and the fluid phase we have

$$J = \frac{1 - \phi_0}{1 - \phi}. \quad (3.18)$$

Here  $J$  represents the change in volume of the solid skeleton. The solid skeleton includes the solid (tissue in biological applications) and the voids are occupied by the fluid. Note that although the solid is assumed to be incompressible the solid skeleton is able to change in volume, since fluid can enter or leave the solid skeleton.

### 3.1.2 Balance of mass

When no mass change occurs, neither for the solid skeleton or the fluid contained in  $\Omega_t$ , using the Reynolds transport theorem (3.14), the mass balance can be expressed as

$$\begin{aligned} \frac{d^s}{dt} \int_{\Omega_t} (1 - \phi) \rho^s d\Omega_t &= \int_{\Omega_t} \left( \frac{\partial(1 - \phi)\rho^s}{\partial t} + \nabla \cdot ((1 - \phi)\rho^s \mathbf{v}^s) \right) d\Omega_t, \\ \frac{d^f}{dt} \int_{\Omega_t} \phi \rho^f d\Omega_t &= \int_{\Omega_t} \left( \frac{\partial \phi \rho^f}{\partial t} + \nabla \cdot (\phi \rho^f \mathbf{v}^f) \right) d\Omega_t. \end{aligned}$$

Thus, the balance of mass for the solid is given by (see Boer [2005, eqn. (8.57)])

$$\frac{\partial(1 - \phi)\rho^s}{\partial t} + \nabla \cdot ((1 - \phi)\rho^s \mathbf{v}^s) = 0 \quad \text{in } \Omega_t, \quad (3.19)$$

where  $\mathbf{v}^s$  is the velocity vector of the solid. Similarly, the balance of mass for the fluid is given by

$$\frac{\partial \phi \rho^f}{\partial t} + \nabla \cdot (\phi \rho^f \mathbf{v}^f) = g \quad \text{in } \Omega_t, \quad (3.20)$$

where  $\mathbf{v}^f$  is the velocity vector of the fluid and  $g$  is a general source or sink term. Adding (3.19) and (3.20), and noting that  $\rho^s$  and  $\rho^f$  are assumed to be constants, we arrive at the continuity or mass balance equation of the mixture (see Boer [2005, eqn. (8.59)])

$$\nabla \cdot ((1 - \phi)\mathbf{v}^s) + \nabla \cdot (\phi \mathbf{v}^f) = g \quad \text{in } \Omega_t. \quad (3.21)$$

### 3.1.3 Balance of momentum

The balance law of linear momentum for each individual constituent is given by

$$\frac{d^\alpha}{dt} \int_{\Omega_t} \hat{\rho}^\alpha \mathbf{v}^\alpha d\Omega_t = \int_{\Omega_t} \nabla \cdot \boldsymbol{\sigma}^\alpha + \hat{\rho}^\alpha \mathbf{f} + \hat{\mathbf{p}}^\alpha d\Omega_t. \quad (3.22)$$

Here  $\boldsymbol{\sigma}^\alpha$  is the Cauchy stress tensor of the  $\alpha$  constituent,  $\mathbf{f}$  is a volume force acting on the constituents, and  $\hat{\mathbf{p}}^\alpha$  are interaction forces representing frictional interactions between the solid and fluid, defined later in section 7.2.1. Using the Reynolds transport theorem (3.14), and the balance of mass of the solid (3.19) and fluid (3.20), we get (see Boer [2005, eqn. (4.15)])

$$\nabla \cdot \boldsymbol{\sigma}^\alpha + \hat{\rho}^\alpha \mathbf{f} + \hat{\mathbf{p}}^\alpha = \hat{\rho}^\alpha \mathbf{a}^\alpha + \mathbf{v}^\alpha \left( \frac{d^\alpha \hat{\rho}^\alpha}{dt} + \hat{\rho}^\alpha \nabla \cdot \mathbf{v}^\alpha \right) \quad \text{in } \Omega_t, \quad (3.23)$$

where  $\mathbf{a}^\alpha$  are acceleration vectors of the constituents. The acceleration of the solid and fluid are given by (see Boer [2005, eqn. (3.7,3.9)]),

$$\mathbf{a}^f = \frac{d^f \mathbf{v}^f}{dt} = \frac{\partial}{\partial t} \mathbf{v}^f + (\nabla \mathbf{v}^f) \mathbf{v}^f. \quad (3.24)$$

$$\mathbf{a}^s(\mathbf{x}, t)|_{\mathbf{x}=\chi(\mathbf{X}, t)} = \frac{\partial^2}{\partial t^2} \boldsymbol{\chi}(\mathbf{X}, t). \quad (3.25)$$

We also have that by the action reaction law, each constituent exerts an opposite interaction force on the other constituent

$$\hat{\mathbf{p}}^s + \hat{\mathbf{p}}^f = 0. \quad (3.26)$$

### 3.1.4 Constitutive relations

The interaction force is given by (see Coussy [2004, eqn. (3.49)])

$$\hat{\mathbf{p}}^s = -\hat{\mathbf{p}}^f = -p\nabla\phi + \phi^2\boldsymbol{\kappa}^{-1} \cdot (\mathbf{v}^f - \mathbf{v}^s), \quad (3.27)$$

where  $\boldsymbol{\kappa}$  is the (dynamic) permeability tensor, which includes the dynamic viscosity of the fluid i.e.  $\boldsymbol{\kappa} = \mu_f^{-1}\mathbf{k}$ , where  $\mu_f$  and  $\mathbf{k}$  are the fluid viscosity and the intrinsic permeability tensor, respectively. The fluid pressure is denoted by  $p$ . According to Coussy [2004], the first term,  $p\nabla\phi$ , accounts for the pressure effect resulting from the variation of the section offered to the fluid flow, and the second term,  $\phi^2\boldsymbol{\kappa}^{-1} \cdot (\mathbf{v}^f - \mathbf{v}^s)$ , describes the viscous resistance opposed by the shear stress to the fluid flow from the drag at the internal walls of the porous network. This particular choice for the interaction force means that the momentum balance for the fluid flow can later be reduced to the well known Darcy law. The permeability tensor in the current configuration is given by

$$\boldsymbol{\kappa} = J^{-1}\mathbf{F}\boldsymbol{\kappa}_0(\chi)\mathbf{F}^T, \quad (3.28)$$

where  $\boldsymbol{\kappa}_0(\chi)$  is the permeability in the reference configuration, which may be chosen to be some (nonlinear) function dependent on the deformation. Examples of deformation dependent permeability tensors for biological tissues can be found in [Holmes and Mow, 1990, Kowalczyk and Kleiber, 1994, Lai and Mow, 1980]. The solid stress tensor is given by the effective stress

principle (see eqn. (8.62) in Boer [2005]),

$$\boldsymbol{\sigma}^s = \boldsymbol{\sigma}_e^s - (1 - \phi)\mathbf{I}p, \quad (3.29)$$

where  $\boldsymbol{\sigma}_e^s$  is the effective stress tensor given by

$$\boldsymbol{\sigma}_e^s = \frac{1}{J}\mathbf{F} \cdot 2\frac{\partial W(\chi)}{\partial \mathbf{C}} \cdot \mathbf{F}^T. \quad (3.30)$$

Here  $W(\chi)$  denotes a strain-energy law (hyperelastic Helmholtz energy functional) dependent on the deformation of the solid. The fluid stress tensor can be written as (see Boer [2005, eqn. (8.63)])

$$\boldsymbol{\sigma}^f = \boldsymbol{\sigma}_{vis}^f - \phi\mathbf{I}p, \quad (3.31)$$

where  $\boldsymbol{\sigma}_{vis}^f$  denotes the viscous stress tensor of the fluid, given by (see Boer [2005, eqn. (6.145)])

$$\boldsymbol{\sigma}_{vis}^f = \mu_f\phi(\nabla\mathbf{v}_f + (\nabla\mathbf{v}_f)^T - \frac{2}{3}\nabla \cdot \mathbf{v}_f), \quad (3.32)$$

where  $\mu_f$  is the dynamic viscosity of the fluid.

### 3.1.5 Momentum balance of the mixture

The balance of linear momentum for the mixture can be obtained by adding the momentum of all its constituents, the solid and the fluid. Using the constitutive relations, and then adding the momentum balance equation (3.23) with  $\alpha = s$  and  $\alpha = f$  we get

$$\hat{\rho}^s \mathbf{a}^s + \hat{\rho}^f \mathbf{a}^f + \mathbf{v}^s \left( \frac{d^s \hat{\rho}^s}{dt} + \hat{\rho}^s \nabla \cdot \mathbf{v}^s \right) + \mathbf{v}^f \left( \frac{d^f \hat{\rho}^f}{dt} + \hat{\rho}^f \nabla \cdot \mathbf{v}^f \right) = \nabla \cdot (\boldsymbol{\sigma}_e + \boldsymbol{\sigma}_{vis} - p \mathbf{I}) + \rho \mathbf{f} \quad \text{in } \Omega_t.$$

After some calculations (see Chapelle et al. [2010, section 3.2] for details) we get

$$\hat{\rho}^s \mathbf{a}^s + \hat{\rho}^f \mathbf{a}^f = \nabla \cdot (\boldsymbol{\sigma}_e + \boldsymbol{\sigma}_{vis} - p \mathbf{I}) + \rho \mathbf{f} \quad \text{in } \Omega_t. \quad (3.33)$$

### 3.1.6 Momentum balance of the fluid

The momentum equation for the fluid flow can be identified from (3.23) with  $\alpha = f$  as

$$\hat{\rho}^f \mathbf{a}^f + \mathbf{v}^f \left( \frac{d^f \hat{\rho}^f}{dt} + \hat{\rho}^f \nabla \cdot \mathbf{v}^f \right) = \nabla \cdot (\boldsymbol{\sigma}_{vis}^f - \phi p \mathbf{I}) + \hat{\rho}^f \mathbf{f} + p \nabla \phi - \phi^2 \boldsymbol{\kappa}^{-1} (\mathbf{v}^f - \mathbf{v}^s) \quad \text{in } \Omega_t. \quad (3.34)$$

### 3.1.7 Summary of the general poroelasticity model

We define the boundary  $\partial \Omega_t = \Gamma_d \cup \Gamma_t$  for the mixture (solid skeleton), and  $\partial \Omega_t = \Gamma_p \cup \Gamma_f$  for the fluid, with an outward pointing unit normal  $\mathbf{n}$ . The strong problem for the full mixture theory model is to find  $\chi(\mathbf{X}, t)$ ,  $\mathbf{v}^f(\mathbf{x}, t)$

and  $p(\mathbf{x}, t)$  such that

$$\hat{\rho}^s \mathbf{a}^s + \hat{\rho}^f \mathbf{a}^f = \nabla \cdot (\boldsymbol{\sigma}_e + \boldsymbol{\sigma}_{vis} - p\mathbf{I}) + \rho \mathbf{f} \quad \text{in } \Omega_t, \quad (3.35a)$$

$$\begin{aligned} \hat{\rho}^f \mathbf{a}^f + \mathbf{v}^f \left( \frac{d^f \hat{\rho}^f}{dt} + \hat{\rho}^f \nabla \cdot \mathbf{v}^f \right) &= \nabla \cdot (\boldsymbol{\sigma}_{vis}^f - \phi p \mathbf{I}) \\ &+ p \nabla \phi - \phi \boldsymbol{\kappa}^{-1} (\mathbf{v}^f - \mathbf{v}^s) + \hat{\rho}^f \mathbf{f} \quad \text{in } \Omega_t, \end{aligned} \quad (3.35b)$$

$$\nabla \cdot ((1 - \phi) \mathbf{v}^s) + \nabla \cdot (\phi \mathbf{v}^f) = g \quad \text{in } \Omega_t, \quad (3.35c)$$

$$\boldsymbol{\chi} = \mathbf{X} + \mathbf{u}_D \quad \text{on } \Gamma_d, \quad (3.35d)$$

$$(\boldsymbol{\sigma}_e + \boldsymbol{\sigma}_{vis} - p\mathbf{I}) \mathbf{n} = \mathbf{t}_N \quad \text{on } \Gamma_t, \quad (3.35e)$$

$$\mathbf{v}^f = \mathbf{v}_D^f \quad \text{on } \Gamma_f, \quad (3.35f)$$

$$\mu_f \phi \frac{\partial \mathbf{v}^f}{\partial \mathbf{n}} - \phi p \mathbf{n} = \hat{\mathbf{s}} \quad \text{on } \Gamma_p, \quad (3.35g)$$

$$\boldsymbol{\chi}(0) = \mathbf{X} + \mathbf{u}^0, \quad \mathbf{v}^s(0) = \mathbf{v}^{s0}, \quad \mathbf{v}^f(0) = \mathbf{v}^{f0} \quad \text{in } \Omega_0. \quad (3.35h)$$

We have summarized all the variables and corresponding equations in table 3.1.

Unknown	Notation	Equation
<b>Primary variables</b>		<b>Primary equations (general model)</b>
Motion of the solid	$\chi$	$\hat{\rho}^s \mathbf{a}^s + \hat{\rho}^f \mathbf{a}^f = \nabla \cdot (\boldsymbol{\sigma}_e + \boldsymbol{\sigma}_{vis} - p\mathbf{I}) + \rho \mathbf{f}$ (3.33)
Fluid velocity	$\mathbf{v}^f$	$\hat{\rho}^f \mathbf{a}^f + \mathbf{v}^f \left( \frac{d^f \hat{\rho}^f}{dt} + \hat{\rho}^f \nabla \cdot \mathbf{v}^f \right) = \nabla \cdot (\boldsymbol{\sigma}_{vis}^f - \phi p \mathbf{I}) + p \nabla \phi - \phi \boldsymbol{\kappa}^{-1}(\mathbf{v}^f - \mathbf{v}^s) + \hat{\rho}^f \mathbf{f}$ (3.34)
Pressure of the fluid	$p$	$\nabla \cdot ((1 - \phi) \mathbf{v}^s) + \nabla \cdot (\phi \mathbf{v}^f) = g$ (3.21)
<b>Secondary variables</b>		<b>Secondary equations</b>
Deformation gradient tensor	$\mathbf{F}$	$\mathbf{F} = \frac{\partial}{\partial \mathbf{X}} \chi(\mathbf{X}, t)$ (7.2)
Right Cauchy-Green tensor	$\mathbf{C}$	$\mathbf{C} = \mathbf{F}^T \mathbf{F}$ (3.3)
Jacobian	$J$	$J = \det(\mathbf{F})$ (3.4)
Velocity of the solid	$\mathbf{v}^s$	$\mathbf{v}^s(\mathbf{x}, t) _{\mathbf{x}=\chi(\mathbf{X}, t)} = \frac{\partial}{\partial t} \chi(\mathbf{X}, t)$ (3.11)
Acceleration of the solid	$\mathbf{a}^s$	$\mathbf{a}^s(\mathbf{x}, t) _{\mathbf{x}=\chi(\mathbf{X}, t)} = \frac{\partial^2}{\partial t^2} \chi(\mathbf{X}, t)$ (3.25)
Acceleration of the fluid	$\mathbf{a}^f$	$\mathbf{a}^f = \frac{\partial}{\partial t} \mathbf{v}^f + (\nabla \mathbf{v}^f) \mathbf{v}^f$ (3.24)
Porosity	$\phi$	$\phi = 1 - \frac{1 - \phi_0}{J}$ (7.4)
Mixture density	$\rho$	$\rho = \rho^s(1 - \phi) + \rho^f \phi$ (7.3)
Eulerian solid density	$\hat{\rho}_s$	$\hat{\rho}^s = \rho^s(1 - \phi)$ (3.16)
Eulerian fluid density	$\hat{\rho}_f$	$\hat{\rho}^f = \rho^f \phi$ (3.17)
<b>Constitutive variables</b>		<b>Constitutive equations</b>
Solid elastic stress tensor	$\boldsymbol{\sigma}_e$	$\boldsymbol{\sigma}_e^s = \frac{1}{J} \mathbf{F} \cdot 2 \frac{\partial W(\chi)}{\partial \mathbf{C}} \cdot \mathbf{F}^T$ (7.8)
Fluid viscous stress tensor	$\boldsymbol{\sigma}_{vis}$	$\boldsymbol{\sigma}_{vis}^f = \mu_f \phi (\nabla \mathbf{v}_f + (\nabla \mathbf{v}_f)^T - \frac{2}{3} \nabla \cdot \mathbf{v}_f)$ (3.31)
Permeability tensor	$\boldsymbol{\kappa}$	$\boldsymbol{\kappa} = J^{-1} \mathbf{F} \boldsymbol{\kappa}_0 \mathbf{F}^T$ (7.6)

Table 3.1: Recapitulating the unknowns and equations of the general poroelasticity model.

### 3.1.8 Simplification and reformulation of the model

To arrive at the quasi-static, fully saturated, incompressible three-field large deformation poroelasticity model, we will now ignore inertia forces (left hand side of (3.33) and (3.34)), and ignore the viscous shear stress in the fluid ( $\boldsymbol{\sigma}_{vis}^f$  in (3.34)). After making these assumptions, and rewriting the equations in terms of the fluid flux, given by

$$\mathbf{z} = \phi(\mathbf{v}^f - \mathbf{v}^s), \quad (3.36)$$

the resulting problem is to find  $\chi(\mathbf{X}, t)$ ,  $\mathbf{z}(\mathbf{x}, t)$  and  $p(\mathbf{x}, t)$  such that

$$-\nabla \cdot (\boldsymbol{\sigma}_e - p\mathbf{I}) = \rho\mathbf{f} \quad \text{in } \Omega_t, \quad (3.37a)$$

$$\kappa^{-1}\mathbf{z} + \nabla p = \rho^f\mathbf{f} \quad \text{in } \Omega_t, \quad (3.37b)$$

$$\nabla \cdot (\mathbf{v}^s + \mathbf{z}) = g \quad \text{in } \Omega_t, \quad (3.37c)$$

$$\chi = \mathbf{X} + \mathbf{u}_D \quad \text{on } \Gamma_d, \quad (3.37d)$$

$$(\boldsymbol{\sigma}_e - p\mathbf{I})\mathbf{n} = \mathbf{t}_N \quad \text{on } \Gamma_t, \quad (3.37e)$$

$$\mathbf{z} \cdot \mathbf{n} = q_D \quad \text{on } \Gamma_f, \quad (3.37f)$$

$$p = p_D \quad \text{on } \Gamma_p, \quad (3.37g)$$

$$\chi(0) = \mathbf{X} + \mathbf{u}^0, \quad \text{in } \Omega_0. \quad (3.37h)$$

This is the large deformation we will consider from here onwards.

## 3.2 Linear poroelasticity

We will now assume small deformations to yield a linear model of poroelasticity. This model is often referred to Biot's model in the geomechanics community and contains some additional terms. We will introduce the Biot model here, for use with a 2D cantilever bracket problem later tested in section 5.4.4, and to highlight that any subsequent theory developed in later chapters can be extended to the full Biot model. The governing equations of the Biot model, with displacement  $\mathbf{u}$ , fluid flux  $\mathbf{z}$ , and pressure  $p$  as primary variables are summarized below:

$$-\nabla \cdot \sigma = \mathbf{f} \quad \text{in } \Omega \times (0, T), \quad (3.38a)$$

$$\kappa^{-1}\mathbf{z} + \nabla p = \mathbf{b} \quad \text{in } \Omega \times (0, T), \quad (3.38b)$$

$$\nabla \cdot \mathbf{z} + \frac{\partial}{\partial t}(\alpha \nabla \cdot \mathbf{u} + c_0 p) = g \quad \text{in } \Omega \times (0, T), \quad (3.38c)$$

$$\mathbf{u} = \mathbf{u}_D \quad \text{on } \Gamma_D \times (0, T), \quad (3.38d)$$

$$\sigma \mathbf{n} = \mathbf{t}_N \quad \text{on } \Gamma_N \times (0, T), \quad (3.38e)$$

$$p = p_D \quad \text{on } \Gamma_P \times (0, T), \quad (3.38f)$$

$$\mathbf{z} \cdot \mathbf{n} = q_D \quad \text{on } \Gamma_F \times (0, T), \quad (3.38g)$$

$$\mathbf{u}(0, \cdot) = \mathbf{u}^0, \quad p(0) = p^0, \quad \text{in } \Omega, \quad (3.38h)$$

where  $\sigma$  is the total stress tensor given by  $\sigma = \lambda \text{tr}(\epsilon(\mathbf{u}))\mathbf{I} + 2\mu_s \epsilon(\mathbf{u}) - \alpha p \mathbf{I}$ , with the linear strain tensor defined as  $\epsilon(\mathbf{u}) = \frac{1}{2} (\nabla \mathbf{u} + (\nabla \mathbf{u})^T)$ ,  $g$  is the fluid source term,  $\mathbf{f}$  is the body force on the mixture, and  $\mathbf{b}$  is the body

force on the fluid. Here  $\Omega$  is a bounded domain in  $\mathbb{R}^2$  or  $\mathbb{R}^3$ , and for the purpose of defining boundary conditions,  $\partial\Omega = \Gamma_D + \Gamma_N$  for displacement and stress boundary conditions and  $\partial\Omega = \Gamma_P + \Gamma_F$  for pressure and flux boundary conditions, with outward pointing unit normal  $\mathbf{n}$ . The parameters along with a description are given in Table 6.3.

Parameter	
Lamé's first parameter	$\lambda$ ,
Lamé's second parameter (shear modulus)	$\mu_s$ ,
Dynamic permeability tensor	$\kappa$ ,
Biot-Willis constant	$\alpha$ ,
Constrained specific storage coefficient	$c_0$ .

Table 3.2: Poroelasticity parameters.

We have also set  $\kappa = \mu_f^{-1}\mathbf{k}$ , where  $\mu_f$  and  $\mathbf{k}$  are the fluid viscosity and the permeability tensor, respectively. A derivation and more detailed explanation of these equations can be found in Phillips and Wheeler [2007a] and Showalter [2000]. In this work we will mainly consider a simplification of the full Biot model (3.38), by setting  $\alpha = 1$  and  $c_0 = 0$ . This yields the following fully

saturated and incompressible model:

$$-(\lambda + \mu_s)\nabla(\nabla \cdot \mathbf{u}) - \mu_s\nabla^2\mathbf{u} + \nabla p = \mathbf{f} \quad \text{in } \Omega \times (0, T), \quad (3.39a)$$

$$\kappa^{-1}\mathbf{z} + \nabla p = \mathbf{b} \quad \text{in } \Omega \times (0, T), \quad (3.39b)$$

$$\nabla \cdot (\mathbf{u}_t + \mathbf{z}) = g \quad \text{in } \Omega \times (0, T), \quad (3.39c)$$

$$\mathbf{u} = \mathbf{u}_D \quad \text{on } \Gamma_D \times (0, T), \quad (3.39d)$$

$$\sigma\mathbf{n} = \mathbf{t}_N \quad \text{on } \Gamma_N \times (0, T), \quad (3.39e)$$

$$p = p_D \quad \text{on } \Gamma_P \times (0, T), \quad (3.39f)$$

$$\mathbf{z} \cdot \mathbf{n} = q_D \quad \text{on } \Gamma_F \times (0, T), \quad (3.39g)$$

$$\mathbf{u}(0, \cdot) = \mathbf{u}^0 \quad \text{in } \Omega. \quad (3.39h)$$

This model is the small deformation version of the simplified and reformulated large deformation poroelasticity model (7.5). The extension of the theoretical results presented in Chapter 5 to the full Biot equations (3.38), with  $\alpha \in \mathbb{R}_{>0}$  and  $c_0 \in \mathbb{R}_{>0}$  is straightforward. In the analysis, the constant  $\alpha$  would just get absorbed by a general constant  $C$ . When  $c_0 > 0$ , an additional pressure term is introduced into the mass conservation equation. Since this term is coercive, it only improves the stability of the system.

# Chapter 4

## Finite element method

### 4.1 Introduction

A large proportion of the mathematical models in science and engineering take the form of differential equations. Only in the simplest cases, or under strong assumptions, is it possible to find exact analytical solutions to the equations in the model. Numerical methods are an established means of solving differential equations that are of practical interest in a variety of applied problems. Finite difference, finite volume and finite element methods are the most widely used types of such methods. Their basic idea is replacing the original infinite-dimensional problem by a finite-dimensional approximation, which is, generally speaking, easier to solve. Finite element methods are based on weakening the restrictions on the solution space in the continuous setting, and searching for the approximate solution in the subspace

which spans basis functions supported on small regions inside the domain. These methods are well-suited to solving problems on complex domains, and are therefore widely used in practical applications. In this work we consider only finite element methods (FEMs) for solving partial differential equations. This chapter comprises an overview of several theoretical and practical aspects of classical FEMs. The theory and notation presented here are essential in developing the techniques that form the core of this thesis. Most of the work presented in this chapter is based on work already presented in Arthurs [2012], Asner [2013], Bernabeu [2011], Burman and Hansbo [2007].

## 4.2 Functional analysis

Let  $\Omega$  be a bounded domain in  $\mathbb{R}^2$  or  $\mathbb{R}^3$ , and  $\partial\Omega$  be the associated boundary. The space of square integrable functions is then given by

$$L^2(\Omega) = \left\{ u : \int_{\Omega} |u(x)|^2 dx < \infty \right\},$$

with norm

$$\|u\|_{0,\Omega} = \left\{ \int_{\Omega} |u(x)|^2 dx \right\}^{1/2}.$$

This space is equipped with the inner product

$$(u, v) := \int_{\Omega} u(x)v(x)dx,$$

such that  $\|u\|_{0,\Omega} = (u, v)^{1/2}$ . Throughout this thesis we shall frequently refer to the Hilbertian Sobolev spaces  $H^1(\Omega)$  and  $H^2(\Omega)$ . The definitions of these are as follows:

$$H^1(\Omega) = \left\{ u \in L^2(\Omega) : \frac{\partial u}{\partial x_j} \in L_2(\Omega), j = 1, \dots, n, \right\}$$

$$H^2(\Omega) = \left\{ u \in L^2(\Omega) : \frac{\partial u}{\partial x_j} \in L_2(\Omega), j = 1, \dots, n, \frac{\partial^2 u}{\partial x_i \partial x_j} \in L_2(\Omega), i, j = 1, \dots, n \right\}$$

The corresponding norms are defined as

$$\|u\|_{1,\Omega} = \left\{ \|u\|_{0,\Omega}^2 + \sum_{j=1}^n \left\| \frac{\partial u}{\partial x_j} \right\|_{0,\Omega}^2 \right\}^{1/2},$$

$$\|u\|_{2,\Omega} = \left\{ \|u\|_{0,\Omega}^2 + \sum_{j=1}^n \left\| \frac{\partial u}{\partial x_j} \right\|_{0,\Omega}^2 + \sum_{i,j=1}^n \left\| \frac{\partial^2 u}{\partial x_i \partial x_j} \right\|_{0,\Omega}^2 \right\}^{1/2}.$$

We also define the divergence space

$$H_{div}(\Omega) = \left\{ \mathbf{v} \in L^2(\Omega) : \nabla \cdot \mathbf{v} \in L^2(\Omega) \right\}.$$

The set of functions of  $L^2(\partial\Omega)$  which are traces of functions of  $H^1(\Omega)$  constitutes a subspace of  $L^2(\partial\Omega)$  denoted by  $H^{1/2}(\partial\Omega)$ . We will also briefly use linear and bounded functionals (dual spaces) of  $H^1$ ,  $H^{1/2}$  and  $H_{div}$ , which will be denoted by  $H^{-1}$ ,  $H^{-1/2}$  and  $H_{div}^{-1}$ , respectively. Finally we define the

following continuous time-dependent norms

$$\|v\|_{L^2(L^2)} = \left( \int_0^T \|v(\cdot, t_n)\|_{0,\Omega}^2 dt \right)^{1/2},$$

$$\|v\|_{L^\infty(L^2)} = \sup \left\{ \|v(\cdot, t_n)\|_{0,\Omega} : t_n \in [0, T] \right\},$$

and their discrete in time counterparts

$$\|v\|_{l^2(L^2)} = \left( \sum_{n=0}^N \Delta t_n \|v(\cdot, t_n)\|_{0,\Omega}^2 \right)^{1/2},$$

$$\|v\|_{l^\infty(L^2)} = \max \left\{ \|v(\cdot, t_n)\|_{0,\Omega} : t_n \in [0, t_1, t_2, \dots, T] \right\}.$$

Definitions for slightly different combinations of temporal and spatial norms such as  $\|v\|_{H^1(L^2)}$  are straightforward adaptations of the above. Additional shorthand notation we will be introduced throughout the thesis as is needed.

### 4.3 Model problem

It is instructive to begin at a simple level and proceed by incrementally adding to the complexity of the equations we are discretising when explaining the use of the FEM, so we begin by considering the heat equation problem: given

$T > 0$ , for  $t \in [0, T]$  ind  $u(x, t)$  such that

$$\frac{\partial u}{\partial t} - \nabla \cdot \nabla u = 0 \quad \text{in } \Omega_t, \quad (4.1a)$$

$$\mathbf{n} \cdot \nabla u = g_N \quad \text{on } \Gamma_N, \quad (4.1b)$$

$$u = g_D \quad \text{on } \Gamma_D, \quad (4.1c)$$

$$u(x, 0) = u^0(x) \quad \text{in } \Omega. \quad (4.1d)$$

Here  $\Omega$  is a bounded domain in  $\mathbb{R}^2$  or  $\mathbb{R}^3$ , with boundary  $\partial\Omega = \Gamma_N \cup \Gamma_D$ , that has an outward pointing unit normal  $\mathbf{n}$ . The initial condition is given by  $u^0(x)$ . In the case where  $g_N = 0$ , system (4.1) can describe the evolution of heat in an object with geometry described by  $\Omega$ , where we have perfect thermal insulation on  $\Gamma_N$  and fixed temperature distributions given by the function  $g_D$  defined on the boundary due to some part of the environment with fixed temperature contacting the object along  $\Gamma_D$ .

### 4.3.1 Weak formulation

The strong form of (4.1) requires  $u$  to be at least twice differentiable. To weaken the regularity restrictions we multiply equation (4.1a) by an arbitrary function  $v$ , called a test function, and integrate over  $\Omega$ :

$$\left( \frac{\partial u}{\partial t}, v \right) - (\nabla \cdot \nabla u, v) = 0.$$

Applying the divergence theorem, this equation can be rewritten:

$$\begin{aligned} \left( \frac{\partial u}{\partial t}, v \right) - (\nabla u \cdot \mathbf{n}, v)_{\partial\Omega} + (\nabla u, \nabla v) \\ = \left( \frac{\partial u}{\partial t}, v \right) - (\nabla u \cdot \mathbf{n}, v)_{\Gamma_D} - (g_N, v)_{\Gamma_N} + (\nabla u, \nabla v) = 0. \end{aligned}$$

Here  $(\cdot, \cdot)_{\Gamma_N}$  and  $(\cdot, \cdot)_{\Gamma_D}$  denote the inner product taken over  $\Gamma_N$  and  $\Gamma_D$ , respectively. Taking note of the Dirichlet condition (4.1c), and letting  $v = 0$  on  $\Gamma_D$ , we arrive at the following equation:

$$\left( \frac{\partial u}{\partial t}, v \right) + (\nabla u, \nabla v) = (g_N, v)_{\Gamma_N}.$$

Note that in this equation the second derivatives of  $u$  need not exist. With that in mind, both the solution and the test functions can come from the space  $H^1(\Omega)$ , as long as they satisfy the appropriate Dirichlet boundary conditions. For convenience we will use the notation  $X_D = \{v \in H^1(\Omega) | v = u_D \text{ on } \Gamma_D\}$  and  $X_0 = \{v \in H^1(\Omega) | v = 0 \text{ on } \Gamma_D\}$ . The weak formulation of (4.1a) is as follows: Find  $u \in X_D$  such that

$$\left( \frac{\partial u}{\partial t}, v \right) + (\nabla u, \nabla v) = (g_N, v)_{\Gamma_N} \quad \forall v \in X_0. \quad (4.2)$$

### 4.3.2 Time discretisation

We also need to choose a method of treating the time derivative. In this work, we do so using Euler difference quotients, and so we make the approximation

$u_t(x, t + \Delta t) \approx \frac{u(x, t + \Delta t) - u(x, t)}{\Delta t}$  for some constant time step  $\Delta t$ . We write  $u(x)^n$  for the temporally-semidiscrete approximation to  $u(x, n\Delta t)$ , and our numerical scheme will yield approximations at times  $t = 0, \Delta t, 2\Delta t, \dots, T$ . Inserting this difference quotient and assuming that  $\Delta T$  divides  $T$ , Equation (4.3) becomes: for  $n = 1, 2, \dots, \frac{T}{\Delta t}$ , find  $u^n \in X_D$  such that

$$(u^n, v) + \Delta t (\nabla u, \nabla v) = (g_N, v)_{\Gamma N} + (u^{n-1}, v) \quad \forall v \in X_0. \quad (4.3)$$

### 4.3.3 Finite element discretisation

In order to solve this problem numerically, we must make it finite dimensional by discretising it suitably. The finite element approximation space is constructed as follows: first, the problem domain is partitioned into small element domains, and second, the element is defined by prescribing for each element domain a set of nodes and nodal values, and defining suitable basis functions on these, for example, as piecewise-linear basis functions.

Element domains are normally shaped as triangles or squares in  $\mathbb{R}^2$ , tetrahedra or hexahedra in  $\mathbb{R}^3$ . All the nodes, edges and faces of element domains constitute the problem mesh. Defining local basis functions completes the finite element space. For a rigorous definition of finite elements, and a description of different types of elements we refer to Brenner and Scott [2008].

Let  $\mathcal{T}^h$  be a partition of  $\Omega$  into non-overlapping elements  $K$ . We denote by  $h$  the size of the largest element in  $\mathcal{T}^h$  and assume that the partition is quasi-uniform. On the given partition  $\mathcal{T}^h$  we then define the following finite

element spaces, to solve the model problem

$$X_{hD} = \{u \in C^0(\Omega) : u|_K \in P_1(K); u = u_D \text{ on } \Gamma_D; \forall K \in \mathcal{T}^h\},$$

$$X_{h0} = \{u \in C^0(\Omega) : u|_K \in P_1(K); u = 0 \text{ on } \Gamma_D; \forall K \in \mathcal{T}^h\},$$

where  $P_1(K)$  is the space of linear polynomials on  $K$ , and  $C^0(\Omega)$  is the space of continuous functions on  $\Omega$ . The discretised problem, for each time step, is to find  $u_h^n \in X_{hD}$  such that

$$(u_h^n, v_h) + \Delta t (\nabla u_h, \nabla v_h) = (g_N, v_h)_{\Gamma N} + (u_h^{n-1}, v_h) \quad \forall v_h \in X_{h0}. \quad (4.4)$$

We now choose the Lagrangian basis  $\{\phi_1, \phi_2, \dots, \phi_m\}$  of  $X^h$  defined by the nodal values at the nodes  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ , namely

$$\phi_i(\mathbf{x}_j) = \delta_{i,j} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases},$$

We observe that a basis of  $X_{h0}$  can be constructed by removing  $\phi_i$  with  $\mathbf{x}_i \in \Gamma_D$  from the basis of  $X_h$ . Let us assume that the indices of such basis functions are  $1, \dots, m$ , and therefore  $X_{h0} = \text{span}\{\phi_1, \dots, \phi_m\}$ . The finite-dimensional weak problem (4.4) is equivalent to: Find  $u_h^n \in X_{hD}$  such that

$$(u_h^n, \phi_i) + \Delta t (\nabla u_h, \nabla \phi_i) = (g_N, \phi_i)_{\Gamma N} + (u_h^{n-1}, \phi_i) \quad \forall i = 1, \dots, m. \quad (4.5)$$

Any function from  $X_h$  can be presented in the form of a basis expansion. Let this basis expansion for  $u_h$  be

$$u_h^n = \sum_{i=1}^m u_i^n \phi_i,$$

with  $u_i^n = u_h^n(\mathbf{x}_i)$ . We define the vector of nodal values to be  $\mathbf{u}^n = [u_1^n, \dots, u_m^n]$ . Substituting this expression into (4.5), we finally obtain a linear system which we can solve for  $\mathbf{u}^n$ :

$$(\mathbf{M} + \Delta t \mathbf{A}) \mathbf{u}^n = \mathbf{M} \mathbf{u}^{n-1} + \mathbf{g}, \quad (4.6)$$

where we have defined the following matrices and vectors:

$$\mathbf{A} = [\mathbf{a}_{ij}], \quad \mathbf{a}_{ij} = \int_{\Omega} \nabla \phi_i \cdot \nabla \phi_j,$$

$$\mathbf{M} = [\mathbf{m}_{ij}], \quad \mathbf{m}_{ij} = \int_{\Omega} \phi_i \cdot \phi_j,$$

$$\mathbf{g} = [\mathbf{g}_i], \quad \mathbf{g}_i = \int_{\Gamma_N} g_N \cdot \phi_i,$$

$$\mathbf{u}^n = (u_1^n, \dots, u_m^n)^T.$$

The linear system of equations (4.6) can be solved using standard methods such as Gaussian elimination.

## 4.4 Mixed methods

Before considering the discretisation of the poroelasticity equations in chapter ?? we first consider the problems of Darcy and Stokes flow. Solving the three-field poroelasticity problem is essentially equivalent to coupling the Stokes equations (elasticity of the porous mixture) with the Darcy equations (fluid flow through pores), with a modified incompressibility constraint that combines the divergence of the displacement velocity and the fluid flux. Mixed methods refer to the discretisation of different variable using different finite elements. We begin with a unified formulation of both the Darcy and Stokes flow equations:

$$\mathbf{A}(\mathbf{u}) + \nabla p = \mathbf{f} \quad \text{in } \Omega_t, \tag{4.7a}$$

$$\nabla \cdot \mathbf{u} = 0 \quad \text{in } \Omega_t, \tag{4.7b}$$

where  $\mathbf{u}$  denotes the velocity vector and  $p$  the pressure and  $\mathbf{f} \in [L^2(\Omega)]^d$ , with  $d = 2, 3$ . For simplicity we assume Dirichlet conditions on the boundary, that is,  $\mathbf{u} = 0$  on  $\Gamma_D$  for Stokes and  $\mathbf{u} \cdot \mathbf{n} = 0$  on for Darcy. For the choice of  $\mathbf{A}$  we focus on two cases

- $\mathbf{A}(\mathbf{u}) := \kappa^{-1} \mathbf{I} \mathbf{u}$ , corresponding to Darcy's equation.
- $\mathbf{A}(\mathbf{u}) := -2\mu_f \nabla \cdot \epsilon(\mathbf{u})$ , corresponding to Stokes equation.

In order to formulate our finite element method we first introduce the weak formulation of problem (4.7). We introduce the spaces

$$W^D = \{\mathbf{v} \in H_{div}(\Omega) : \mathbf{v} \cdot \mathbf{n} = 0 \text{ on } \partial\Omega\},$$

$$W^S = \{\mathbf{v} \in [H_1(\Omega)]^d : \mathbf{v} = 0 \text{ on } \Gamma_D\},$$

and

$$L_0^2 = \left\{ q \in L_2(\Omega) : \int_{\Gamma} q \, dx = 0 \right\}.$$

We denote the product space  $W^X \times L_0^2$  by  $\mathcal{W}^X$  where  $X$  is chosen to be  $D$  for the Darcy equation or  $S$  for the Stokes equation. We also define the following norm on  $\mathcal{W}^X$ :

$$\|(\mathbf{u}, p)\|_{\mathcal{W}^X}^2 = \|\mathbf{u}\|_{l, \Omega}^2 + \|\nabla \cdot \mathbf{u}\|_{0, \Omega}^2 + \|p\|_{0, \Omega}^2,$$

with  $l = 0$  for Darcy and  $l = 1$  for Stokes. Let  $a(\mathbf{u}, \mathbf{v})$  be the bilinear form corresponding to the weak formulation of  $A(\mathbf{u})$  and consider the combined bilinear form

$$B[(\mathbf{u}, p), (\mathbf{v}, q)] = a(\mathbf{u}, \mathbf{v}) - (p, \nabla \cdot \mathbf{v}) + (q, \nabla \cdot \mathbf{u}).$$

The continuous weak formulation of (4.7) is now to find  $(\mathbf{u}, p) \in \mathcal{W}^X$  such that

$$B[(\mathbf{u}, p), (\mathbf{v}, q)] = (\mathbf{f}, \mathbf{v}) \quad \forall (\mathbf{u}, p) \in \mathcal{W}^X.$$

By considering the discrete subspace  $\mathcal{W}_h^X \in \mathcal{W}^X$ , we arrive at the following discrete formulation of the problem: find  $(\mathbf{u}_h, p_h) \in \mathcal{W}_h^X$  such that:

$$B[(\mathbf{u}_h, p_h), (\mathbf{v}_h, q_h)] = (\mathbf{f}, \mathbf{v}_h) \quad \forall (\mathbf{v}_h, q_h) \in \mathcal{W}_h^X.$$

To ensure stability and convergence of the discretisation, the discrete subspace (mixed element) has to be chosen such that the following discrete inf-sup condition [Babuška, 1971] is fulfilled. Let  $\gamma > 0$  be a constant independent of any mesh parameters.

$$\gamma \|(\mathbf{u}_h, p_h)\|_{\mathcal{W}_h^X} \leq \sup_{(\mathbf{v}_h, q_h) \in \mathcal{W}_h^X} \frac{B_h[(\mathbf{u}_h, p_h), (\mathbf{v}_h, q_h)]}{\|(\mathbf{v}, q)\|_{\mathcal{W}_h^X}} \quad \forall (\mathbf{u}, p) \in \mathcal{W}_h^X. \quad (4.8)$$

Establishing this condition ensures wellposedness of the discretization so that the linear system arising from the fully discrete method is non-singular and can be solved using standard methods. It is not trivial to prove (4.8) for different finite element combinations. This has been a major research topics for many decades, and zillions of papers have been published. In table 4.1 we have documented some popular standard finite element pairs for solving the Stokes and Darcy equations, and outlined whether these satisfy (4.8), and therefore yield a stable and optimally converging method or not. Note that many other possible discretisations exists.

The naive choice of piecewise linear finite elements for both the velocities and the pressure (P1-P1) or piecewise linear finite elements for the velocities and piecewise constants (P1-P0) for the pressure results in an ill posed

Mixed element	Stokes	Darcy
P1-P1	✗	✗
P2-P1	✓	✗
P1-P1+stab	✓	✓
P1-P0	✗	✗
RT-P0	✗	✓
P1-P0+stab	✓	✓

Table 4.1: Possible finite element combinations for Stokes and Darcy flow, showing whether a particular choice of elements is stable and optimally converging or not.

discretizations [Burman and Hansbo, 2007]. Intuitively, this is because the velocity space is not rich enough to constrain the pressures, thus resulting in spurious pressure oscillations. A detailed explanation of this along with some worked examples can be found in Elman et al. [2005, section 5.3]. The Taylor-Hood element (P2-P1) is a commonly used element for the Stokes equations. However for the Darcy equations this element does not converge at the right order and fails to converge for the divergence of the velocities [Burman and Hansbo, 2007]. The Raviart-Thomas element (RT-P0), first proposed in Raviart and Thomas [1977] is a divergence free element, often used to solve the Darcy equations. Velocities are required to have continuous normal components across element interfaces, whereas tangential components are discontinuous. Pressure fields are discontinuous and must not be of too high order, otherwise the inf-sup condition is violated [Masud and Hughes, 2002]. More details on how to construct this element are given in Quarteroni and Valli [2008]. However this element is not able to control  $H^1$  velocities,

and therefore can not be used to solve the Stokes equations. When the finite element discretisation is based on a discrete subspace that does not satisfy the discrete inf-sup condition (4.8), a procedure aiming at stabilizing the discrete system may be accomplished. The philosophy of stabilized methods is to strengthen formulations by adding an extra term, often to the mass conservation equation, so that discrete approximations, which would otherwise be unstable, become stable and convergent [Masud and Hughes, 2002]. Numerous stabilization techniques exist. To stabilize the equal order piecewise linear pair, a polynomial pressure projection has been proposed that results in a stable element for both the Stokes and Darcy equations (P1-P1+stab) Bochev and Dohrmann [2006]. Also, a pressure jump stabilization that uses a piecewise constant pressure approximation and is stable and optimally converging for both the Stokes and Darcy equation has been analysed by Burman and Hansbo [2007].

# Chapter 5

## Stabilized low-order finite element approximation for linear three-field poroelasticity

### 5.1 Weak formulation

Before presenting the weak formulation of (3.39), we first need to introduce some notation and bilinear forms. We introduce the space  $L_0^2(\Omega) := \{q \in L^2(\Omega) : \int_{\Omega} q \, dx = 0\}$ , and define

$$\mathcal{L}(\Omega) := \begin{cases} L^2(\Omega) & \text{if } \Gamma_N \cup \Gamma_P \neq \emptyset \\ L_0^2(\Omega) & \text{if } \Gamma_N \cup \Gamma_P = \emptyset \end{cases}$$

We let  $\mathbf{W}^E = \{\mathbf{u} \in [H^1(\Omega)]^d : \mathbf{u} = \mathbf{u}_D \text{ on } \Gamma_D\}$ ,  $\mathbf{W}^D = \{\mathbf{z} \in H_{div}(\Omega) : \mathbf{z} \cdot \mathbf{n} = q_D \text{ on } \Gamma_F\}$ . We also define the mixed solution space  $\mathcal{W}^X = \{\mathbf{W}^E \times \mathbf{W}^D \times \mathcal{L}(\Omega)\}$ . For the test functions we define the spaces  $\mathbf{W}_0^E = \{\mathbf{v} \in [H^1(\Omega)]^d : \mathbf{v} = 0 \text{ on } \Gamma_D\}$ ,  $\mathbf{W}_0^D = \{\mathbf{w} \in H_{div}(\Omega) : \mathbf{w} \cdot \mathbf{n} = 0 \text{ on } \Gamma_F\}$ , and  $\mathcal{V}^X = \{\mathbf{W}_0^E \times \mathbf{W}_0^D \times \mathcal{L}(\Omega)\}$ . With  $\mathbf{u} \in \mathbf{W}^E$  and  $\mathbf{v} \in \mathbf{W}_0^E$ , we define the bilinear form

$$a(\mathbf{u}, \mathbf{v}) = \int_{\Omega} 2\mu_s(\epsilon(\mathbf{u}) : \epsilon(\mathbf{v})) + \lambda(\nabla \cdot \mathbf{u})(\nabla \cdot \mathbf{v}) \, dx,$$

which corresponds to the elasticity part of the mixture momentum equation (3.39a). This bilinear form is continuous such that

$$a(\mathbf{u}, \mathbf{v}) \leq C_c \|\mathbf{u}\|_{1,\Omega} \|\mathbf{v}\|_{1,\Omega} \quad \forall \mathbf{u}, \mathbf{v} \in [H^1(\Omega)]^d. \quad (5.1)$$

In addition, using Korn's inequality [Brenner and Scott, 2008, ?], and  $\int_{\Omega} \lambda(\nabla \cdot \mathbf{v})(\nabla \cdot \mathbf{v}) \geq 0$  we have

$$\|\mathbf{v}\|_{a,\Omega}^2 = a(\mathbf{v}, \mathbf{v}) \geq 2\mu_s \|\epsilon(\mathbf{v})\|_{0,\Omega}^2 \geq C_k \|\mathbf{v}_h\|_{1,\Omega}^2 \quad \forall \mathbf{v} \in \mathbf{W}_0^E. \quad (5.2)$$

We also have

$$\lambda_{min}^{-1} \|\mathbf{w}\|_{0,\Omega}^2 \geq (\kappa^{-1} \mathbf{w}, \mathbf{w}) \geq \lambda_{max}^{-1} \|\mathbf{w}\|_{0,\Omega}^2 \quad \forall \mathbf{w} \in \mathbf{W}_0^D, \quad (5.3)$$

since  $\kappa$  is assumed to be a symmetric and strictly positive definite tensor, there exists eigenfunctions  $\lambda_{min}, \lambda_{max} > 0$  such that  $\forall \mathbf{x} \in \Omega, \lambda_{min}\|\eta\|_{0,\Omega} \leq \eta^t \kappa(\mathbf{x})\eta \leq \lambda_{max}\|\eta\|_{0,\Omega} \quad \forall \eta \in \mathbb{R}^d$ .

### 5.1.1 Continuous weak formulation

We now multiply the strong form of the problem (3.39) by test functions  $(\mathbf{v}, \mathbf{w}, q) \in \mathcal{V}^X$  and integrate to yield the continuous weak problem, which is to find  $\mathbf{u}(x, t) \in \mathbf{W}^E$ ,  $\mathbf{z}(x, t) \in \mathbf{W}^D$ , and  $p(x, t) \in \mathcal{L}(\Omega)$  for any time  $t \in [0, T]$  such that:

$$a(\mathbf{u}, \mathbf{v}) - (p, \nabla \cdot \mathbf{v}) = (\mathbf{f}, \mathbf{v}) + (\mathbf{t}_N, \mathbf{v})_{\Gamma_N} \quad \forall \mathbf{v} \in \mathbf{W}_0^E, \quad (5.4a)$$

$$(\kappa^{-1}\mathbf{z}, \mathbf{w}) - (p, \nabla \cdot \mathbf{w}) = (\mathbf{b}, \mathbf{w}) - (p_D, \mathbf{w} \cdot \mathbf{n})_{\Gamma_P} \quad \forall \mathbf{w} \in \mathbf{W}_0^D, \quad (5.4b)$$

$$(\nabla \cdot \mathbf{u}_t, q) + (\nabla \cdot \mathbf{z}, q) = (g, q) \quad \forall q \in \mathcal{L}(\Omega). \quad (5.4c)$$

We will assume the following regularity requirements on the data:

$$\mathbf{f} \in C^1([0, T]; (H^{-1}(\Omega))^d), \quad (5.5a)$$

$$\mathbf{b} \in C^1([0, T]; H_{div}^{-1}(\Omega)), \quad (5.5b)$$

$$g \in C^0([0, T]; (L^2(\Omega))^d), \quad (5.5c)$$

$$\mathbf{u}_D \in C^1([0, T]; H^{1/2}(\Gamma_D)), \quad (5.5d)$$

$$\mathbf{t}_N \in C^1([0, T]; H^{-1/2}(\Gamma_N)), \quad (5.5e)$$

$$q_D \in C^0([0, T]; TrW), \quad (5.5f)$$

$$p_D \in C^0([0, T]; L^2(\Gamma_P)), \quad (5.5g)$$

where  $TrW := \{\mathbf{w} \cdot \mathbf{n}|_{\Gamma_f} : \mathbf{w} \in H_{div}(\Omega)\}$ . For the initial conditions we require that  $\mathbf{u}^0 \in (H^1(\Omega))^d$ . The well-posedness of the continuous two-field formulation has been proven by Showalter [2000]. Lipnikov [2002] proves well-posedness for the continuous three-field formulation (6.1). In this work we also establish the well-posedness of (6.1) as a result of the energy estimates proven in section 5.2.3, see remark 2.

**Remark 1** Since the above resulting system of equations is linear, for ease within the analysis, we will assume all boundary conditions are homogeneous. Hence,  $\mathcal{W}^X = \mathcal{V}^X$ .

## 5.2 Fully-discrete model

We discretize in space using the following finite elements. Let  $\mathcal{T}^h$  be a partition of  $\Omega$  into non-overlapping elements  $K$ . We denote by  $h$  the size of the largest element in  $\mathcal{T}^h$  and assume that the partition is quasi-uniform. On the given partition  $\mathcal{T}^h$  we then define the following finite element spaces

$$\begin{aligned}\mathbf{W}_h^E &= \left\{ \mathbf{u}_h \in C^0(\Omega) : \mathbf{u}_h|_K \in P_1(K); \mathbf{u}_h = 0 \text{ on } \Gamma_D; \forall K \in \mathcal{T}^h \right\}, \\ \mathbf{W}_h^D &= \left\{ \mathbf{z}_h \in C^0(\Omega) : \mathbf{z}_h|_K \in P_1(K); \mathbf{z}_h \cdot \mathbf{n} = 0 \text{ on } \Gamma_F; \forall K \in \mathcal{T}^h \right\}, \\ Q_h &= \begin{cases} \left\{ p_h : p_h|_K \in P_0(K); \forall K \in \mathcal{T}^h \right\} & \text{if } \Gamma_N \cup \Gamma_P \neq \emptyset \\ \left\{ p_h : p_h|_K \in P_0(K); \int_{\Omega} p_h = 0; \forall K \in \mathcal{T}^h \right\} & \text{if } \Gamma_N \cup \Gamma_P = \emptyset \end{cases},\end{aligned}$$

where  $P_0(K)$  and  $P_1(K)$  are the spaces of constant and linear polynomials on  $K$ , respectively.

The discretisation in time is given by partitioning  $[0, T]$  into  $N$  evenly spaced non-overlapping regions  $(t_{n-1}, t_n]$ ,  $n = 1, 2, \dots, N$ , where  $t_n - t_{n-1} = \Delta t$ . For any sufficiently smooth function  $v(t, x)$  we define  $v^n(x) = v(t_n, x)$  and the discrete time derivative by  $v_{\delta t}^n := \frac{v^n - v^{n-1}}{\Delta t}$ .

The fully discrete weak problem is now to find, for  $n = 1, 2, \dots, N$ ,  $\mathbf{u}_h^n \in$

$\mathbf{W}_h^E$ ,  $\mathbf{z}_h^n \in \mathbf{W}_h^D$  and  $p_h^n \in Q_h$  such that:

$$a(\mathbf{u}_h^n, \mathbf{v}_h) - (p_h^n, \nabla \cdot \mathbf{v}_h) = (\mathbf{f}^n, \mathbf{v}_h) + (\mathbf{t}_N, \mathbf{v}_h)_{\Gamma_N} \quad \forall \mathbf{v}_h \in \mathbf{W}_h^E, \quad (5.6a)$$

$$(\kappa^{-1} \mathbf{z}_h^n, \mathbf{w}_h) - (p_h^n, \nabla \cdot \mathbf{w}_h) = (\mathbf{b}^n, \mathbf{w}_h) - (p_D, \mathbf{w}_h \cdot \mathbf{n})_{\Gamma_P} \quad \forall \mathbf{w}_h \in \mathbf{W}_h^D, \quad (5.6b)$$

$$(\nabla \cdot \mathbf{u}_{\delta t, h}, q_h) + (\nabla \cdot \mathbf{z}_h^n, q_h) + J(p_{\delta t, h}, q_h) = (g^n, q_h) \quad \forall q_h \in Q_h. \quad (5.6c)$$

Along with the following initial approximations

$$a(\mathbf{u}_h^0, \mathbf{v}) = a(\mathbf{u}^0, \mathbf{v}) \quad \forall \mathbf{v}_h \in \mathbf{W}_{h0}^E, \quad (5.7a)$$

$$J(p_h^0, q_h) = J(p^0, q_h) \quad \forall q_h \in Q_h, \quad (5.7b)$$

where  $p^0 \in \mathcal{L}(\Omega)$ . The stabilization term is given by, see Burman and Hansbo [2007],

$$J(p, q) = \delta \sum_K \int_{\partial K \setminus \partial \Omega} h_{\partial K} [p][q] \, ds. \quad (5.8)$$

Here  $\delta$  is a stabilization parameter that is independent of  $h$  and  $\Delta t$ . We will see in the numerical results, section 6.2 that the convergence is not sensitive to  $\delta$ . Here  $h_{\partial K}$  denotes the size (diameter) of an element edge in 2D or face in 3D, and  $[\cdot]$  is the jump across an edge or face (taken on the interior edges only).

### 5.2.1 Auxiliary results

The stabilization term gives rise to the semi-norm  $|q|_{J,\Omega} := J(q, q)^{1/2}$ . Throughout this work, we will let  $C$  denote a generic positive constant, whose value may change from instance to instance, but is independent of any mesh parameters. Using standard scaling arguments, Cauchy-Schwarz and the triangle inequality the following bounds for the stabilization term hold.

$$|p_h|_{J,\Omega} \leq C\|p_h\|_{0,\Omega} \quad \text{and} \quad J(q_h, p_h) \leq |q_h|_{J,\Omega}|p_h|_{J,\Omega}, \quad \forall p_h, q_h \in Q_h. \quad (5.9)$$

Furthermore, for any  $q \in H^1(\Omega)$ ,

$$J(q, p) = 0, \quad \forall p \in \mathcal{L}(\Omega), \quad (5.10)$$

see lemma 1.23 in Di Pietro and Ern [2011]).

We will now give some approximation results that will be useful later.

Let  $\pi_h^1 : H^1(\Omega) \rightarrow \mathbf{W}_h^E$  and  $\pi_h^0 : L^2(\Omega) \rightarrow Q_h$  be Clément projections, see ?.

**Lemma 5.2.1** *For all  $\mathbf{v} \in (H^2(\Omega))^d$  and  $q \in H^1(\Omega)$  the interpolation operators satisfy: For  $s = 0, 1$*

$$\|\mathbf{v} - \pi_h^1 \mathbf{v}\|_{s,\Omega} \leq Ch^{2-s} \|\mathbf{v}\|_{2,\Omega}, \quad (5.11)$$

$$\|q - \pi_h^0 q\|_{0,\Omega} \leq Ch \|q\|_{1,\Omega}, \quad (5.12)$$

$$|q - \pi_h^0 q|_{J,\Omega} \leq Ch \|q\|_{1,\Omega}. \quad (5.13)$$

**Proof 1** *The first two results are standard. The final result is obtained by using the element error estimate provided in Verfürth [1998] and then summing over all elements.*

Due to the surjectivity of the divergence operator, for every  $p \in L^2(\Omega)$  there exists a function  $\mathbf{v}_p \in (H^1(\Omega))^d$  such that  $\nabla \cdot \mathbf{v}_p = -p$  and  $\|\mathbf{v}_p\|_{1,\Omega} \leq c\|p\|_{0,\Omega}$ . We assume that the projection,  $\pi_h^1 \mathbf{v}_p$ , is stable such that

$$\|\pi_h^1 \mathbf{v}_p\|_{1,\Omega} \leq \hat{c}\|p\|_{0,\Omega}. \quad (5.14)$$

Furthermore, for any element  $K \in \mathcal{T}^h$

$$\|\mathbf{v}_p - \pi_h^1 \mathbf{v}_p\|_{L^2(K)} \leq Ch\|\mathbf{v}_p\|_{H^1(\omega_K)}, \quad (5.15)$$

where  $\omega_K$  is the union of all elements  $J \in \mathcal{T}^h$  such that  $\overline{K} \cap \overline{J} \neq \emptyset$ .

Combining the above with the trace inequality, see lemma 3.1 in Verfürth [1998],

$$\|(\mathbf{v}_p - \pi_h^1 \mathbf{v}_p) \cdot \mathbf{n}\|_{0,\partial K}^2 \leq C\|\mathbf{v}_p - \pi_h^1 \mathbf{v}_p\|_{0,K}(h^{-1}\|\mathbf{v}_p - \pi_h^1 \mathbf{v}_p\|_{0,K} + \|\mathbf{v}_p - \pi_h^1 \mathbf{v}_p\|_{1,K}), \quad (5.16)$$

we obtain

$$\|(\mathbf{v}_p - \pi_h^1 \mathbf{v}_p) \cdot \mathbf{n}\|_{0,\partial K}^2 \leq Ch\|\mathbf{v}_p\|_{H^1(\omega_K)}^2. \quad (5.17)$$

Taking into account  $\|\mathbf{v}_p\|_{1,\Omega} \leq c\|p\|_{0,\Omega}$ , we may write

$$\sum_K \int_{\partial K} h^{-1}|(\mathbf{v}_p - \pi_h^1 \mathbf{v}_p) \cdot \mathbf{n}|^2 ds \leq c_t \|p\|_{0,\Omega}^2. \quad (5.18)$$

We define the fully discrete finite element approximation for all time to be the piecewise constant in time functions  $\mathbf{u}_h(t, \mathbf{x}) := \mathbf{u}_h^n(\mathbf{x})$  for  $t \in (t_{n-1}, t_n]$ ,  $\mathbf{z}_h$  and  $p_h$  are defined similarly. For such piecewise continuous in time functions,  $v$ , the norms  $L^2([0, T]; X)$  satisfy

$$\|v\|_{L^2([0, T]; X)} = \int_0^T \|v(s, \cdot)\|_X^2 ds = \sum_{n=1}^N \Delta t \|v^n\|_X^2,$$

where  $X$  is any given function space over  $\Omega$ .

For all  $\mathbf{v} \in H^2([0, T]; [L^2(\Omega)]^d)$

$$\sum_{n=1}^N \Delta t \left\| \mathbf{v}_{\delta t}^n - \frac{\partial \mathbf{v}}{\partial t}(t^n, \cdot) \right\|_{0,\Omega}^2 \leq \Delta t^2 \int_0^T \|\mathbf{v}_{tt}\|_{0,\Omega}^2 ds. \quad (5.19)$$

For all  $[\mathbf{v}, \mathbf{w}, q] \in [(H^1(\Omega))^d \times H_{div}(\Omega) \times L^2(\Omega)]$  we define the norm

$$\|[\mathbf{v}, \mathbf{w}, q]\|_A^2 := \|\mathbf{v}\|_{1,\Omega}^2 + \Delta t^2 \|\nabla \cdot \mathbf{w}\|_{0,\Omega}^2 + \Delta t \|\mathbf{w}\|_{0,\Omega}^2 + \|q\|_{0,\Omega}^2 + |q|_{J,\Omega}^2. \quad (5.20)$$

For all  $[\mathbf{v}, \mathbf{w}, q] \in [L^\infty([0, T]; (H^1(\Omega))^d) \times L^2([0, T]; H_{div}(\Omega)) \times L^2([0, T]; L^2(\Omega))]$  the norm

$$\|[\mathbf{v}, \mathbf{w}, q]\|_B^2 := \|\mathbf{v}\|_{L^\infty(H^1)}^2 + \|\mathbf{w}\|_{L^2(L^2)}^2 + \|q\|_{L^2(L^2)}^2, \quad (5.21)$$

### 5.2.2 Existence and uniqueness of the fully-discrete model at a time step

The well-posedness of the the fully-discretized system (5.6) with  $\delta = 0$  and the additional term arising from  $c_0 > 0$  is shown by Phillips and Wheeler [2007b]. Lipnikov [2002] shows the more general case for  $c_0 \geq 0$ . As the permeability tends to zero and the porous mixture becomes impermeable, the three-field linear poroelasticity tends to a mixed linear elasticity problem. The choice of the mixed finite element space then also needs to be stable for the mixed linear elasticity problem [see, Haga et al., 2012]. However the discretization methods presented in Phillips and Wheeler [2007b] and Lipnikov [2002] both use piecewise linear approximations for displacements and mixed low-order Raviart Thomas elements for the fluid flux and pressure variables. It is well known that a piecewise linear approximations for displacements and piecewise constant approximation for the pressure is not a stable element combination for the mixed linear elasticity problem. We hypothesise that this is why the method presented by Phillips and Wheeler [2007b] experiences numerical instabilities, as is shown in Phillips and Wheeler [2008] and Phillips and Wheeler [2009]. Our proposed stabilized method is stable for both the Darcy problem (as the elasticity coefficients tend to infinity) and the mixed linear elasticity problem (as the permeability tends to zero), and therefore is stable for all permeabilities and elasticity coefficients.

Combining the fully discrete balance equations (5.6a), (5.6b) and (5.6c),

after first multiplying (5.6b) and (5.6c) by  $\Delta t$ , gives the equivalent problem;

For  $n = 1, 2, \dots, n$  find  $(\mathbf{u}_h, \mathbf{z}_h, p_h)$  such that

$$\begin{aligned} B_h^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (\mathbf{v}_h, \mathbf{w}_h, q_h)] &= (\mathbf{f}^n, \mathbf{v}_h) + (\mathbf{t}_N, \mathbf{v}_h)_{\Gamma_N} + \Delta t(\mathbf{b}^n, \mathbf{w}_h) - \Delta t(p_D, \mathbf{w}_h \cdot \mathbf{n})_{\Gamma_P} + \Delta t(g^n, q_h) \\ &\quad + (\nabla \cdot \mathbf{u}_h^{n-1}, q_h) + J(p_h^{n-1}, q_h) \quad \forall (\mathbf{v}_h, \mathbf{w}_h, q_h) \in \mathcal{W}_h^X, \end{aligned} \quad (5.22)$$

where

$$\begin{aligned} B_h^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (\mathbf{v}_h, \mathbf{w}_h, q_h)] &= a(\mathbf{u}_h^n, \mathbf{v}_h) + \Delta t(\kappa^{-1} \mathbf{z}_h^n, \mathbf{w}_h) - (p_h^n, \nabla \cdot \mathbf{v}_h) - \Delta t(p_h^n, \nabla \cdot \mathbf{w}_h) \\ &\quad + (\nabla \cdot \mathbf{u}_h^n, q_h) + \Delta t(\nabla \cdot \mathbf{z}_h^n, q_h) + J(p_h^n, q_h). \end{aligned} \quad (5.23)$$

The linear form satisfies the following continuity property

$$|B_h^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (\mathbf{v}_h, \mathbf{w}_h, q_h)]| \leq C \|(\mathbf{u}_h^n, \mathbf{z}_h^n, p_h^n)\|_A \|(\mathbf{v}_h, \mathbf{w}_h, q_h)\|_A.$$

We apply Babuska's theory [Babuška, 1971] to show well-posedness (existence and uniqueness) of this discretized system at a particular time step. This requires us to prove a discrete inf-sup type result (Theorem 5.2.2) for the combined bilinear form (5.23).

**Theorem 5.2.2** *Let  $\gamma > 0$  be a constant independent of any mesh parameters. Then the finite element formulation (5.6) satisfies the following discrete*

*inf-sup condition*

$$\gamma \|\|(\mathbf{u}_h^n, \mathbf{z}_h^n, p_h^n)\|\|_A \leq \sup_{(\mathbf{v}_h, \mathbf{w}_h, q_h) \in \mathcal{V}_h^X} \frac{B_h^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (\mathbf{v}_h, \mathbf{w}_h, q_h)]}{\|\|(\mathbf{v}_h, \mathbf{w}_h, q_h)\|\|_A} \quad \forall (\mathbf{u}_h, \mathbf{z}_h, p_h) \in \mathcal{W}_h^X. \quad (5.24)$$

Hence, given a solution at the previous time step the linear system arising from the fully discrete method for the subsequent time step is non-singular.

The following proof follows ideas presented by Burman and Hansbo [2007].

## Proof 2

**Step 1, bounding  $\|\mathbf{u}_h^n\|_{1,\Omega}$ ,  $\Delta t^{1/2}\|\mathbf{z}_h^n\|_{0,\Omega}$ , and  $|p_h^n|_{J,\Omega}$ .**

Choose  $(\mathbf{v}_h, \mathbf{w}_h, q_h) = (\mathbf{u}_h^n, \mathbf{z}_h^n, p_h^n)$ , then using (5.2) and (5.3), we obtain,

$$\begin{aligned} B_h^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (\mathbf{u}_h, \mathbf{z}_h, p_h)] &= a(\mathbf{u}_h^n, \mathbf{u}_h^n) + \Delta t(\kappa^{-1}\mathbf{z}_h^n, \mathbf{z}_h^n) + J(p_h^n, p_h^n) \\ &\geq C_k \|\mathbf{u}_h^n\|_{1,\Omega}^2 + \lambda_{max}^{-1} \Delta t \|\mathbf{z}_h^n\|_{0,\Omega}^2 + |p_h^n|_{J,\Omega}^2. \end{aligned} \quad (5.25)$$

**Step 2, bounding  $\|p_h^n\|_{0,\Omega}$ .**

Choose  $(\mathbf{v}_h, \mathbf{w}_h, q_h) = (\pi_h^1 \mathbf{v}_{p_h^n}, 0, 0)$  and add  $0 = \|p_h^n\|_{0,\Omega}^2 + (p_h^n, \nabla \cdot \mathbf{v}_{p_h^n})$  to obtain

$$B_h^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (\pi_h^1 \mathbf{v}_{p_h^n}, 0, 0)] = a(\mathbf{u}_h^n, \pi_h^1 \mathbf{v}_{p_h^n}) + \|p_h^n\|_{0,\Omega}^2 + (p_h^n, \nabla \cdot (\mathbf{v}_{p_h^n} - \pi_h^1 \mathbf{v}_{p_h^n})). \quad (5.26)$$

Focusing on the third term in (5.26) only, we apply the divergence theorem and split the integral over local elements to get

$$(p_h^n, \nabla \cdot (\mathbf{v}_{p_h^n} - \pi_h^1 \mathbf{v}_{p_h^n})) = \sum_K \int_{\partial K} p_h^n (\mathbf{v}_{p_h^n} - \pi_h^1 \mathbf{v}_{p_h^n}) \cdot \mathbf{n} \, ds = \sum_K \frac{1}{2} \int_{\partial K} [p_h^n] (\mathbf{v}_{p_h^n} - \pi_h^1 \mathbf{v}_{p_h^n}) \cdot \mathbf{n} \, ds.$$

We thus have

$$B_h^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (\pi_h^1 \mathbf{v}_{p_h^n}, 0, 0)] = \|p_h^n\|_{0,\Omega}^2 + a(\mathbf{u}_h^n, \pi_h^1 \mathbf{v}_{p_h^n}) + \sum_K \frac{1}{2} \int_{\partial K} [p_h^n] (\mathbf{v}_{p_h^n} - \pi_h^1 \mathbf{v}_{p_h^n}) \cdot \mathbf{n} \, ds.$$

Now first applying the Cauchy-Schwarz inequality and (5.1) on the right hand side to get

$$\begin{aligned} B_h^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (\pi_h^1 \mathbf{v}_{p_h^n}, 0, 0)] &\geq \|p_h^n\|_{0,\Omega}^2 - C_c \|\mathbf{u}_h^n\|_{1,\Omega} \|\pi_h^1 \mathbf{v}_{p_h^n}\|_{1,\Omega} \\ &\quad - \sum_K \frac{1}{2} \left( \int_{\partial K} (h^{1/2} [p_h^n])^2 \, ds \right)^{1/2} \cdot \left( \int_{\partial K} (h^{-1/2} (\mathbf{v}_{p_h^n} - \pi_h^1 \mathbf{v}_{p_h^n}) \cdot \mathbf{n})^2 \, ds \right)^{1/2}. \end{aligned}$$

Now apply Young's inequality and (5.14) to obtain

$$\begin{aligned} B_h^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (\pi_h^1 \mathbf{v}_{p_h^n}, 0, 0)] &\geq \|p_h^n\|_{0,\Omega}^2 - \frac{C_c^2}{2\epsilon} \|\mathbf{u}_h^n\|_{1,\Omega}^2 - \frac{\epsilon \hat{C}}{2} \|p_h^n\|_{0,\Omega}^2 \\ &\quad - \frac{1}{2\epsilon\delta} J(p_h^n, p_h^n) - \frac{\epsilon}{2} \sum_K \int_{\partial K} h^{-1} |(\mathbf{v}_{p_h^n} - \pi_h^1 \mathbf{v}_{p_h^n}) \cdot \mathbf{n}|^2 \, ds. \end{aligned}$$

Applying (5.18) we obtain

$$B_h^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (\pi_h^1 \mathbf{v}_{p_h^n}, 0, 0)] \geq -\frac{C_c^2}{2\epsilon} \|\mathbf{u}_h^n\|_{1,\Omega}^2 + \left(1 - (\hat{c} + c_t) \frac{\epsilon}{2}\right) \|p_h^n\|_{0,\Omega}^2 - \frac{1}{2\epsilon\delta} |p_h^n|_{J,\Omega}^2. \quad (5.27)$$

**Step 3, bounding  $\Delta t \|\nabla \cdot \mathbf{z}_h^n\|_{0,\Omega}$ .**

Choosing  $(\mathbf{v}_h, \mathbf{w}_h, q_h) = (0, 0, \Delta t \nabla \cdot \mathbf{z}_h^n)$  yields

$$B_h^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (0, 0, \Delta t \nabla \cdot \mathbf{z}_h^n)] = (\nabla \cdot \mathbf{u}_h^n, \Delta t \nabla \cdot \mathbf{z}_h^n) + \Delta t^2 \|\nabla \cdot \mathbf{z}_h^n\|_{0,\Omega}^2 + J(p_h^n, \Delta t \nabla \cdot \mathbf{z}_h^n).$$

We bound the first term using the Cauchy-Schwarz inequality followed by Young's inequality such that

$$(\nabla \cdot \mathbf{u}_h^n, \Delta t \nabla \cdot \mathbf{z}_h^n) \leq \frac{C_p}{2\epsilon} \|\mathbf{u}_h^n\|_{1,\Omega}^2 + \frac{\epsilon \Delta t^2}{2} \|\nabla \cdot \mathbf{z}_h^n\|_{0,\Omega}^2.$$

We can also bound the third term as before using the Cauchy-Schwarz inequality followed by Young's inequality such that

$$\begin{aligned} J(p_h^n, \Delta t \nabla \cdot \mathbf{z}_h^n) &\leq \frac{1}{2\epsilon} J(p_h^n, p_h^n) + \frac{\epsilon \Delta t^2}{2} J(\nabla \cdot \mathbf{z}_h^n, \nabla \cdot \mathbf{z}_h^n) = \frac{1}{2\epsilon} J(p_h^n, p_h^n) + \epsilon \delta \Delta t^2 \sum_K \int_{\partial K} |h^{1/2} \nabla \cdot \mathbf{z}_h^n|^2 ds \\ &\leq \frac{1}{2\epsilon} J(p_h^n, p_h^n) + \epsilon \delta c_z \Delta t^2 \|\nabla \cdot \mathbf{z}_h^n\|_{0,\Omega}^2. \end{aligned} \quad (5.28)$$

Where we have used the fact that  $\nabla \cdot \mathbf{z}_h^n \in Q_h$  in conjunction with (5.9). This

yields

$$B_h^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (0, 0, \Delta t \nabla \cdot \mathbf{z}_h^n)] \geq (1 - \epsilon \delta c_z - \frac{\epsilon}{2}) \Delta t^2 \|\nabla \cdot \mathbf{z}_h^n\|_{0,\Omega}^2 - \frac{1}{2\epsilon} |p_h^n|_{J,\Omega}^2 - \frac{C_p}{2\epsilon} \|\mathbf{u}_h^n\|_{1,\Omega}^2. \quad (5.29)$$

**Combining all the steps.** Finally we can combine (5.25), (5.27) and (5.29) to get control over all the norms by choosing  $(\mathbf{v}_h, \mathbf{w}_h, q_h) = (\beta \mathbf{u}_h^n + \pi_h^1 \mathbf{v}_{p_h^n}, \beta \mathbf{z}_h^n, \beta p_h^n + \Delta t \nabla \cdot \mathbf{z}_h^n)$ , which yields

$$\begin{aligned} B_h^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (\beta \mathbf{u}_h^n + \pi_h^1 \mathbf{v}_{p_h^n}, \beta \mathbf{z}_h^n, \beta p_h^n + \Delta t \nabla \cdot \mathbf{z}_h^n)] &\geq \\ (\beta C_k - \frac{C_c^2 + C_p}{2\epsilon}) \|\mathbf{u}_h^n\|_{1,\Omega}^2 + \beta \lambda_{max}^{-1} \Delta t \|\mathbf{z}_h^n\|_{0,\Omega}^2 + \left(1 - \epsilon \delta c_z - \frac{\epsilon}{2}\right) \Delta t^2 \|\nabla \cdot \mathbf{z}_h^n\|_{0,\Omega}^2 \\ + \left(1 - (\hat{c} + c_t) \frac{\epsilon}{2}\right) \|p_h^n\|_{0,\Omega}^2 + \left(\beta - \frac{1}{2\epsilon} - \frac{1}{2\epsilon\delta}\right) |p_h^n|_{J,\Omega}^2, \end{aligned} \quad (5.30)$$

where we can choose

$$\beta \geq \max \left[ \frac{C_c^2 + C_p}{2\epsilon C_k} + \frac{1 - \bar{C}\epsilon}{C_k}, \lambda_{max} (1 - \bar{C}\epsilon), \frac{1}{2\epsilon} + \frac{1}{2\epsilon\delta} + 1 - \bar{C}\epsilon \right],$$

with  $\bar{C} = \max \left[ \frac{\hat{c} + c_t}{2}, \delta c_z - \frac{1}{2} \right]$ . This yields

$$B_h^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (\beta \mathbf{u}_h^n + \pi_h^1 \mathbf{v}_{p_h^n}, \beta \mathbf{z}_h^n, \beta p_h^n + \nabla \cdot \mathbf{z}_h^n)] \geq (1 - \bar{C}\epsilon) \|\|(\mathbf{u}_h^n, \mathbf{z}_h^n, p_h^n)\|\|_A^2.$$

To complete the proof we need to take  $\epsilon$  sufficiently small and show that there exists a constant  $C$  such that  $\|\|(\mathbf{u}_h^n, \mathbf{z}_h^n, p_h^n)\|\|_A \geq C \|\|(\mathbf{v}_h, \mathbf{w}_h, q_h)\|\|_A$ , with  $(\mathbf{v}_h, \mathbf{w}_h, q_h) = (\beta \mathbf{u}_h^n + \pi_h^1 \mathbf{v}_{p_h^n}, \beta \mathbf{z}_h^n, \beta p_h^n + \Delta t \nabla \cdot \mathbf{z}_h^n)$ . Using the triangle

inequality and (5.14) we obtain

$$\begin{aligned} \left\| (\beta \mathbf{u}_h^n + \pi_h^1 \mathbf{v}_{p_h^n}, \beta \mathbf{z}_h^n, \beta p_h^n + \Delta t \nabla \cdot \mathbf{z}_h^n) \right\|_A^2 &\leq C \left( \beta^2 \|\mathbf{u}_h^n\|_{1,\Omega}^2 + \|\pi_h^1 \mathbf{v}_{p_h^n}\|_{1,\Omega}^2 + \Delta t^2 (1+\beta)^2 \|\nabla \cdot \mathbf{z}_h^n\|_{0,\Omega}^2 \right. \\ &\quad \left. + \beta^2 \Delta t \|\mathbf{z}_h^n\|_{0,\Omega}^2 + \beta^2 \|p_h^n\|_{0,\Omega}^2 + \beta^2 |p_h^n|_{J,\Omega}^2 + \Delta t^2 |\nabla \cdot \mathbf{z}_h^n|_{J,\Omega}^2 \right) \\ &\leq C \|(\mathbf{u}_h^n, \mathbf{z}_h^n, p_h^n)\|_A^2, \end{aligned}$$

as desired.

### 5.2.3 Energy estimate of the fully-discrete model

In this section we construct two new combined bilinear forms,  $B_{\delta t,h}^n$  (lemmas 5.2.3 and 5.2.4) and  $\mathcal{B}_h^n$  (lemmas 5.2.5 and 5.2.6). Lemmas 5.2.3 and 5.2.4 respectively develop a bound on  $B_{\delta t,h}^n$ , and an upper energy bound on the combined finite element formulation (5.31) with specially chosen test functions which combine to bound the displacements and pressure at time  $T = N\Delta t$  in the  $H^1$ -norm and  $J$ -norm respectively, and the pressures and fluxes in  $\Omega \times (0, T]$  in the  $L^2(L^2)$  norm. Lemmas 5.2.5 and 5.2.6 respectively develop a lower bound and an upper energy bound on (5.31) with specially chosen test functions which together bound the divergence of the flux in  $\Omega \times (0, T]$  in the  $L^2(L^2)$  norm. Since  $N$  is arbitrary, lemmas 5.2.4 and 5.2.6 can then be combined to bound the displacements and pressures in  $l^\infty$  in  $(0, T]$  and the  $H^1$  and  $J$ -norms respectively in  $\Omega$ . Pressure, flux and the divergence of the flux are also all bounded in  $\Omega \times (0, T]$  in the  $L^2(L^2)$  norm.

Adding (5.6a), (5.6b) and (5.6c), and assuming  $p_D = 0$  on  $\Gamma_p$  and  $\mathbf{t}_N = 0$

on  $\Gamma_t$ , we get the following

$$B_{\delta t,h}^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (\mathbf{v}_h, \mathbf{w}_h, q_h)] = (\mathbf{f}^n, \mathbf{v}_h) + (\mathbf{b}^n, \mathbf{w}_h) + (g^n, q_h) \quad \forall (\mathbf{v}_h, \mathbf{w}_h, q_h) \in \mathcal{W}_h^X, \quad (5.31)$$

where

$$\begin{aligned} B_{\delta t,h}^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (\mathbf{v}_h, \mathbf{w}_h, q_h)] &= a(\mathbf{u}_h^n, \mathbf{v}_h) + (\kappa^{-1} \mathbf{z}_h^n, \mathbf{w}_h) - (p_h^n, \nabla \cdot \mathbf{v}_h) - (p_h^n, \nabla \cdot \mathbf{w}_h) \\ &\quad + (\nabla \cdot \mathbf{u}_{\delta t,h}^n, q_h) + (\nabla \cdot \mathbf{z}_h^n, q_h) + J(p_{\delta t,h}^n, q_h). \end{aligned} \quad (5.32)$$

+

**Lemma 5.2.3**  $(\mathbf{u}_h, \mathbf{z}_h, p_h)$  satisfies

$$\begin{aligned} \sum_{n=1}^N \Delta t B_{\delta t,h}^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (\mathbf{u}_{\delta t,h}^n + \pi_h^1 \mathbf{v}_{p_h^n}, \mathbf{z}_h^n, p_h^n)] + \|\mathbf{u}_h^0\|_{1,\Omega}^2 + |p_h^0|_{J,\Omega}^2 + \|\mathbf{u}_h\|_{L^2(H^1)}^2 + \|p_h\|_{L^2(J)}^2 \geq \\ \left( \|\mathbf{u}_h^N\|_{1,\Omega}^2 + |p_h^N|_{J,\Omega}^2 + \|\mathbf{z}_h\|_{L^2(L^2)}^2 + \|p_h\|_{L^2(L^2)}^2 \right). \end{aligned}$$

**Proof 3** For  $n = 1, 2, \dots, N$  we choose  $(\mathbf{v}_h, \mathbf{w}_h, q_h) = (\mathbf{u}_{\delta t,h}^n + \pi_h^1 \mathbf{v}_{p_h^n}, \mathbf{z}_h^n, p_h^n)$  in (5.32), multiplying by  $\Delta t$ , and summing over all time steps, we get

$$\begin{aligned} \sum_{n=1}^N \Delta t B_{\delta t,h}^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (\mathbf{u}_{\delta t,h}^n + \pi_h^1 \mathbf{v}_{p_h^n}, \mathbf{z}_h^n, p_h^n)] &= \sum_{n=1}^N \Delta t a(\mathbf{u}_h^n, \mathbf{u}_{\delta t,h}^n) + \sum_{n=1}^N \Delta t J(p_{\delta t,h}^n, p_h^n) + \sum_{n=1}^N \Delta t \kappa^{-1} \\ &\quad + \sum_{n=1}^N \Delta t a(\mathbf{u}_h^n, \pi_h^1 \mathbf{v}_{p_h^n}) - \sum_{n=1}^N \Delta t (p_h^n, \nabla \cdot \pi_h^1 \mathbf{v}_{p_h^n}). \end{aligned} \quad (5.33)$$

By telescoping out the first two terms on the righthand side, using (5.3) on

the third, and applying (5.33) to the final two terms we obtain the inequality

$$\begin{aligned} & \sum_{n=0}^N \Delta t B_{\delta t, h}^n [(\mathbf{u}_h, \mathbf{z}_h, p_h), (\mathbf{u}_{\delta t, h}^n + \pi_h^1 \mathbf{v}_{p_h^n}, \mathbf{z}_h^n, p_h^n)] + \frac{C_c}{2} \|\mathbf{u}_h^0\|_{1, \Omega}^2 + \frac{C_c}{2\epsilon} \|\mathbf{u}_h\|_{L^2(H^1)}^2 + \frac{1}{4\epsilon} \|p_h\|_{L^2(J)}^2 + \frac{1}{2} |p_h^0|_{J, \Omega}^2 \\ & \geq \frac{C_k}{2} \|\mathbf{u}_h^N\|_{1, \Omega}^2 + \frac{1}{2} |p_h^N|_{J, \Omega}^2 + \lambda_{max}^{-1} \|\mathbf{z}_h\|_{L^2(L^2)}^2 + (1 - C\epsilon) \|p_h\|_{L^2(L^2)}^2. \end{aligned} \quad (5.34)$$

Finally, choosing  $\epsilon$  sufficiently small completes the proof.

**Lemma 5.2.4**  $(\mathbf{u}_h, \mathbf{z}_h, p_h)$  satisfies

$$\|\mathbf{u}_h^N\|_{1, \Omega}^2 + |p_h^N|_{J, \Omega}^2 + \|\mathbf{z}_h\|_{L^2(L^2)}^2 + \|p_h\|_{L^2(L^2)}^2 \leq C(T).$$

**Proof 4** For  $n = 1, 2, \dots, N$  we choose  $(\mathbf{v}_h, \mathbf{w}_h, q_h) = (\mathbf{u}_{\delta t, h}^n + \pi_h^1 \mathbf{v}_{p_h^n}, \mathbf{z}_h^n, p_h^n)$  in (5.40), multiplying by  $\Delta t$ , and summing yields

$$\begin{aligned} \sum_{n=1}^N \Delta t B_{\delta t, h}^n [(\mathbf{u}_h^n, \mathbf{z}_h^n, p_h^n), (\mathbf{u}_{\delta t, h}^n + \pi_h^1 \mathbf{v}_{p_h^n}, \mathbf{z}_h^n, p_h^n)] &= \sum_{n=1}^N \Delta t (\mathbf{f}^n, \mathbf{u}_{\delta t, h}^n + \pi_h^1 \mathbf{v}_{p_h^n}) \\ &\quad + \sum_{n=1}^N \Delta t (\mathbf{b}^n, \mathbf{z}_h^n) + \sum_{n=1}^N \Delta t (g^n, p_h^n). \end{aligned}$$

Let us note that the standard result, for any  $\epsilon > 0$

$$\sum_{n=1}^N \Delta t (\mathbf{f}^n, \mathbf{u}_{\delta t, h}^n) \leq C \left( \frac{1}{2\epsilon} \left( \|\mathbf{f}^0\|_{0, \Omega}^2 + \|\mathbf{f}^N\|_{0, \Omega}^2 + \|\mathbf{f}_t\|_{L^2(L^2)}^2 \right) + \frac{\epsilon}{2} \left( \|\mathbf{u}_h^0\|_{0, \Omega}^2 + \|\mathbf{u}_h^N\|_{0, \Omega}^2 + \|\mathbf{u}_h\|_{L^2(L^2)}^2 \right) \right) \quad (5.35)$$

Now using the above, lemma 5.2.3, the Cauchy-Schwarz and Young's inequality

ities, choosing  $\epsilon$  sufficiently small, and noting (5.14), we arrive at

$$\begin{aligned} \|\mathbf{u}_h^N\|_{1,\Omega}^2 + |p_h^N|_{J,\Omega}^2 + \|\mathbf{z}_h\|_{L^2(L^2)}^2 + \|p_h\|_{L^2(L^2)}^2 &\leq C \left( \|\mathbf{u}_h\|_{L^2(H^1)}^2 + \|p_h\|_{L^2(J)}^2 + \|\mathbf{f}^N\|_{0,\Omega}^2 \right. \\ &+ \|\mathbf{f}_t\|_{L^2(L^2)}^2 + \|\mathbf{u}_h^0\|_{0,\Omega}^2 + |p_h^0|_{J,\Omega}^2 + \|\mathbf{f}^0\|_{L^2(L^2)}^2 + \|\mathbf{f}\|_{L^2(L^2)}^2 + \|\mathbf{b}\|_{L^2(L^2)}^2 + \|g\|_{L^2(L^2)}^2 \left. \right). \end{aligned}$$

Using (5.5a), (5.5b), (5.5c) and (5.7) to bound the third term and upwards on the righthand side we obtain

$$\|\mathbf{u}_h^N\|_{1,\Omega}^2 + |p_h^N|_{J,\Omega}^2 + \|\mathbf{z}_h\|_{L^2(L^2)}^2 + \|p_h\|_{L^2(L^2)}^2 \leq C \left( 1 + \|\mathbf{u}_h\|_{L^2(H^1)}^2 + \|p_h\|_{L^2(J)}^2 \right).$$

Upon applying the Gronwall lemma to the above inequality we obtain the desired result.

To get a bound for the fluid flux in its natural  $H_{div}(\Omega)$  norm we now define the bilinear form  $\mathcal{B}_h^n$ . We first show how we derive  $\mathcal{B}_h^n$  from the fully-discrete weak form (5.6), for which we know that a solution  $(\mathbf{u}_h, \mathbf{z}_h, p_h)$  exists for test functions  $(\mathbf{v}_h, \mathbf{w}_h, q_h) \in \mathcal{V}_h^X$ . Adding (5.6a) and (5.6b), assuming  $p_D = 0$  on  $\Gamma_p$  and  $\mathbf{t}_N = 0$  on  $\Gamma_t$ , and summing we have

$$\begin{aligned} \sum_{n=1}^N a(\mathbf{u}_h^n, \mathbf{v}_h) + \sum_{n=1}^N (\kappa^{-1} \mathbf{z}_h^n, \mathbf{w}_h) - \sum_{n=1}^N (p_h^n, \nabla \cdot \mathbf{v}_h) - \sum_{n=1}^N (p_h^n, \nabla \cdot \mathbf{w}_h) \\ = \sum_{n=1}^N (\mathbf{f}^n, \mathbf{v}_h) + \sum_{n=1}^N (\mathbf{b}^n, \mathbf{w}_h) \quad \forall (\mathbf{v}_h, \mathbf{w}_h, q_h) \in \mathcal{V}_h^X. \quad (5.36) \end{aligned}$$

For the purposes of this proof we now introduce initial conditions for the

fluid flux and the pressure,  $\mathbf{z}^0 \in H_{div}(\Omega)$  and  $p^0 \in \mathcal{L}(\Omega)$  respectively. We also define their projections into their respective finite element spaces by  $\mathbf{z}_h^0 := \pi_h^0 \mathbf{z}^0$  and  $p_h^0 := \pi_h^0 p^0$ .

Adding (5.6a) and (5.6b), and summing from 0 to  $N - 1$ , we have

$$\begin{aligned} & \sum_{n=1}^N a(\mathbf{u}_h^{n-1}, \mathbf{v}_h) + \sum_{n=1}^N (\kappa^{-1} \mathbf{z}_h^{n-1}, \mathbf{w}_h) - \sum_{n=1}^N (p_h^{n-1}, \nabla \cdot \mathbf{v}_h) - \sum_{n=1}^N (p_h^{n-1}, \nabla \cdot \mathbf{w}_h) \\ &= \sum_{n=1}^N (\mathbf{f}^{n-1}, \mathbf{v}_h) + \sum_{n=1}^N (\mathbf{b}^{n-1}, \mathbf{w}_h) \quad \forall (\mathbf{v}_h, \mathbf{w}_h, q_h) \in \mathcal{V}_h^X. \end{aligned} \quad (5.37)$$

Taking (5.6c), multiplying by  $\Delta t$ , and summing we have

$$\sum_{n=1}^N \Delta t(\nabla \cdot \mathbf{u}_{\delta t,h}^n, q_h) + \sum_{n=1}^N \Delta t(\nabla \cdot \mathbf{z}_h^n, q_h) + \sum_{n=1}^N \Delta t J(p_{\delta t,h}^n, q_h) = \sum_{n=1}^N \Delta t(g^n, q_h) \quad \forall (\mathbf{v}_h, \mathbf{w}_h, q_h) \in \mathcal{V}_h^X. \quad (5.38)$$

Now adding (5.36) and (5.38), and subtracting (5.37) we get

$$\begin{aligned} \sum_{n=1}^N \Delta t \mathcal{B}_h^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (\mathbf{v}_h, \mathbf{w}_h, q_h)] &= \sum_{n=1}^N \Delta t(\mathbf{f}_{\delta t}^n, \mathbf{v}_h) + \sum_{n=1}^N \Delta t(\mathbf{b}_{\delta t}^n, \mathbf{w}_h) \\ &\quad + \sum_{n=1}^N \Delta t(g^n, q_h) \quad \forall (\mathbf{v}_h, \mathbf{w}_h, q_h) \in \mathcal{V}_h^X \end{aligned}$$

where

$$\begin{aligned} \mathcal{B}_h^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (\mathbf{v}_h, \mathbf{w}_h, q_h)] &= a(\mathbf{u}_{\delta t,h}^n, \mathbf{v}_h) + (\kappa^{-1} \mathbf{z}_{\delta t,h}^n, \mathbf{w}_h) - (p_{\delta t,h}^n, \nabla \cdot \mathbf{v}_h) - (p_{\delta t,h}^n, \nabla \cdot \mathbf{w}_h) \\ &\quad + (\nabla \cdot \mathbf{u}_{\delta t,h}^n, q_h) + (\nabla \cdot \mathbf{z}_h^n, q_h) + J(p_{\delta t,h}^n, q_h). \end{aligned} \quad (5.40)$$

**Lemma 5.2.5** For all  $\beta > \beta^* > 0$ ,  $(\mathbf{u}_h, \mathbf{z}_h, p_h)$  satisfies

$$\begin{aligned} & \sum_{n=1}^N \Delta t \mathcal{B}_h^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (\beta \mathbf{u}_{\delta t, h}^n + \pi_h^1 \mathbf{v}_p, \beta \mathbf{z}_h^n, \beta p_{\delta t, h}^n + \nabla \cdot \mathbf{z}_h^n)] + \|\mathbf{z}_h^0\|_{0, \Omega}^2 \geq \\ & C \left( \|\mathbf{u}_{\delta t, h}\|_{L^2(H^1)}^2 + \|\mathbf{z}_h^N\|_{0, \Omega}^2 + \|p_{\delta t, h}\|_{L^2(L^2)}^2 + \|p_{\delta t, h}\|_{L^2(J)}^2 + \|\nabla \cdot \mathbf{z}_h\|_{L^2(L^2)}^2 \right). \end{aligned}$$

**Proof 5** For  $n = 1, 2, \dots, N$  we choose  $(\mathbf{v}_h, \mathbf{w}_h, q_h) = (\beta \mathbf{u}_{\delta t, h}^n + \pi_h^1 \mathbf{v}_p, \beta \mathbf{z}_h^n, \beta p_{\delta t, h}^n + \nabla \cdot \mathbf{z}_h^n)$  in (5.39)

$$\begin{aligned} & \sum_{n=1}^N \Delta t \mathcal{B}_h^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (\beta \mathbf{u}_{\delta t, h}^n + \pi_h^1 \mathbf{v}_p, \beta \mathbf{z}_h^n, \beta p_{\delta t, h}^n + \nabla \cdot \mathbf{z}_h^n)] \\ & = \sum_{n=1}^N \Delta t a(\mathbf{u}_{\delta t, h}^n, \beta \mathbf{u}_{\delta t, h}^n) + \sum_{n=1}^N \Delta t \kappa^{-1}(\mathbf{z}_{\delta t, h}^n, \beta \mathbf{z}_h^n) + \sum_{n=1}^N \Delta t (\nabla \cdot \mathbf{z}_h^n, \nabla \cdot \mathbf{z}_h^n) + \sum_{n=1}^N \Delta t (\mathbf{u}_{\delta t, h}^n, \nabla \cdot \mathbf{z}_h^n) \\ & + \sum_{n=1}^N \Delta t J(p_{\delta t, h}^n, \nabla \cdot \mathbf{z}_h^n) + \sum_{n=1}^N \Delta t J(p_{\delta t, h}^n, \beta p_{\delta t, h}^n) + \sum_{n=1}^N \Delta t a(\mathbf{u}_{\delta t, h}^n, \pi_h^1 \mathbf{v}_p) - \sum_{n=1}^N \Delta t (p_{\delta t, h}^n, \nabla \cdot \pi_h^1 \mathbf{v}_p). \end{aligned} \tag{5.41}$$

For all  $\epsilon > 0$  using (5.2), (5.3), the Cauchy-Schwarz, Young's and Poincaré inequalities, (5.9) on  $\nabla \cdot \mathbf{z}_h^n$ , and an approach similar to step 2 in the proof of Theorem 5.2.2 for the final two terms on the righthand side, we obtain

$$\begin{aligned} & \sum_{n=1}^N \Delta t \mathcal{B}_h^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (\beta \mathbf{u}_{\delta t, h}^n + \pi_h^1 \mathbf{v}_p, \beta \mathbf{z}_h^n, \beta p_{\delta t, h}^n + \nabla \cdot \mathbf{z}_h^n)] \geq \left( \beta C_k - \frac{C_p + C_c}{2\epsilon} \right) \|\mathbf{u}_{\delta t, h}\|_{L^2(H^1)}^2 \\ & + \frac{\beta \lambda_{max}^{-1}}{2} \|\mathbf{z}_h^N\|_{0, \Omega}^2 + \left( \beta - \frac{3}{4\epsilon} \right) \|p_{\delta t, h}\|_{L^2(J)}^2 + (1 - \epsilon(1 + c_z)) \|\nabla \cdot \mathbf{z}_h\|_{L^2(L^2)}^2 \\ & - \frac{\beta \lambda_{min}^{-1}}{2} \|\mathbf{z}_h^0\|_{0, \Omega}^2 + (1 - C\epsilon) \|p_{\delta t, h}\|_{L^2(L^2)}^2. \end{aligned} \tag{5.42}$$

Finally choosing  $\epsilon$  sufficiently small and  $\beta \geq \max \left[ \frac{C_p}{2C_k\epsilon}, \frac{3}{4\epsilon} \right]$  completes the proof.

The following Lemma shows the divergence control of the fluid flow.

**Lemma 5.2.6**  $\mathbf{z}_h$  obtained from (5.39) satisfies

$$\|\nabla \cdot \mathbf{z}_h\|_{L^2(L^2)}^2 \leq C.$$

**Proof 6** For  $n = 1, 2, \dots, N$  we choose  $(\mathbf{v}_h, \mathbf{w}_h, q_h) = (\beta \mathbf{u}_{\delta t,h}^n + \pi_h^1 \mathbf{v}_{p_h^n}, \beta \mathbf{z}_h^n, \beta p_{\delta t,h}^n + \nabla \cdot \mathbf{z}_h^n)$  in (5.39) yielding

$$\begin{aligned} \sum_{n=1}^N \Delta t \mathcal{B}_h^n[(\mathbf{u}_h^n, \mathbf{z}_h^n, p_h^n), (\beta \mathbf{u}_{\delta t,h}^n + \pi_h^1 \mathbf{v}_{p_h^n}, \mathbf{z}_h^n, \beta p_{\delta t,h}^n + \nabla \cdot \mathbf{z}_h^n)] &= \sum_{n=1}^N \Delta t(\mathbf{f}_{\delta t}^n, \beta \mathbf{u}_{\delta t,h}^n + \pi_h^1 \mathbf{v}_{p_h^n}) + \sum_{n=1}^N \Delta t(\mathbf{b}_t^n, \\ &\quad + \sum_{n=1}^N \Delta t(g^n, \beta p_{\delta t,h}^n + \nabla \cdot \mathbf{z}_h^n). \end{aligned}$$

Using lemma 5.2.5, the Cauchy-Schwarz and Young's inequalities, and (5.14), along with ideas already presented in the proof of lemma 5.2.4

$$\begin{aligned} \|\mathbf{u}_{\delta t,h}\|_{L^2(H^1)}^2 + \|p_{\delta t,h}\|_{L^2(L^2)}^2 + \|p_{\delta t,h}\|_{L^2(J)}^2 + \|\mathbf{z}_h^N\|_{0,\Omega}^2 + \|\nabla \cdot \mathbf{z}_h\|_{L^2(L^2)}^2 \\ \leq C \left( \|\mathbf{f}_t\|_{L^2(L^2)}^2 + \|\mathbf{b}_t\|_{L^2(L^2)}^2 + \|p_h\|_{L^2(L^2)}^2 + \|\mathbf{z}_h\|_{L^2(L^2)}^2 + \|g\|_{L^2(L^2)}^2 \right). \end{aligned}$$

Finally, using Lemma 5.2.4 to bound  $\|p_h\|_{L^2(L^2)}$ , applying a Gronwall lemma, and using the regularity assumptions (5.5a), (5.5b) and (5.5c), we obtain the desired result.

**Theorem 5.2.7** *The solution to the fully-discrete problem (5.6) satisfies the energy estimate*

$$\|\mathbf{u}_h\|_{L^\infty(H^1)}^2 + \|p_h\|_{L^\infty(J)}^2 + \|\mathbf{z}_h\|_{L^2(L^2)}^2 + \|p_h\|_{L^2(L^2)}^2 + \|\nabla \cdot \mathbf{z}_h\|_{L^2(L^2)}^2 \leq C.$$

**Proof 7** *The proof follows from combining lemma 5.2.4 and lemma 5.2.6, and noting that these lemmas hold for all time steps  $n = 0, 1, \dots, N$ . This then gives the desired discrete in time  $l^\infty$  bounds.*

**Remark 2** *Having proven Theorem 5.2.7, it is now a standard calculation to show that the discrete Galerkin approximation converges weakly, as  $\Delta t, h \rightarrow 0$ , to the continuous problem with respect to continuous versions of the norms of the energy estimate in Theorem 5.2.7. This in turn shows that the continuous variational problem is well-posed. Due to the linearity of the variational form and noting that  $|\mathbf{v}|_{J,\Omega} \rightarrow 0$  as  $h \rightarrow 0$ , these calculations are straight forward and closely follow the existence and uniqueness proofs presented in Ženíšek [1984] and Barucq et al. [2005] for the linear two-field Biot problem and a nonlinear Biot problem, respectively.*

### 5.3 A-priori error analysis

We now derive an a-priori error estimate for the fully-discrete model. We start by introducing some notation, and then give a Galerkin orthogonality result which will form the corner stone of the error analysis. Lemma 5.3.1

provides a Galerkin orthogonality result obtained by comparing continuous and discrete weak forms. Lemma 5.2.1 is a standard approximation result for projections. Lemma 5.3.2 bounds the auxiliary errors for displacement, flux and pressure in the appropriate norms. Lemma 5.56 bounds the auxiliary error for the divergence of the flux. Since Lemmas 5.3.2 and 4.4 bound the auxiliary errors at the same order as the projection errors, combining projection and auxiliary errors in Theorem ?? provides an optimal error estimate.

We define the finite element error functions

$$\mathbf{e}_\mathbf{u} := \mathbf{u} - \mathbf{u}_h, , \quad \mathbf{e}_\mathbf{z} := \mathbf{z} - \mathbf{z}_h, \quad e_p := p - p_h.$$

We introduce the following projection errors:

$$\eta_\mathbf{u} := \mathbf{u} - \pi_h^1 \mathbf{u}, \quad \eta_\mathbf{z} := \mathbf{z} - \pi_h^1 \mathbf{z}, \quad \eta_p := p - \pi_h^0 p,$$

where we have assumed  $\mathbf{z}(t_n, \cdot) \in (H^1(\Omega))^d$ .

Auxiliary errors:

$$\theta_\mathbf{u}^n(\cdot) := \pi_h^1 \mathbf{u}(t_n, \cdot) - \mathbf{u}_h^n(\cdot), \quad \theta_\mathbf{z}^n(\cdot) := \pi_h^1 \mathbf{z}(t_n, \cdot) - \mathbf{z}_h^n(\cdot), \quad \theta_p^n(\cdot) := \pi_h^0 p(t_n, \cdot) - p_h^n(\cdot), \quad (5.43)$$

and time-discretization errors:

$$\rho_{\mathbf{u}}^n(\cdot) := \frac{\mathbf{u}(t_n, \cdot) - \mathbf{u}(t_{n-1}, \cdot)}{\Delta t} - \frac{\partial \mathbf{u}(t_n, \cdot)}{\partial t}, \quad \rho_p^n := \frac{p(t_n, \cdot) - p(t_{n-1}, \cdot)}{\Delta t} - \frac{\partial p(t_n, \cdot)}{\partial t}. \quad (5.44)$$

### 5.3.1 Galerkin orthogonality

We now give a Galerkin orthogonality type argument for analysing the difference between the fully-discrete approximation and the true solution. For this we introduce the continuous counterpart of the fully-discrete combined weak form (5.31) given by

$$B^n[(\mathbf{u}, \mathbf{z}, p), (\mathbf{v}, \mathbf{w}, q)] = (\mathbf{f}(t_n, \cdot), \mathbf{v}) + (\mathbf{b}(t_n), \mathbf{w}) + (g(t_n, \cdot), q) \quad \forall (\mathbf{v}, \mathbf{w}, q) \in \mathcal{V}^X$$

where

$$\begin{aligned} B^n[(\mathbf{u}, \mathbf{z}, p), (\mathbf{v}, \mathbf{w}, q)] &= a(\mathbf{u}(t_n, \cdot), \mathbf{v}) + \kappa^{-1}(\mathbf{z}(t_n, \cdot), \mathbf{w}) - (p(t_n, \cdot), \nabla \cdot \mathbf{v}) \\ &\quad - (p(t_n, \cdot), \nabla \cdot \mathbf{w}) + (\nabla \cdot \mathbf{u}_t(t_n, \cdot), q) + (\nabla \cdot \mathbf{z}(t_n, \cdot), q) \end{aligned}$$

**Lemma 5.3.1** *Assuming  $(\mathbf{u}(t_n, \cdot), \mathbf{z}(t_n, \cdot), p(t_n, \cdot)) \in [H^1(\Omega)]^d \times H_{div}(\Omega) \times (H^1(\Omega) \cap \mathcal{L}(\Omega))$*

$$B_{\delta t, h}^n[(\mathbf{e}_{\mathbf{u}}, \mathbf{e}_{\mathbf{z}}, e_p), (\mathbf{v}_h, \mathbf{w}_h, q_h)] = (\nabla \cdot \rho_{\mathbf{u}}^n, q_h) + J(\rho_p^n, q_h) \quad \forall (\mathbf{v}_h, \mathbf{w}_h, q_h) \in \mathcal{V}_h^X.$$

**Proof 8** *Subtracting the discrete weak form (5.31) from the continuous weak*

form (5.45), we obtain

$$B^n[(\mathbf{u}, \mathbf{z}, p), (\mathbf{v}_h, \mathbf{w}_h, q_h)] - B_{\delta t, h}^n[(\mathbf{u}_h, \mathbf{z}_h, p_h), (\mathbf{v}_h, \mathbf{w}_h, q_h)] = 0, \quad \forall (\mathbf{v}_h, \mathbf{w}_h, q_h) \in \mathcal{V}_h^X.$$

Now add  $J(p_t(t_n, \cdot), q) = 0$  to the left hand side, see (5.10). Finally add  $(\nabla \cdot (\mathbf{u}_{\delta t}(t_n, \cdot) - \mathbf{u}_t(t_n, \cdot)), q) + J(p_{\delta t}(t_n, \cdot) - p_t(t_n, \cdot), q)$  to the left and the righthand side to obtain the desired result.

### 5.3.2 Auxiliary error estimates

**Lemma 5.3.2** Assuming  $\mathbf{u} \in H^2(0, T; [L^2(\Omega)]^d) \cap H^1(0, T; [H^2(\Omega)]^d)$ ,  $\mathbf{z} \in L^2(0, T; [H^1(\Omega)]^d)$  and  $p \in H^2(0, T; H^1(\Omega) \cap \mathcal{L}(\Omega))$ , then the finite element solution (5.6) satisfies the error estimate

$$\|[\theta_{\mathbf{u}}, \theta_{\mathbf{z}}, \theta_p]\|_B^2 + \|\theta_p\|_{L^\infty(J)}^2 \leq C(T)(h^2 + \Delta t^2). \quad (5.47)$$

**Proof 9** Using Lemma 5.3.1 and choosing  $\mathbf{v}_h^n = \theta_{\delta t, \mathbf{u}}^n + \pi_h^1 \mathbf{v}_{p_h^n}$ ,  $\mathbf{w}_h^n = \theta_{\mathbf{z}}^n$ ,  $q_h^n = \theta_p^n$ , we get

$$B_{\delta t, h}^n[(\theta_{\mathbf{u}}^n + \eta_{\mathbf{u}}^n, \theta_{\mathbf{z}}^n + \eta_{\mathbf{z}}^n, \theta_p^n + \eta_p^n), (\theta_{\delta t, \mathbf{u}}^n + \pi_h^1 \mathbf{v}_{p_h^n}, \theta_{\mathbf{z}}^n, \theta_p^n)] = (\nabla \cdot \rho_{\mathbf{u}}^n, \theta_p^n) + J(\rho_p^n, \theta_p^n)$$

Rearranging gives

$$\begin{aligned} B_{\delta t,h}^n[(\theta_{\mathbf{u}}^n, \theta_{\mathbf{z}}^n, \theta_p^n), (\theta_{\delta t,\mathbf{u}}^n + \pi_h^1 \mathbf{v}_{p_h^n}, \theta_{\mathbf{z}}^n, \theta_p^n)] &= (\nabla \cdot \rho_{\mathbf{u}}^n, \theta_p^n) + J(\rho_p^n, \theta_p^n) \\ &\quad - B_{\delta t,h}^n[(\eta_{\mathbf{u}}^n, \eta_{\mathbf{z}}^n, \eta_p^n), (\theta_{\delta t,\mathbf{u}}^n + \pi_h^1 \mathbf{v}_{p_h^n}, \theta_{\mathbf{z}}^n, \theta_p^n)] \end{aligned}$$

Expanding the righthand side, noting that  $(\eta_p^n, \nabla \cdot (\theta_{\delta t,\mathbf{u}}^n + \pi_h^1 \mathbf{v}_p)) = 0$ ,  $(\eta_p^n, \nabla \cdot \theta_{\mathbf{z}}^n) = 0$ , multiplying both sides by  $\Delta t$  and summing gives

$$\sum_{n=1}^N \Delta t B_{\delta t,h}^n[(\theta_{\mathbf{u}}^n, \theta_{\mathbf{z}}^n, \theta_p^n), (\theta_{\delta t,\mathbf{u}}^n + \pi_h^1 \mathbf{v}_{p_h^n}, \theta_{\mathbf{z}}^n, \theta_p^n)] = \Phi_1 + \Phi_2 + \Phi_3 + \Phi_4 + \Phi_5 + \Phi_6 \quad (\text{5.48})$$

where

$$\begin{aligned} \Phi_1 &:= - \sum_{n=1}^N \Delta t a(\eta_{\mathbf{u}}^n, \theta_{\delta t,\mathbf{u}}^n), & \Phi_2 &:= - \sum_{n=1}^N \Delta t (\kappa^{-1}(\eta_{\mathbf{z}}^n, \theta_{\mathbf{z}}^n)), & \Phi_3 &:= - \sum_{n=1}^N \Delta t a(\eta_{\mathbf{u}}^n, \pi_h^1 \mathbf{v}_p), \\ \Phi_4 &:= - \sum_{n=1}^N \Delta t J(\eta_{\delta t,p}^n, \theta_p^n), & \Phi_5 &:= \sum_{n=1}^N \Delta t (\nabla \cdot \rho_{\mathbf{u}}^n, \theta_p^n), & \Phi_6 &:= \sum_{n=1}^N \Delta t J(\rho_p^n, \theta_p^n), \\ \Phi_7 &:= - \sum_{n=1}^N \Delta t (\theta_p^n, \nabla \cdot (\eta_{\delta t,\mathbf{u}}^n + \eta_{\mathbf{z}}^n)). \end{aligned}$$

We now individually consider the terms on the right hand side of (5.48):

To bound the first quantity, we use (5.19), Lemma 5.2.1, the triangle,

Cauchy-Schwarz and Young's inequalities,  $\theta_{\mathbf{u}}^0 = 0$ , and (5.1),

$$\begin{aligned}
\Phi_1 &= - \sum_{n=1}^N a(\eta_{\mathbf{u}}^n, \theta_{\mathbf{u}}^n - \theta_{\mathbf{u}}^{n-1}) \\
&= -a(\eta_{\mathbf{u}}^N, \theta_{\mathbf{u}}^N) + \sum_{n=1}^N a(\eta_{\mathbf{u}}^n - \eta_{\mathbf{u}}^{n-1}, \theta_{\mathbf{u}}^{n-1}) \\
&= -a(\eta_{\mathbf{u}}^N, \theta_{\mathbf{u}}^N) + \Delta t \sum_{n=1}^N a \left( (I - \pi_h^1) \left( \rho_{\mathbf{u}}^n + \frac{\partial \mathbf{u}(t_n, \cdot)}{\partial t} \right), \theta_{\mathbf{u}}^{n-1} \right) \\
&\leq \epsilon C \|\theta_{\mathbf{u}}^N\|_{1,\Omega}^2 + \frac{Ch^2}{\epsilon} \|\mathbf{u}^N\|_{2,\Omega}^2 + \epsilon C \|\theta_{\mathbf{u}}\|_{L^2(H^1)}^2 + \frac{Ch^2}{2\epsilon} \|\mathbf{u}_t\|_{L^2(H^2)}^2 + \frac{C\Delta t^2}{2\epsilon} \|\mathbf{u}_{tt}\|_{L^2(H^1)}^2.
\end{aligned} \tag{5.49}$$

Next, using (5.3), Young's inequality, (5.14) and Lemma 5.2.1,

$$\Phi_2 \leq \frac{\epsilon}{2} \|\theta_{\mathbf{z}}\|_{L^2(L^2)}^2 + \frac{\lambda_{\min}^{-2} h^2}{2\epsilon} \|\mathbf{z}\|_{L^2(H^1)}^2. \tag{5.50}$$

Using (5.1), Young's inequality and Lemma 5.2.1,

$$\Phi_3 \leq \frac{\epsilon}{2} \|\pi_h^1 \mathbf{v}_{p_h^n}\|_{L^2(H^1)}^2 + \frac{C}{2\epsilon} \|\eta_{\mathbf{u}}\|_{L^2(H^1)}^2 \leq \frac{\epsilon \hat{c}^2}{2} \|\theta_p\|_{L^2(L^2)}^2 + \frac{Ch^2}{2\epsilon} \|\mathbf{u}\|_{L^2(H^2)}^2. \tag{5.51}$$

The bound on  $\Phi_4$  is obtained using a similar argument to the bound on  $\Phi_1$ ,

$$\Phi_4 \leq \epsilon \|\theta_p\|_{L^2(J)}^2 + \frac{h^2}{2\epsilon} \|p_t\|_{L^2(H^1)}^2 + \frac{\Delta t^2}{2\epsilon} \|p_{tt}\|_{L^2(H^1)}^2. \tag{5.52}$$

Using the Cauchy-Schwarz and Young's inequalities and lemma 5.2.1,

$$\Phi_5 \leq \frac{\epsilon}{2} \|\theta_p\|_{L^2(L^2)}^2 + \frac{\Delta t^2}{2\epsilon} \|\mathbf{u}_{tt}\|_{L^2(L^2)}^2 \text{ and } \Phi_6 \leq \frac{\epsilon}{2} \|\theta_p\|_{L^2(J)}^2 + \frac{\Delta t^2}{2\epsilon} \|p_{tt}\|_{L^2(J)}^2 \quad (5.53)$$

Finally, using the Cauchy-Schwarz and Young's inequalities, and a similar argument to the bound on  $\Phi_1$ ,

$$\Phi_7 \leq \frac{3\epsilon}{2} \|\theta_p\|_{L^2(L^2)}^2 + \frac{h^2}{2\epsilon} \|\mathbf{u}_t\|_{L^2(H^2)}^2 + \frac{\Delta t^2}{2\epsilon} \|\mathbf{u}_{tt}\|_{L^2(H^1)}^2 + \frac{h^2}{2\epsilon} \|\mathbf{z}\|_{L^2(H^2)}^2 \quad (5.54)$$

Combining these bounds with an application of coercivity Lemma 5.2.3 to (5.48), noting the assumed regularity of the continuous solution and choosing  $\epsilon$  sufficiently small, gives

$$\|\theta_{\mathbf{u}}^N\|_{1,\Omega}^2 + |\theta_p^N|_{J,\Omega}^2 + \|\theta_{\mathbf{z}}\|_{L^2(L^2)}^2 + \|\theta_p\|_{L^2(L^2)}^2 \leq C \left( \|\theta_{\mathbf{u}}\|_{L^2(H^1)}^2 + \|\theta_p\|_{L^2(J)}^2 + h^2 + \Delta t^2 \right). \quad (5.55)$$

An application of Gronwall's lemma gives

$$\|\theta_{\mathbf{u}}^N\|_{1,\Omega}^2 + |\theta_p^N|_{J,\Omega}^2 + \|\theta_{\mathbf{z}}\|_{L^2(L^2)}^2 + \|\theta_p\|_{L^2(L^2)}^2 \leq C(T) (h^2 + \Delta t^2).$$

Because the above holds for all time steps  $n = 0, 1, \dots, N$ , we can get the desired  $L^\infty$  bounds to complete the proof of the theorem.

We now present an a-priori auxillary error estimate of the fluid flux, in its natural  $H\text{div}$  norm.

**Lemma 5.3.3** Assuming  $\mathbf{u} \in H^2(0, T; [H^1(\Omega)]^d) \cap H^1(0, T; [H^2(\Omega)]^d)$ ,  $\mathbf{z} \in L^2(0, T; [H^2(\Omega)]^d)$  and  $p \in H^2(0, T; J \cap \mathcal{L}(\Omega)) \cap H^1(0, T; H^1(\Omega))$ , then the finite element solution (5.6) satisfies the auxillary error estimate

$$\|\nabla \cdot \theta_{\mathbf{z}}\|_{L^2(L^2)}^2 \leq C(T)(h^2 + \Delta t^2). \quad (5.56)$$

**Proof 10** Similarly to the approach taken in obtaining (5.39) we may easily obtain the following identity

$$\sum_{n=1}^N \Delta t \mathcal{B}_h^n[(\theta_{\mathbf{u}}^n, \theta_{\mathbf{z}}^n, \theta_p^n), (\beta \theta_{\delta t, \mathbf{u}}^n + \pi_h^1 \mathbf{v}_{\theta_{\delta t, p}^n}, \beta \theta_{\mathbf{z}}^n, \beta \theta_{\delta t, p}^n + \nabla \cdot \theta_{\mathbf{z}}^n)] = \Psi_1 + \Psi_2 + \Psi_3 + \Psi_4 + \Psi_5 + \Psi_6, \quad (5.57)$$

where

$$\begin{aligned} \Psi_1 &= - \sum_{n=1}^N \Delta t a(\eta_{\delta t, \mathbf{u}}^n, \beta \theta_{\delta t, \mathbf{u}}^n + \pi_h^1 \mathbf{v}_{\theta_{\delta t, p}^n}), & \Psi_2 &= - \sum_{n=1}^N \Delta t (\nabla \cdot (\eta_{\delta t, \mathbf{u}}^n + \eta_{\mathbf{z}}^n), \nabla \cdot \theta_{\mathbf{z}}^n + \beta \theta_{\delta t, p}^n), \\ \Psi_3 &= \sum_{n=1}^N \Delta t J(\eta_{\delta t, p}^n, \beta \theta_{\delta t, p}^n + \nabla \cdot \theta_{\mathbf{z}}^n), & \Psi_4 &= - \sum_{n=1}^N \Delta t (\kappa^{-1}(\eta_{\delta t, \mathbf{z}}^n, \beta \theta_{\mathbf{z}}^n)), \\ \Psi_5 &= \sum_{n=1}^N \Delta t J(\rho_p^n, \beta \theta_{\delta t, p}^n + \nabla \cdot \theta_{\mathbf{z}}^n), & \Psi_6 &= \sum_{n=1}^N \Delta t (\nabla \cdot \rho_{\mathbf{u}}^n, \beta \theta_{\delta t, p}^n + \nabla \cdot \theta_{\mathbf{z}}^n). \end{aligned}$$

We now bound the terms on the right hand side of (5.57) using machinery developed during the previous proof:

To bound the first quantity, we use (5.19), Lemma 5.2.1, the triangle, Cauchy-Schwarz and Young's inequalities,  $\theta_{\mathbf{u}}^0 = 0$ , (5.14), (5.9), (5.1), and

*Lemma 5.2.1,*

$$\Psi_1 \leq \frac{C\epsilon}{2} \|\theta_{\delta t, \mathbf{u}}\|_{L^2(H^1)}^2 + \frac{\hat{c}^2\epsilon}{2} \|\theta_{\delta t, p}\|_{L^2(L^2)}^2 + \frac{Ch^2}{2\epsilon} \|\mathbf{u}_t\|_{L^2(H^2)}^2 + \frac{C}{2\epsilon} \Delta t^2 \|\mathbf{u}_{tt}\|_{L^2(H^8)}^2$$

$$\Psi_2 \leq \epsilon \|\nabla \cdot \theta_{\mathbf{z}}\|_{L^2(L^2)}^2 + \epsilon \|\theta_{\delta t, p}\|_{L^2(L^2)}^2 + \frac{Ch^2}{2\epsilon} \left( \|\mathbf{u}_t\|_{L^2(H^2)}^2 + \|\mathbf{z}\|_{L^2(H^2)}^2 \right) + \frac{C}{2\epsilon} \Delta t^2 \|\mathbf{u}_{tt}\|_{L^2(H^9)}^2$$

$$\Psi_3 \leq \epsilon C \|\nabla \cdot \theta_{\mathbf{z}}\|_{L^2(L^2)}^2 + \epsilon \|\theta_{\delta t, p}^n\|_{L^2(J)}^2 + \frac{Ch^2}{2\epsilon} \|p_t\|_{L^2(H^1)}^2 + \frac{C}{2\epsilon} \Delta t^2 \|p_{tt}\|_{L^2(J)}^2$$

$$\Psi_4 \leq \epsilon \|\theta_{\mathbf{z}}\|_{L^2(L^2)}^2 + \frac{Ch^2}{2\epsilon} \|\mathbf{z}_t\|_{L^2(H^1)}^2 + \frac{C}{2\epsilon} \Delta t^2 \|\mathbf{z}_{tt}\|_{L^2(L^2)}^2, \quad (5.61)$$

$$\Psi_5 \leq \epsilon \|\theta_{\delta t, p}\|_{L^2(J)}^2 + \epsilon C \|\nabla \cdot \theta_{\mathbf{z}}\|_{L^2(L^2)}^2 + \frac{C\Delta t^2}{2\epsilon} \|p_{tt}\|_{L^2(J)}^2 \quad (5.62)$$

$$\Psi_6 \leq \epsilon \|\theta_{\delta t, p}\|_{L^2(L^2)}^2 + \epsilon \|\nabla \cdot \theta_{\mathbf{z}}\|_{L^2(L^2)}^2 + \frac{C}{2\epsilon} \Delta t^2 \|\mathbf{u}_{tt}\|_{L^2(H^1)}^2. \quad (5.63)$$

We can now combine the individual bounds (5.58), (5.59), (5.60), (5.61), (5.62), and (5.63), with the coercivity result Lemma 5.2.5, choosing  $\beta$  sufficiently large, use the assumption  $\theta_{\mathbf{z}}^0 = 0$ , the assumed regularity of  $\mathbf{u}, \mathbf{z}$  and  $p$ , and choosing  $\epsilon$  sufficiently small to obtain

$$\|\theta_{\mathbf{z}}^N\|_{0,\Omega}^2 + \|\nabla \cdot \theta_{\mathbf{z}}\|_{L^2(L^2)}^2 \leq C \|\theta_{\mathbf{z}}\|_{L^2(L^2)}^2 + C(h^2 + \Delta t^2).$$

Applying Gronwall's lemma, we get the desired result.

**Theorem 5.3.4** Assuming  $\mathbf{u} \in H^2(0, T; [L^2(\Omega)]^d) \cap H^1(0, T; [H^2(\Omega)]^d)$ ,  $\mathbf{z} \in L^2(0, T; [H^1(\Omega)]^d)$  and  $p \in H^2(0, T; H^1(\Omega) \cap \mathcal{L}(\Omega))$ , then the finite element solution (5.6) satisfies the error estimate

$$\|\mathbf{e}_\mathbf{u}, \mathbf{e}_\mathbf{z}, e_p\|_B^2 \leq C(h^2 + \Delta t^2).$$

Assuming  $\mathbf{u} \in H^2(0, T; [H^1(\Omega)]^d) \cap H^1(0, T; [H^2(\Omega)]^d)$ ,  $\mathbf{z} \in L^2(0, T; [H^2(\Omega)]^d)$  and  $p \in H^2(0, T; J \cap \mathcal{L}(\Omega)) \cap H^1(0, T; H^1(\Omega))$ , then the finite element solution (5.6) satisfies the error estimate

$$\|\mathbf{e}_\mathbf{u}, \mathbf{e}_\mathbf{z}, e_p\|_B^2 + \|\nabla \cdot \mathbf{e}_\mathbf{z}\|_{L^2(L^2)}^2 \leq C(h^2 + \Delta t^2).$$

**Proof 11** We first write the errors as  $\mathbf{e}_\mathbf{u}^n = \eta_\mathbf{u}^n + \theta_\mathbf{u}^n$ , and similarly for the other variables. Using lemma 5.2.1 we can bound the projection errors, and using lemma 5.3.2 and lemma 5.56 we can bound the auxillary errors to give the desired result.

## 5.4 Numerical Results

In this section, we present both two-dimensional and three-dimensional numerical experiments with analytical solutions to verify the theoretical convergence rates of the fully-discrete finite element method developed in this paper. We also test our method on the popular 2D cantilever bracket problem

to make sure our method is able to overcome the spurious pressure oscillations often experienced in poroelastic simulations. Finally, a 3D unconfined compression problem is presented that highlights the added mass effect of the method for different choices of the stabilization parameter  $\delta$ .

### 5.4.1 Implementation

For the implementation we used the C++ library libmesh [Kirk et al., 2006], and the multi-frontal direct solver mumps [Amestoy et al., 2000] to solve the resulting linear system. To solve the full Biot model problem (3.38), we need to solve the following linear system at each time step:

$$\begin{bmatrix} \mathbf{A} & 0 & \alpha\mathbf{B}^T \\ 0 & \mathbf{M} & \mathbf{B}^T \\ -\alpha\mathbf{B} & -\Delta t\mathbf{B} & c_0\mathbf{Q} + \mathbf{J} \end{bmatrix} \begin{bmatrix} \mathbf{u}^n \\ \mathbf{z}^n \\ \mathbf{p}^n \end{bmatrix} = \begin{bmatrix} \mathbf{r} \\ \mathbf{s} \\ \Delta t \mathbf{g} - \mathbf{B}\mathbf{u}^{n-1} + c_0\mathbf{Q}\mathbf{p}^{n-1} + \mathbf{J}\mathbf{p}^{n-1} \end{bmatrix},$$

where we have defined the following matrices and vectors:

$$\mathbf{A} = [\mathbf{a}_{ij}], \quad \mathbf{a}_{ij} = \int_{\Omega} 2\mu_s \nabla \phi_i : \nabla \phi_j + \lambda(\nabla \cdot \phi_i)(\nabla \cdot \phi_j),$$

$$\mathbf{M} = [\mathbf{m}_{ij}], \quad \mathbf{m}_{ij} = \int_{\Omega} \kappa^{-1} \phi_i \cdot \phi_j,$$

$$\mathbf{B} = [\mathbf{b}_{ij}], \quad \mathbf{b}_{ij} = - \int_{\Omega} \psi_i \nabla \cdot \phi_j,$$

$$\mathbf{Q} = [\mathbf{q}_{ij}], \quad \mathbf{q}_{ij} = \int_{\Omega} \psi_i \cdot \psi_j,$$

$$\mathbf{J} = [\mathbf{j}_{ij}], \quad \mathbf{j}_{ij} = \delta \sum_K \int_{\partial k \setminus \partial \Omega} h_{\partial K}[\psi_i][\psi_j] \, ds,$$

$$\mathbf{r} = [\mathbf{r}_i], \quad \mathbf{r}_i = \int_{\Omega} \mathbf{f}_i \cdot \phi_i + \int_{\Gamma_t} \mathbf{t}_{Ni} \cdot \phi_i,$$

$$\mathbf{s} = [\mathbf{s}_i], \quad \mathbf{s}_i = \int_{\Omega} \mathbf{b}_i \cdot \phi_i - \int_{\Gamma_p} p_D \phi_i \cdot \mathbf{n},$$

$$\mathbf{g} = [\mathbf{g}_i], \quad \mathbf{g}_i = \int_{\Omega} g \psi_i.$$

Here  $\phi_i$  are vector valued linear basis functions such that the displacement vector can be written as  $\mathbf{u}^n = \sum_{i=1}^{n_u} \mathbf{u}_i^n \phi_i$ , with  $\sum_{i=1}^{n_u} \mathbf{u}_i^n \phi_i \in \mathbf{W}_h^E$ . Similarly for the relative fluid vector we have  $\mathbf{z}^n = \sum_{i=1}^{n_z} \mathbf{z}_i^n \phi_i$ , with  $\sum_{i=1}^{n_z} \mathbf{z}_i^n \phi_i \in \mathbf{W}_h^D$ . The scalar valued constant basis functions  $\psi_i$  are used to approximate the pressure, such that  $\mathbf{p}^n = \sum_{i=1}^{n_p} p_i^n \psi_i$ , with  $\sum_{i=1}^{n_p} p_i^n \psi_i \in Q_h$ .

#### 5.4.2 2D convergence study

We adapt a test problem used in Burman and Hansbo [2007] to verify the convergence of our method. To simplify the analytical solution we chose parameters so that the equilibrium equations are simplified as below, yet retain all the analytical properties and difficulties of the original problem. Any conclusions drawn from this convergence study therefore also apply to the original system of equations (3.38). The analytical solution of the pressure is given by  $p = \sin(2\pi x) \sin(2\pi y) \sin(2\pi t)$ , with  $t \in [0, 0.25]$ . For this test problem the domain,  $\Omega$ , is the unit square. The simplified system of equations is

given by

$$-\nabla^2 \mathbf{u} + \nabla p = 0 \quad \text{in } \Omega, \quad (5.64\text{a})$$

$$\mathbf{z} + \nabla p = 0 \quad \text{in } \Omega, \quad (5.64\text{b})$$

$$\nabla \cdot (\mathbf{u}_t + \mathbf{z}) = g \quad \text{in } \Omega, \quad (5.64\text{c})$$

$$\mathbf{u}(t) = \mathbf{u}_D \quad \text{on } \Gamma_d, \quad (5.64\text{d})$$

$$\mathbf{z}(t) \cdot \mathbf{n} = q_D \quad \text{on } \Gamma_f, \quad (5.64\text{e})$$

$$\mathbf{u}(0) = 0 \quad \text{in } \Omega, \quad (5.64\text{f})$$

$$p(0) = 0 \quad \text{in } \Omega. \quad (5.64\text{g})$$

where the displacement boundary condition, the fluid flux boundary condition, and the source term are calculated to be

$$\mathbf{u}_D = \begin{pmatrix} -\frac{1}{4}\pi \cos(2\pi x) \sin(2\pi y) \sin(2\pi t) \\ -\frac{1}{4}\pi \sin(2\pi x) \cos(2\pi y) \sin(2\pi t) \end{pmatrix},$$

$$q_D = \begin{pmatrix} -2\pi \cos(2\pi x) \sin(2\pi y) \sin(2\pi t) \\ -2\pi \sin(2\pi x) \cos(2\pi y) \sin(2\pi t) \end{pmatrix} \cdot \mathbf{n},$$

$$g = 2\pi \sin(2\pi x) \sin(2\pi y) \cos(2\pi t) + 8\pi^2 \sin(2\pi x) \sin(2\pi y) \sin(2\pi t).$$

We also set the pressure to zero at  $[0, 0]$  to remove the null space. Figure 5.1a illustrates what happens when the stabilization parameter  $\delta$  is not chosen large enough ( $\delta = 0.1$ ), resulting in a spurious pressure solution. For this

test problem we have found that a value of  $\delta = 1$  is sufficiently large to result in a smooth pressure solution (see Figure 5.1b). The value of  $\delta$  required to produce a stable solution depends on the geometry and material parameters of the particular problem under investigation, but is independent of any mesh parameters. In Figure 5.2, we show the convergence of the method in relevant norms for each variable, with  $\delta = 1, 10, 100$ . The rates of convergence agree with the theoretically derived error estimates. Figure 5.3 and Figure 5.4 show the deformation, pressure, and fluid flux in the  $x$  and  $y$  directions at time  $t = 0.125$  and  $t = 0.25$ , respectively.

The correct form of the of the stabilization term is not trivial due to the time dependence in the poroelastic equations. The correct form has been derived using the previously presented numerical analysis. In Figure 5.5 we illustrate the convergence of the pressure error for three possible stabilization forms. As previously shown, the proposed stabilization  $J(p_{\delta t,h}, q_h)$  yields a stable solution and optimal convergence rate. The naive approach of inserting the stabilization  $J(p_h, q_h)$ , taken straight from Burman and Hansbo [2007], results in the solution becoming unstable after the first refinement step. This is because the stabilization becomes relatively small as  $\Delta t$  decreases. To overcome this issue one could chose to scale the stabilization, and try  $\frac{1}{\Delta t}J(p_h, q_h)$ . Although this stabilization now stays stable during refinement, it does not converge at an optimal rate. (We should change the stabilization to  $J(p_{\delta t,h}, q_h)$  the first time is introduced in the introduction)

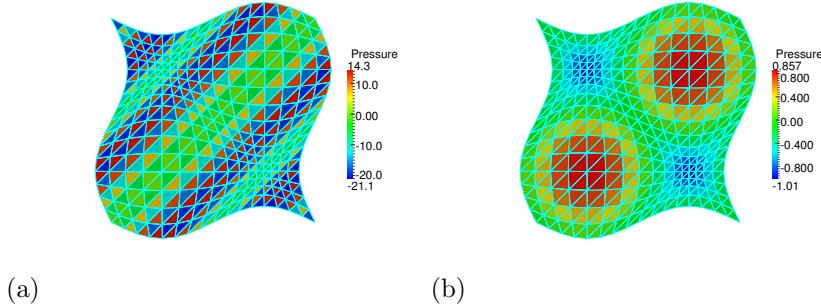


Figure 5.1: (a) Unstable pressure field, caused by not choosing the stabilization parameter  $\delta$  large enough, with  $\delta = 0.1$ , at  $t = 0.25$ . (b) Stable pressure field, with  $\delta = 1$  at  $t = 0.25$ .

### 5.4.3 3D convergence study

We extend the previous test problem to a unit cube. We now set the analytical pressure to

$p = \sin(2\pi x) \sin(2\pi y) \sin(2\pi z) \sin(2\pi t)$ , which gives the following displacement boundary condition, fluid flux boundary condition, and right hand side source term

$$\mathbf{u}_D = \begin{pmatrix} -\frac{1}{6}\pi \cos(2\pi x) \sin(2\pi y) \sin(2\pi z) \sin(2\pi t) \\ -\frac{1}{6}\pi \sin(2\pi x) \cos(2\pi y) \sin(2\pi z) \sin(2\pi t) \\ -\frac{1}{6}\pi \sin(2\pi x) \sin(2\pi y) \cos(2\pi z) \sin(2\pi t) \end{pmatrix},$$

$$q_d = \begin{pmatrix} -2\pi \cos(2\pi x) \sin(2\pi y) \sin(2\pi z) \sin(2\pi t) \\ -2\pi \sin(2\pi x) \cos(2\pi y) \sin(2\pi z) \sin(2\pi t) \\ -2\pi \sin(2\pi x) \sin(2\pi y) \cos(2\pi z) \sin(2\pi t) \end{pmatrix} \cdot \mathbf{n},$$

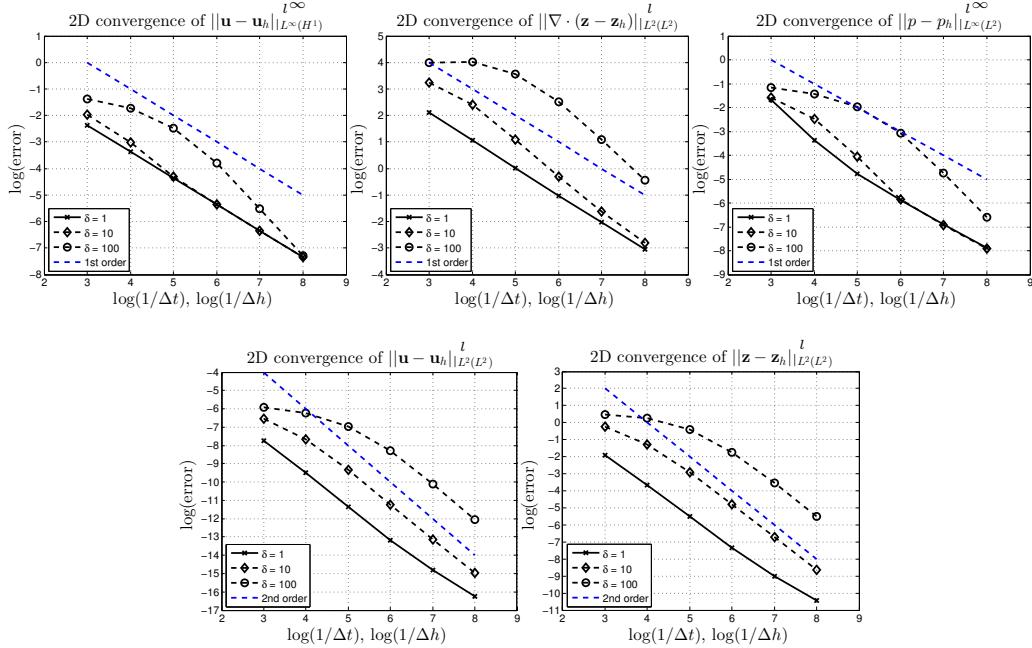


Figure 5.2: Convergence of the displacement, fluid flux, and pressure errors in their respective norms of the simplified poroelastic 2D test problem with different (stable) values for the stabilization parameter  $\delta$ .

$$g = 2\pi \sin(2\pi x) \sin(2\pi y) \sin(2\pi z) \cos(2\pi t) + 12\pi^2 \sin(2\pi x) \sin(2\pi y) \sin(2\pi z) \sin(2\pi t).$$

We also set the pressure to zero at  $[0, 0, 0]$  to remove the null space. In Figure 5.6, we show the 3D convergence of the method in relevant norms for each variable, with  $\delta = 0.001, 0.01, 0.1$ . The rates of convergence agree with the theoretically derived error estimates. We have observed that for 3D problems  $\delta$  can be chosen to be very small compared to 2D problems, making the effect of the stabilization term negligible to the system. This can be explained by the improved ratio of solid displacement and fluid flux nodes

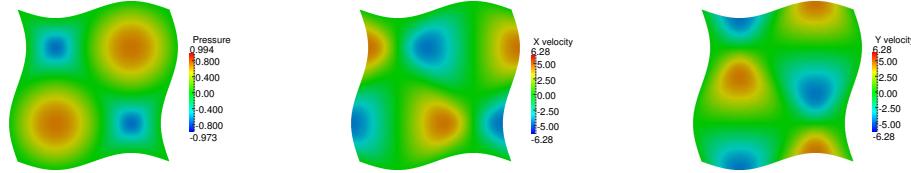


Figure 5.3: Pressure (left),  $x$ -fluid flux (middle), and  $y$ -fluid flux (right) at  $t = 0.125$ .

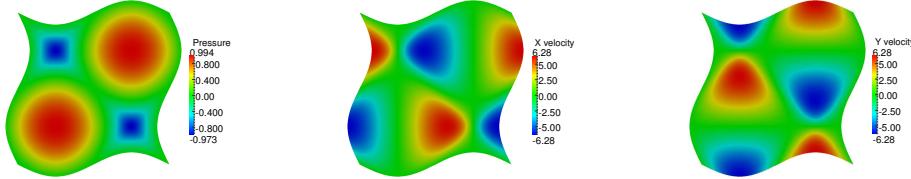


Figure 5.4: Pressure (left),  $x$ -fluid flux (right), and  $y$ -fluid flux (middle) at  $t = 0.25$ .

to pressure nodes in 3 dimensions, easing the LBB condition, the violation of which explains the spurious pressure oscillations found in unstable finite element formulations.

#### 5.4.4 2D cantilever bracket problem

It has been shown in Phillips and Wheeler [2009] that continuous Galerkin methods (CG/mixed) developed in Phillips and Wheeler [2007a,b] are susceptible to spurious pressure oscillations. The cause of this pressure instability has been attributed to a phenomenon called ‘locking’ by Phillips and Wheeler [2009], who give a discussion of locking in poroelasticity and show how it re-

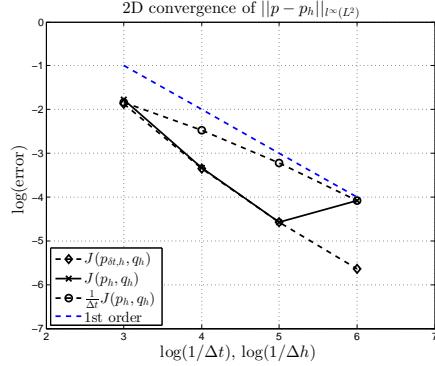


Figure 5.5: Convergence of the pressure error for three different stabilization forms, with  $\delta = 1$ .

lates to the locking phenomenon found in plane linear elasticity problems. A more recent paper by Haga et al. [2012] also investigates the cause of these pressure oscillations, they suggest that for low permeabilities the pressure oscillations are caused by a violation of the inf-sup (LBB) condition. Various methods which approximate the displacement using discontinuous and nonconforming elements have been proposed to overcome this problem [see, e.g., Li and Li, 2012, Liu, 2004, Phillips and Wheeler, 2008, Yi, 2013].

In this example, we consider the 2D cantilever bracket problem. The same

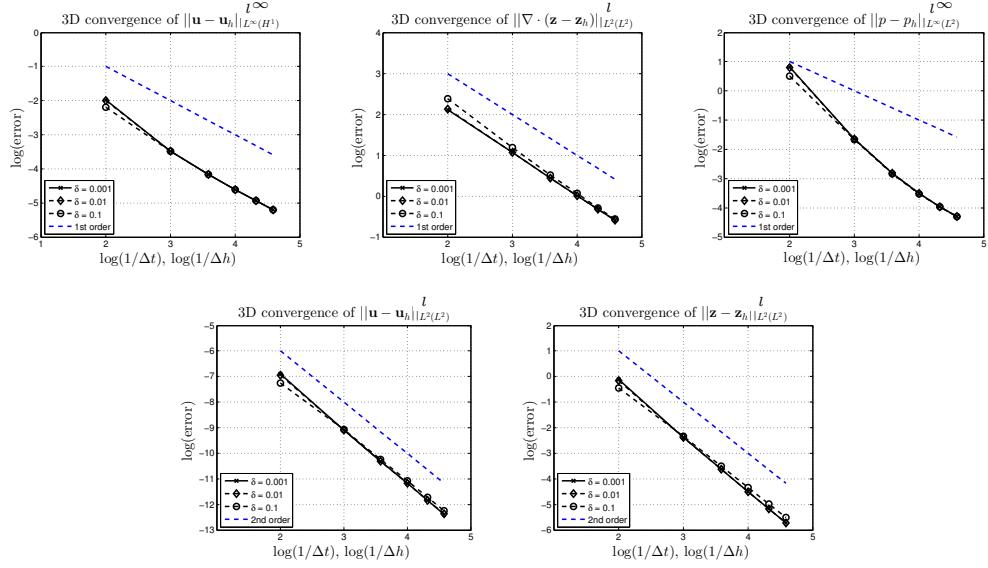


Figure 5.6: Convergence of the displacement, fluid flux, and pressure errors in their respective norms of the simplified poroelastic 3D test problem with different (stable) values for the stabilization parameter  $\delta$ .

test problem has also been used in Phillips and Wheeler [2009] to showcase the problem of spurious pressure oscillation, and used in Liu [2004] and Yi [2013] to demonstrate that their method is able to overcome these spurious pressure oscillations. The cantilever bracket problem (shown in Figure 5.7a) is solved on a unit square  $[0, 1]^2$ . No-flow flux boundary conditions are applied along all sides, the deformation is fixed ( $\mathbf{u} = 0$ ) along the left hand-side ( $x = 0$ ), and a downward traction force,  $\mathbf{t}_N \cdot \mathbf{n} = -1$ , is applied along the top edge ( $y = 1$ ). The right and bottom sides are traction-free. For this numerical experiment, we set  $\Delta t = 0.001$ ,  $h = 1/96$ ,  $\delta = 5e-6$ , and choose the same material parameters as in Phillips and Wheeler [2009] that typically

cause locking:  $E = 1e+5$ ,  $\nu = 0.4$ ,  $\alpha = 0.93$ ,  $c_0 = 0$ ,  $\kappa = 1e-7$ . The proposed stabilized finite element method yields a smooth pressure solution without any oscillations as is shown in Figure 5.7b.

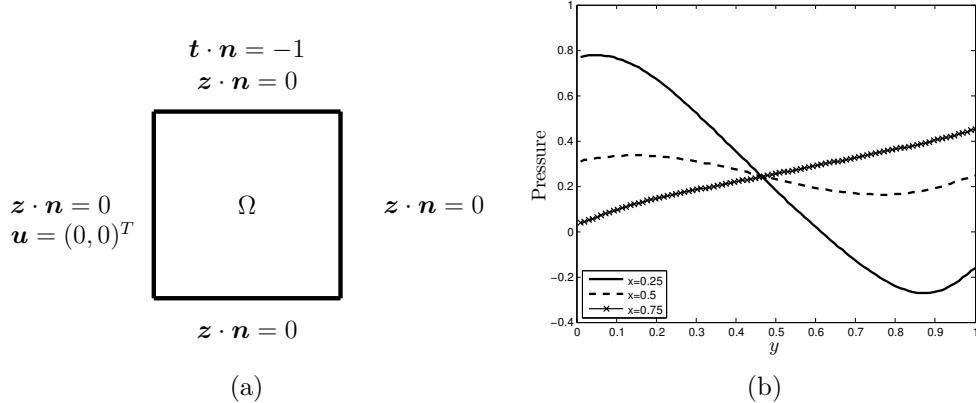


Figure 5.7: (a) Boundary conditions for the cantilever bracket problem. (b) Pressure solution of the cantilever bracket problem at  $t = 0.005$ .

#### 5.4.5 3D unconfined compression stress relaxation

In this test, a cylindrical specimen of porous tissue is exposed to a prescribed displacement in the axial direction while left free to expand radially, see Figure 6.1a. (Note that the two plates are not explicitly modelled in the simulation, but are realised through displacement boundary conditions.) After the initial loading, the displacement is held constant while the tissue relaxes in the radial direction due to interstitial fluid flow through the radial boundary. For the special case of a cylindrical geometry, Armstrong et al. [1984] found a closed-form analytical solution for the radial displacement  $u$

on the porous medium, given by

$$\frac{u}{a}(a, t) = \epsilon_0 \left[ \nu + (1 - 2\nu)(1 - \nu) \sum_{n=1}^{\infty} \frac{\exp(-\alpha_n^2 \frac{H_A k t}{a^2})}{\alpha_n^2 (1 - \nu)^2 - (1 - \nu)} \right]. \quad (5.65)$$

where  $\alpha_n$  are the solutions to the characteristic equation, given by  $J_1(x) - (1 - \nu)xJ_0(x)/(1 - 2\nu) = 0$ , where  $J_0$  and  $J_1$  are Bessel functions,  $\epsilon_0$  is the amplitude of the applied axial strain,  $a$  is the radius of the cylinder, and  $t_g$  is the characteristic time of diffusion (relaxation) given by  $t_g = a^2/Hk$ , where  $H = \lambda + 2\mu_s$  is the aggregate modulus of the elastic solid skeleton, and  $k$  is the permeability. The radial displacement predicted by our implementation (Figure 6.2) using a value of  $\delta = 0.001$  gives a root mean squared error of  $6.7 \times 10^{-4}$  against the analytical solution provided by Armstrong et al. [1984], and yields a stable solution. The same test problem has also been used to verify other poroelastic software such as FEBio [Maas et al., 2012]. The analytical solution available for this test problem describes the displacement of the outer radius which is directly dependent on the amount of mass in the system since the porous medium is assumed to be incompressible and fully saturated. It is therefore an ideal test problem for analyzing the effect that the added stabilization term has on the conservation of mass. In Figure 6.2 we can see that for large values of  $\delta$  the numerical solution loses mass faster and comes to a steady state that has less mass than the analytical solution. This is a clear limitation of the method and the stability parameter therefore

needs to be chosen carefully. However, for 3D problems  $\delta$  can be chosen to be very small so this effect is negligible, as can be seen in Figure 6.2 for a stable value of  $\delta = 0.001$ .

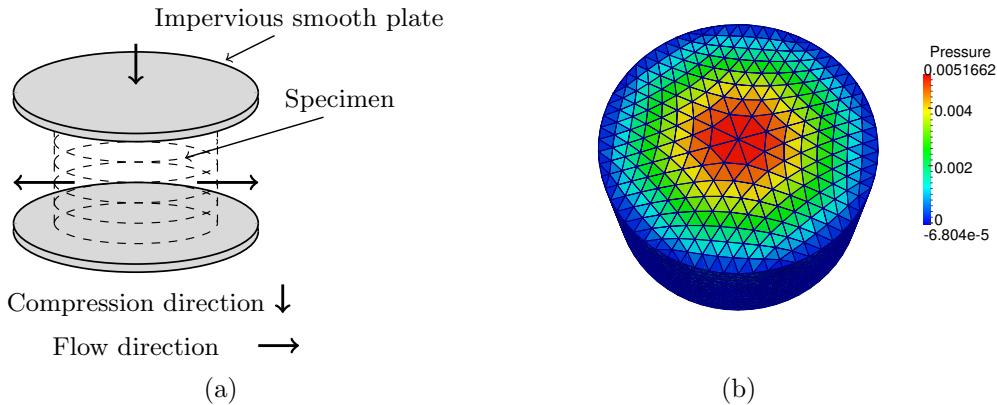


Figure 5.8: (a) Sketch of the test problem. The porous medium is being compressed between two smooth impervious plates. The frictionless plates permit the porous medium to expand in order to conserve volume and then to gradually relax as the fluid seeps out radially. (b) Pressure field solution at  $t = 5s$ , using a mesh with 28160 tetrahedra.

## 5.5 Conclusion

The local pressure jump stabilization method [Burman and Hansbo, 2007] is commonly used to solve the Stokes or Darcy equations using piecewise linear approximations for the velocities, and piecewise constant approximations for pressure variable. The main contribution of this paper has been to extend these ideas to three-field poroelasticity. We have presented a stability result

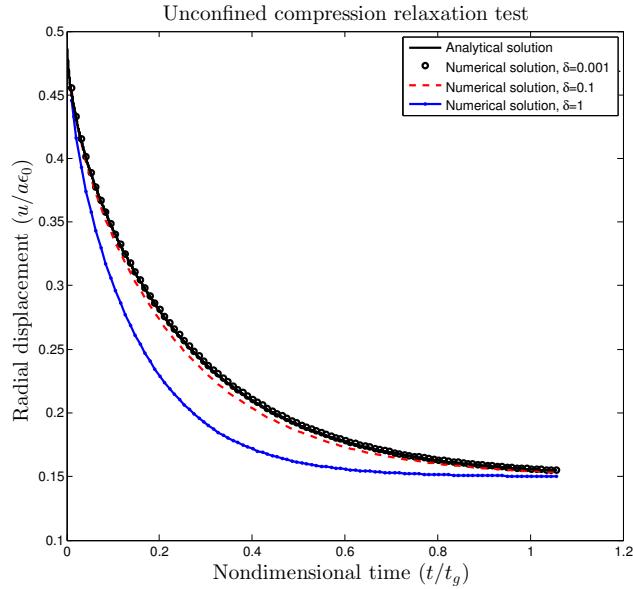


Figure 5.9: Normalized radial displacement versus normalized time calculated using the analytical solution, and using the proposed numerical method with different values of  $\delta$ . At  $t = 0$  the radial expansion is half of the axial compression indicating the instantaneous incompressibility of the poroelastic tissue. The final amount of tissue recoil depends on the intrinsic Poisson ratio of the tissue skeleton.

for the discretized equations that guarantees the existence of a unique solution at each time step, and derived an energy estimate which can be used to prove weak convergence of the solution to the discretized system to the solution to the continuous problem as the mesh parameters tend to zero. We also derived an optimal error estimate which includes an error for the fluid flux in its natural  $H\text{div}$  norm. We have also presented numerical experiments in 2D and 3D that illustrate the convergence of the method, the effectiveness of the method in overcoming spurious pressure oscillations, and the added

mass effect of the stabilization term.

# **Chapter 6**

## **A stabilized nonlinear finite element method for three-field incompressible poroelasticity valid in large deformations**

### **6.0.1 Proposed method**

In Berger et al. [2014a] a stabilized three-field finite element formulation of the linear fully saturated poroelasticity equations has been presented. This method uses the lowest possible approximation order: piecewise constant approximation for the pressure, and piecewise linear continuous elements for the displacements and fluid flux. Berger et al. [2014a] also proved stability

and error estimates for the fully discrete model. In this work we extend the method in [Berger et al., 2014a] to the large deformation (nonlinear) case. Some of the advantages of the proposed method are:

- i The proposed scheme is built on an existing scheme for linear poroelasticity, see Berger et al. [2014a], for which rigorous theoretical results about the stability and optimal convergence have been proven, and numerical experiments have confirmed its ability to overcome spurious pressure oscillations.
- ii Due to the discontinuous pressure approximation, sharp pressure gradients due to changes in material coefficients or boundary layer solutions can be captured reliably, circumventing the need for severe mesh refinement.
- iii Low-order finite element methods can be easily implemented and allow for efficient preconditioning [Ferronato et al., 2010, White and Borja, 2011]. Further, the convergence behavior of iterative solvers can be improved with the addition of a stabilization term Benzi et al. [2005].
- iv Other advantages from using a three-field formulation, as previously discussed, also exist.

The addition of the stabilization term introduces minimal additional computational work, can be assembled locally on each element using standard element information, and leads to a symmetric addition to the original system

matrix, thus preserving any existing symmetry. The effect of the stabilization on the conservation of mass is minimal in 3D, and disappears as the mesh is refined, see Berger et al. [2014a].

### 6.0.2 Overview of the paper

The rest of this paper is organized as follows: in section ??, we introduce the theory of poroelasticity, derive the general model of poroelasticity, and highlight any modelling assumptions required to arrive at the standard quasi-static incompressible model of poroelasticity; in section 6.1 we present the stabilized nonlinear finite element method. Finally in section 6.2, we present some numerical experiments in 3D to verify the method and illustrate its ability to reliably capture steep pressure gradients.

## 6.1 The stabilized nonlinear finite element method

We will now go on to solve the poroelastic model (7.5) using the finite element method. First, in section 6.1.1, we will present the continuous weak form of the problem, and then a linearization of the continuous weak form (section 6.1.2), to yield a Newton-Raphson algorithm (Newton's method) to the nonlinear problem. In section 6.1.3 we will discretise the Newton step and add a stabilization term to ensure that a discrete solution exists. Finally, in section 6.1.4, we will give further details on the implementation of the method.

### 6.1.1 Weak formulation

Instead of solving for the variable  $\chi(\mathbf{X}, t)$  we will, for notational purposes, solve for  $\mathbf{u}(\mathbf{X}, t) = \varphi(\mathbf{X}, t) - \mathbf{X}$ . Before presenting the weak formulation, we first need to introduce some notation and bilinear forms. We introduce the space  $L_0^2(\Omega) := \{q \in L^2(\Omega) : \int_{\Omega} q \, dx = 0\}$ , and define

$$\mathcal{L}(\Omega) := \begin{cases} L^2(\Omega) & \text{if } \Gamma_t \cup \Gamma_p \neq \emptyset \\ L_0^2(\Omega) & \text{if } \Gamma_t \cup \Gamma_p = \emptyset \end{cases}$$

We let  $\mathbf{W}^E = \{\mathbf{u} \in H^1([0, T]; \Omega) : \mathbf{u} = \mathbf{u}_D \text{ on } \Gamma_d\}$ ,  $\mathbf{W}^D = \{\mathbf{z} \in H_{div}([0, T]; \Omega) : \mathbf{z} \cdot \mathbf{n} = q_D \text{ on } \Gamma_f\}$ . We also define the mixed solution space  $\mathcal{W}^X = \{\mathbf{W}^E \times \mathbf{W}^D \times \mathcal{L}(\Omega)\}$ . For the test functions we define the spaces  $\mathbf{W}_0^E = \{\mathbf{v} \in H^1([0, T]; \Omega) : \mathbf{v} = 0 \text{ on } \Gamma_d\}$ ,  $\mathbf{W}_0^D = \{\mathbf{w} \in H_{div}([0, T]; \Omega) : \mathbf{w} \cdot \mathbf{n} = 0 \text{ on } \Gamma_f\}$ , and  $\mathcal{V}^X = \{\mathbf{W}_0^E \times \mathbf{W}_0^D \times \mathcal{L}(\Omega)\}$ . We now multiply the strong form of the problem (7.5) by test functions  $(\mathbf{v}, \mathbf{w}, q) \in \mathcal{V}^X$ , integrate and use the divergence theorem, make use of the identity<sup>1</sup>  $\nabla \cdot (\cdot \boldsymbol{\sigma}_e \mathbf{v}) = \nabla \cdot \boldsymbol{\sigma}_e \cdot \mathbf{v} - \boldsymbol{\sigma}_e : \nabla \mathbf{v}$ , and take into account the symmetry of  $\boldsymbol{\sigma}_e$  to yield the continuous weak problem, which is to find  $\mathbf{u}(\mathbf{X}, t) \in \mathbf{W}^E$ ,  $\mathbf{z}(\mathbf{x}, t) \in \mathbf{W}^D$ , and  $p(\mathbf{x}, t) \in \mathcal{L}(\Omega)$  for any

---

<sup>1</sup>The double contraction of two tensors  $\mathbf{A}$  and  $\mathbf{B}$  is defined in terms of the trace as  $\mathbf{A} : \mathbf{B} = \text{tr}(\mathbf{A}^T \mathbf{B})$ .

time  $t \in [0, T]$  such that:

$$\int_{\Omega_t} \boldsymbol{\sigma}_e : \nabla^S \mathbf{v} \, dv - \int_{\Omega_t} p \nabla \cdot \mathbf{v} \, dv = \int_{\Omega_t} \rho \mathbf{f} \cdot \mathbf{v} \, dv + \int_{\Gamma_t} \mathbf{t}_N \cdot \mathbf{v} \, ds \quad \forall \mathbf{v} \in \mathbf{W}_0^E, \quad (6.1a)$$

$$\int_{\Omega_t} \kappa^{-1} \mathbf{z} \cdot \mathbf{w} \, dv - \int_{\Omega_t} p \nabla \cdot \mathbf{w} \, dv = \int_{\Omega_t} \rho^f \mathbf{f} \cdot \mathbf{w} \, dv - \int_{\Gamma_p} p_D \mathbf{w} \cdot \mathbf{n} \, ds \quad \forall \mathbf{w} \in \mathbf{W}_0^D, \quad (6.1b)$$

$$\int_{\Omega_t} q \nabla \cdot \mathbf{u}_t \, dv + \int_{\Omega_t} q \nabla \cdot \mathbf{z} \, dv = \int_{\Omega_t} g q \, dv \quad \forall q \in \mathcal{L}(\Omega). \quad (6.1c)$$

Here  $\nabla^S \mathbf{v} = \frac{1}{2} (\nabla \mathbf{v} + (\nabla \mathbf{v})^T)$  is the symmetric version of the spatial gradient  $\nabla \mathbf{v}$ .

### 6.1.2 Newton's Method

Since the system of equations (6.1) is highly nonlinear, its solution requires a scheme such as Newton's method. With Newton's method, an improved solution is obtained from a linear approximation of the nonlinear equation at an already computed solution. This first order Taylor expansion corresponds in finite element applications to the linearization of the weak form, and can be obtained by the directional derivative, explained in section 6.1.2. Let  $\mathbf{u} = \{\mathbf{u}, \mathbf{z}, p\}$  denote the solution vector,  $\xi \mathbf{u} = \{\xi \mathbf{u}, \xi \mathbf{z}, \xi p\}$  denote the solution increment vector, and  $\mathbf{v} = \{\mathbf{v}, \mathbf{w}, q\}$  the corresponding vector of test functions. Then the nonlinear system of equations (6.14) can be recast in

the form

$$G(\mathbf{u}, \mathbf{v}) = 0, \quad (6.2)$$

where

$$\begin{aligned} G(\mathbf{u}, \mathbf{v}) &= \int_{\Omega_t} \boldsymbol{\sigma}_e : \nabla^S \mathbf{v} - p \nabla \cdot \mathbf{v} + \kappa^{-1} \mathbf{z} \cdot \mathbf{w} - p \nabla \cdot \mathbf{w} + q \nabla \cdot (\mathbf{u}_{\delta t} + \cdot \mathbf{z}) \, dv \\ &\quad - \int_{\Omega_t} \rho \mathbf{f} \cdot \mathbf{v} + \rho^f \mathbf{f} \cdot \mathbf{w} + gq \, dv - \int_{\Gamma_t} \mathbf{t}_N \cdot \mathbf{v} \, ds + \int_{\Gamma_p} p_D \mathbf{w} \cdot \mathbf{n} \, ds \end{aligned} \quad (6.3)$$

Considering a trial solution  $\bar{\mathbf{u}}$ , equation (6.2) can now be linearized in the direction of an increment  $\xi \mathbf{u}$  at  $\bar{\mathbf{u}}$  as

$$G(\bar{\mathbf{u}}, \mathbf{v}) + DG(\bar{\mathbf{u}}, \mathbf{v})[\xi \mathbf{u}] = 0, \quad (6.4)$$

or

$$DG(\bar{\mathbf{u}}, \mathbf{v})[\xi \mathbf{u}] = -G(\bar{\mathbf{u}}, \mathbf{v}), \quad (6.5)$$

which essentially is Newton's method (see algorithm 2 for the fully discrete version).

## Linearization

In biphasic tissue problems, it is common to approximate the tangent by taking the nonlinear elasticity term as the only nonlinearity present and ignoring the other nonlinearities [Ün and Spilker, 2006, White and Borja,

2008]. The dominant nonlinearity in (6.1) is the elasticity term denoted by

$$E((\mathbf{u}, p), \mathbf{v}) = \int_{\Omega_t} \boldsymbol{\sigma}_e : \nabla^S \mathbf{v} - p \nabla \cdot \mathbf{v} \, dv. \quad (6.6)$$

For Newton's method we require the directional derivative of  $E(\mathbf{u}, \mathbf{v})$  at a particular trial solution  $\bar{\mathbf{u}}$  in the direction  $\xi \mathbf{u}$ , given by (see [Wriggers, 2008, section 3.5.3])

$$DE((\bar{\mathbf{u}}, \bar{p}), \mathbf{v})[\xi \mathbf{u}] = \int_{\Omega_t} \nabla^S \mathbf{v} : \bar{\mathbf{c}} : \nabla^S \xi \mathbf{u} + \bar{\boldsymbol{\sigma}}_e : ((\nabla \xi \mathbf{u})^T \cdot \nabla \mathbf{v}) \, dv, \quad (6.7)$$

where  $\bar{\mathbf{c}}$  is the fourth-order spatial tangent modulus tensor  $\bar{\mathbf{c}}$ , and  $\bar{\boldsymbol{\sigma}}_e$  is the effective (elastic) stress tensor both evaluated at a trial solution  $\bar{\mathbf{u}}$ . Any variable with a bar above it will correspond to it being evaluated at a trial solution and will therefore be considered as a known quantity. In the fully discrete algorithm 2, this trial solution will correspond to the solution of the previous Newton step. The spatial tangent modulus tensor  $\bar{\mathbf{c}}$ , due to its complexity, is described in section A.1. For a detailed explanation and derivation see Bonet and Wood [1997], Wriggers [2008]. The approximate linearization of the nonlinear problem (6.2) is thus given by

$$\begin{aligned} DG(\bar{\mathbf{u}}, \mathbf{v})[\xi \mathbf{u}] &\approx \int_{\Omega_t} \nabla^S \mathbf{v} : \bar{\mathbf{c}} : \nabla^S \xi \mathbf{u} + \bar{\boldsymbol{\sigma}}_e : ((\nabla \xi \mathbf{u})^T \cdot \nabla \mathbf{v}) - \xi p \nabla \cdot \mathbf{v} \, dv \\ &+ \int_{\Omega_t} \bar{\boldsymbol{\kappa}}^{-1} \xi \mathbf{z} \cdot \mathbf{w} - \xi p \nabla \cdot \mathbf{w} \, dv + \int_{\Omega_t} q \nabla \cdot (\frac{\xi \mathbf{u}}{\Delta t} + \cdot \xi \mathbf{z}) \, dv. \end{aligned} \quad (6.8)$$

For a more complete linearization see Ateshian et al. [2010] and Maas et al. [2010], this might be needed when a particularly nonlinear permeability law or large time steps are being used.

### 6.1.3 Discretization of the Newton step

We begin with some standard finite element notation. Let  $\xi_h = \{E_1, E_2, \dots, E_{N_h}\}$  be a subdivision of  $\Omega$ , where  $E_j$  is a  $d$ -simplex. Let  $h_j = \text{diam}(E_j)$  and set  $h = \max\{h_j : j = 1, \dots, M_h\}$ . The mixed finite element space defined on  $\xi_h$  is given as  $\mathcal{W}_h^X = (\mathbf{W}_h^E \times \mathbf{W}_h^D \times Q_h) \subset (\mathbf{W}^E \times \mathbf{W}^D \times \mathcal{L}(\Omega))$ , similarly for the test functions we let  $\mathcal{V}_h^X = (\mathbf{W}_{h0}^E \times \mathbf{W}_{h0}^D \times Q_h) \subset (\mathbf{W}_0^E \times \mathbf{W}_0^D \times \mathcal{L}(\Omega))$ .

That is,

$$\begin{aligned}\mathbf{W}_h^E &= \{\mathbf{u}_h \in C^0(\Omega) : \mathbf{u}_h|_K \in P_1(K); \mathbf{u}_h = \mathbf{u}_D \text{ on } \Gamma_d; \forall K \in \mathcal{T}^h\}, \\ \mathbf{W}_h^D &= \{\mathbf{z}_h \in C^0(\Omega) : \mathbf{z}_h|_K \in P_1(K); \mathbf{z}_h \cdot \mathbf{n} = q_D \text{ on } \Gamma_f; \forall K \in \mathcal{T}^h\}, \\ Q_h &= \begin{cases} \{p_h : p_h|_K \in P_0(K); \forall K \in \mathcal{T}^h\} & \text{if } \Gamma_t \cup \Gamma_p \neq \emptyset \\ \{p_h : p_h|_K \in P_0(K); \int_{\Omega} p_h = 0; \forall K \in \mathcal{T}^h\} & \text{if } \Gamma_t \cup \Gamma_p = \emptyset \end{cases}, \\ \mathbf{W}_{h0}^E &= \{\mathbf{v}_h \in C^0(\Omega) : \mathbf{v}_h|_K \in P_1(K); \mathbf{v}_h = 0 \text{ on } \Gamma_d; \forall K \in \mathcal{T}^h\}, \\ \mathbf{W}_{h0}^D &= \{\mathbf{w}_h \in C^0(\Omega) : \mathbf{w}_h|_K \in P_1(K); \mathbf{w}_h \cdot \mathbf{n} = 0 \text{ on } \Gamma_f; \forall K \in \mathcal{T}^h\},\end{aligned}$$

where  $P_0(K)$  and  $P_1(K)$  are the spaces of constant and linear polynomials on  $K$ , respectively. We will now give a Galerkin conforming mixed finite element method for discretization of (6.1). To discretize the time derivative we use

the fully implicit backward Euler scheme, which we will denote using the shorthand  $v_{\delta t}^n := \frac{v^n - v^{n-1}}{\Delta t}$ . The fully discretized weak problem at each Newton step, to get an update for the approximate solution, is to find  $\xi \mathbf{u}_h \in \mathbf{W}_h^E$ ,  $\xi \mathbf{z}_h \in \mathbf{W}_h^D$  and  $\xi p_h \in Q_h$  such that:

$$\begin{aligned} & \int_{\Omega_t} \nabla^S \mathbf{v}_h : \bar{\boldsymbol{\sigma}} : \nabla^S \xi \mathbf{u}_h + \bar{\boldsymbol{\sigma}}_e : ((\nabla \xi \mathbf{u}_h)^T \cdot \nabla \mathbf{v}_h) - \xi p_h \nabla \cdot \mathbf{v}_h \, dv \\ &= \int_{\Omega_t} \bar{\boldsymbol{\sigma}}_e : \nabla^S \mathbf{v}_h - \bar{p}_h \nabla \cdot \mathbf{v}_h - \bar{\rho} \mathbf{f} \cdot \mathbf{v}_h \, dv - \int_{\Gamma_t} \mathbf{t}_N \cdot \mathbf{v}_h \, ds \quad \forall \mathbf{v}_h \in \mathbf{W}_{h0}^E, \end{aligned} \quad (6.9a)$$

$$\begin{aligned} & \int_{\Omega_t} \bar{\boldsymbol{\kappa}}^{-1} \xi \mathbf{z}_h \cdot \mathbf{w}_h \, dv - \xi p \nabla \cdot \mathbf{w}_h \, dv \\ &= \int_{\Omega_t} \bar{\boldsymbol{\kappa}}^{-1} \bar{\mathbf{z}}_h \cdot \mathbf{w}_h \, dv - \int_{\Omega_t} \bar{p}_h \cdot \nabla \mathbf{w}_h \, dv - \int_{\Omega_t} \bar{\rho}^f \mathbf{f} \cdot \mathbf{w} \, dv + \int_{\Gamma_p} p_D \mathbf{w}_h \cdot \mathbf{n} \quad \forall \mathbf{w}_h \in \mathbf{W}_{h0}^D, \end{aligned} \quad (6.9b)$$

$$\int_{\Omega_t} q_h \nabla \cdot \left( \frac{\xi \mathbf{u}_h}{\Delta t} + \cdot \xi \mathbf{z}_h \right) \, dv = \int_{\Omega_t} q_h \nabla \cdot (\mathbf{u}_{\delta t, h}^- + \cdot \bar{\mathbf{z}}_h) - g q_h \, dv \quad \forall q_h \in Q_h. \quad (6.9c)$$

By adding (6.9a, 6.9b and 6.9c), we can rewrite this using more compact notation

$$DG_h(\bar{\mathbf{u}}_h, \mathbf{v}_h)[\xi \mathbf{u}_h] = -G(\bar{\mathbf{u}}_h, \mathbf{v}_h), \quad (6.10)$$

Note that this system does not satisfy the LBB condition and would therefore yield spurious pressure oscillations. To remove the spurious pressure oscillations and gain coercivity (satisfy the LBB condition) we introduce a

stabilization term, such that

$$DG_h(\bar{\mathbf{u}}_h, \mathbf{v}_h)[\xi \mathbf{u}_h] + J(p_{\delta t,h}, q_h)[\xi \mathbf{u}_h] = -G(\bar{\mathbf{u}}_h, \mathbf{v}_h) - J(p_{\delta t,h}, q_h). \quad (6.11)$$

The stabilization term is given by,

$$J(p_{\delta t,h}, q_h) = \delta \sum_K \int_{\partial K \setminus \partial \Omega} h_{\partial K} [p_{\delta t,h}] [q_h] \, ds.$$

Here  $\delta$  is a penalty parameter that is independent of  $h$  and  $\Delta t$ . It has been shown in the numerical results section of [Berger et al., 2014a], that the convergence is not sensitive to  $\delta$ . The set of all elements is denoted by  $K$ ,  $h_{\partial K}$  denotes the size of an element edge in 2D or face in 3D, and  $[.]$  is the jump across an edge. As an example consider  $[p_h]$ , the jump operator on the piecewise constant pressure. The jump in pressure  $[p_h]$  across an element or face  $E$  adjoining elements  $T$  and  $S$  is defined such that

$$(p_h|_T - p_h|_S)\mathbf{n}_{E,T} = (p_h|_S - p_h|_T)\mathbf{n}_{E,S}.$$

Here  $\mathbf{n}_{E,T}$  is the outward normal from element  $T$ , with respect to edge  $E$ ,  $\mathbf{n}_{E,S}$  is the corresponding inward facing normal, and  $p_h|_T$  and  $p_h|_S$  denote the pressure in element  $T$  and  $S$ , respectively. Finally we can now rewrite the problem of finding the Newton update as

$$\mathcal{K}(\bar{\mathbf{u}}_h, \mathbf{v}_h)[\xi \mathbf{u}_h] = -\mathcal{R}(\bar{\mathbf{u}}_h, \mathbf{v}_h), \quad (6.12)$$

where  $\mathcal{K}(\bar{\mathbf{u}}_h, \mathbf{v}_h)[\xi \mathbf{u}_h] = DG_h(\bar{\mathbf{u}}_h, \mathbf{v}_h)[\xi \mathbf{u}_h] + J(p_{\delta t, h}, q_h)[\xi \mathbf{u}_h]$ , and  $\mathcal{R}(\bar{\mathbf{u}}_h, \mathbf{v}_h) = G(\bar{\mathbf{u}}_h, \mathbf{v}_h) + J(p_{\delta t, h}, q_h)$ .

#### 6.1.4 Newton algorithm

We will now let  $\mathbf{u}_i^n := \{\mathbf{u}_i^n, \mathbf{z}_i^n, p_i^n\}$  denote the fully discrete solution at the  $i$ th step within the Newton method at time  $t^n$ . To solve the nonlinear poroelastic problem using Newton's method at a particular time step, we perform the following steps:

---

**Algorithm 1** Fully discrete Newton's method (for a particular time step  $t^n$ )

---

Initialize  $\mathbf{u}_0^n = \{\mathbf{u}^{n-1}, \mathbf{z}^{n-1}, p^{n-1}\}$

Iteration loop  $i = 0, 1, \dots$  until convergence.

1. Assemble  $\mathbf{R}(\mathbf{u}_i^n, \mathbf{u}^{n-1})$  and  $\mathbf{K}(\mathbf{u}_i^n)$ .
  2. Compute the solution increment from the linear system:  

$$\mathbf{K}(\mathbf{u}_i^n)\xi \mathbf{u}_{i+1}^n = -\mathbf{R}(\mathbf{u}_i^n, \mathbf{u}^{n-1}).$$
  3. Compute the new solution:  $\mathbf{u}_{i+1}^n = \mathbf{u}_i^n + \xi \mathbf{u}_{i+1}^n$ .
  4. <sup>2</sup>Update the mesh,  $\Omega_t = \mathbf{X} + \mathbf{u}_i^n$ .
  5. Test for convergence:  $\|\mathbf{R}(\mathbf{u}_i^n, \mathbf{u}^{n-1})\| > \text{TOL} \rightarrow$  Set  $i = i + 1$  go to 1,  

$$\|\mathbf{R}(\mathbf{u}_i^n, \mathbf{u}^{n-1})\| \leq \text{TOL} \rightarrow$$
  
Stop and move to the next time step.
- 

<sup>2</sup>This method is known in the literature as an updated Lagrange formulation, see e.g. [Bathe et al., 1975] or [Wriggers, 2008, chapter 3], however it should really be called an updated Eulerian formulation, since the current configuration is being updated during each Newton iteration. This method is also used by [Ateshian et al., 2010, White and Borja, 2008] to solve the large deformation poroelasticity problem.

We need to solve the following linear system at each Newton step:

$$\mathbf{K}(\mathbf{u}_i^n)\xi\mathbf{u}_{i+1}^n = -\mathbf{R}(\mathbf{u}_i^n, \mathbf{u}^{n-1}), \quad (6.13)$$

which can be expanded, and written as

$$\begin{bmatrix} \mathbf{K}^e & 0 & \mathbf{B}^T \\ 0 & \mathbf{M} & \mathbf{B}^T \\ -\mathbf{B} & -\Delta t \mathbf{B} & \mathbf{J} \end{bmatrix} \begin{bmatrix} \xi \mathbf{u}_{i+1}^n \\ \xi z_{i+1}^n \\ \xi \mathbf{p}_{i+1}^n \end{bmatrix} = - \begin{bmatrix} \mathbf{r}_1(\mathbf{u}_i^n, p_i^n) \\ \mathbf{r}_2(\mathbf{u}_i^n, \mathbf{z}_i^n, p_i^n) \\ \mathbf{r}_3(\mathbf{u}_i^n, \mathbf{u}^{n-1}, \mathbf{z}_i^n, p_i^n) \end{bmatrix},$$

where we have defined the following matrices:

$$\begin{aligned} \mathbf{K}^e &= [\mathbf{a}_{kl}], \quad \mathbf{k}_{kl}^e = \int_{\Omega_t} \mathbf{B}_k^T \mathbf{D}(\mathbf{u}_i^n) \mathbf{B}_l + (\nabla \boldsymbol{\phi}_k)^T \boldsymbol{\sigma}_e(\mathbf{u}_i^n) \nabla \boldsymbol{\phi}_l \, dv, \\ \mathbf{M} &= [\mathbf{m}_{kl}], \quad \mathbf{m}_{kl} = \int_{\Omega_t} \kappa^{-1}(\mathbf{u}_i^n) \boldsymbol{\phi}_k \cdot \boldsymbol{\phi}_l \, dv, \\ \mathbf{B} &= [\mathbf{b}_{kl}], \quad \mathbf{b}_{kl} = - \int_{\Omega_t} \psi_k \nabla \cdot \boldsymbol{\phi}_l \, dv, \\ \mathbf{J} &= [\mathbf{j}_{kl}], \quad \mathbf{j}_{kl} = \delta \sum_K \int_{\partial k \setminus \partial \Omega_t} h_{\partial K}[\psi_k] [\psi_k] \, ds. \\ \mathbf{r}_1 &= [\mathbf{r}_{1i}], \quad \mathbf{r}_{1i} = \int_{\Omega_t} (\boldsymbol{\sigma}_e(\mathbf{u}_i^n) - p_i^n \mathbf{I}) : \nabla \boldsymbol{\phi}_i - \rho(\mathbf{u}_i^n) \boldsymbol{\phi}_i \cdot \mathbf{f} \, dv - \int_{\Gamma_t} \boldsymbol{\phi}_i \cdot \mathbf{t}_N \, ds, \\ \mathbf{r}_2 &= [\mathbf{r}_{2i}], \quad \mathbf{r}_{2i} = \int_{\Omega_t} \kappa^{-1}(\mathbf{u}_i^n) \boldsymbol{\phi}_i \cdot \mathbf{z}_i^n - p_i^n \nabla \cdot \boldsymbol{\phi}_i - \rho^f(\mathbf{u}_i^n) \boldsymbol{\phi}_i \cdot \mathbf{f} \, dv, \\ \mathbf{r}_3 &= [\mathbf{r}_{3i}], \quad \mathbf{r}_{3i} = \int_{\Omega_t} \psi_i \nabla \cdot (\mathbf{u}_i^n - \mathbf{u}^{n-1}) + \Delta t \psi_i \nabla \cdot \mathbf{z}_i^n - \Delta t \psi_i g \, dv + \delta \sum_K \int_{\partial k \setminus \partial \Omega_t} h_{\partial K}[\psi_i] [p_i^n - p^{n-1}] \, ds \end{aligned}$$

Here  $\phi_k$  are vector valued linear basis functions such that the displacement vector at the  $i$ th iteration can be written as  $\mathbf{u}_i^n = \sum_{k=1}^{n_u} \mathbf{u}_{i,k}^n \phi_k$ , with  $\sum_{k=1}^{n_u} \mathbf{u}_{i,k}^n \phi_k \in \mathbf{W}_h^E$ . Similarly for the fluid flux vector we have  $\mathbf{z}_i^n = \sum_{k=1}^{n_z} \mathbf{z}_{i,k}^n \phi_k$ , with  $\sum_{k=1}^{n_z} \mathbf{z}_{i,k}^n \phi_k \in \mathbf{W}_h^D$ . The scalar valued constant basis functions  $\psi_i$  are used to approximate the pressure, such that  $\mathbf{p}_i^n = \sum_{k=1}^{n_p} p_{i,k}^n \psi_k$ , with  $\sum_{k=1}^{n_p} p_{i,k}^n \psi_k \in Q_h$ . Also to aid the assembly of the fourth order tensor we have adopted the matrix voigt notation. In particular  $\mathbf{D}$  is the matrix form of  $\mathbf{c}$ , and  $\mathbf{B}_k$  is the matrix version of  $\nabla^S \phi_k$ , see (A.3) and (A.4) for details.

### Algorithm to assemble the stabilization matrix

We will give now a procedure for building the stabilization matrix  $\mathbf{J}$ . Let  $\mathcal{T}_h$  be a partition of  $\Omega$  into triangles (quadilaterals) in 2D or tetrahedrals (hexahedra) in 3D, and  $\mathcal{K} \in \mathcal{T}_h$  be an element in this partition. We denote by  $\mathcal{N}_h$  the set of all interpolation nodes related to  $\mathcal{T}_h$ , and by  $\mathcal{N}_h^p(\mathcal{K})$  the set of all pressure nodes that belong to an element  $\mathcal{K}$ . Note that this should only be one node per element because we are using a piecewise constant approximation for the pressure. Also let  $\mathcal{A}_h^{el}(\mathcal{K})$  be the set of neighboring elements  $\mathcal{K}^N \in \mathcal{T}_h$  to  $\mathcal{K}$  and let  $\mathcal{D}(\mathcal{N}_h^p(\mathcal{K}))$  be the degree of freedom associated with node  $\mathcal{N}_h^p(\mathcal{K})$ .

*Algorithm to assemble stabilization matrix  $\mathbf{J}$ :*

**for** every  $\mathcal{K} \in \mathcal{T}_h$  **do**

```

for every  $\mathcal{K}^N \in \mathcal{A}_h^{el}(\mathcal{K})$  do
    compute  $h_{\partial K}$ 
     $i \leftarrow \mathcal{D}(\mathcal{N}_h^p(\mathcal{K}))$ 
     $j \leftarrow \mathcal{D}(\mathcal{N}_h^p(\mathcal{K}^N))$ 
     $\mathbf{J}_{ii} \leftarrow \mathbf{J}_{ii} + (\delta h_{\partial K} \text{ in 2D}, \delta h_{\partial K}^{3/2} \text{ in 3D})$ 
     $\mathbf{J}_{kl} \leftarrow \mathbf{J}_{kl} - (\delta h_{\partial K} \text{ in 2D}, \delta h_{\partial K}^{3/2} \text{ in 3D})$ 
end for
end for

```

### No-flux boundary condition

We introduce a Lagrange multiplier,  $\lambda$ , to enforce the no-flux boundary condition  $\mathbf{z} \cdot \mathbf{n} = 0$  along the boundary  $\Gamma_f$ . Let  $W^f = \{\lambda \in H_{div}(\Gamma_f, \mathbb{R})\}$ . The resulting modified continuous weak-form is now:

$$G((\mathbf{u}, \mathbf{z}, p), (\mathbf{v}, \mathbf{w}, q)) + (\lambda, \mathbf{w} \cdot \mathbf{n})_{\Gamma_f} = 0 \quad \forall (\mathbf{v}, \mathbf{w}, q) \in \mathbf{W}_0^E, \mathbf{W}_0^D, \mathcal{L}(\Omega), \quad (6.14a)$$

$$(\mathbf{z} \cdot \mathbf{n}, \mathbf{l})_{\Gamma_f} = 0, \quad \forall \mathbf{l} \in W^f. \quad (6.14b)$$

The discretization and implementation of this additional constraint is straightforward and results in a linear system with additional degrees of freedom for every node on  $\Gamma_f$ . Note that the added terms  $(\lambda, \mathbf{w} \cdot \mathbf{n})_{\Gamma_f}$  and  $(\mathbf{z} \cdot \mathbf{n}, \mathbf{l})_{\Gamma_f}$  are actually nonlinear since the normal is a function of the displacement.

However we have found that by treating this term as being linear during the linearization process does not affect the convergence of the Newton algorithm. Alternatively this term could be linearized as has been described in detail for the traction boundary condition, see Wriggers [2008, section 4.2.5].

## 6.2 Numerical results

We now present three numerical examples to test the performance of the proposed stabilized finite element method. The first two examples are from mechanobiology and geotechnical applications, while the last is a swelling example that undergoes significant large deformations. For the implementation we used the C++ library libmesh [Kirk et al., 2006], and the multi-frontal direct solver mumps [Amestoy et al., 2000] to solve the resulting linear systems. For the strain energy law we chose a Neo-Hookean law taken from [Wriggers, 2008, eqn. (3.119)],

$$W(\mathbf{C}) = \frac{\mu}{2}(\text{tr}(\mathbf{C}) - 3) + \frac{\Lambda}{4}(J^2 - 1) - (\mu + \frac{\Lambda}{2})\ln J. \quad (6.15)$$

The material parameters  $\mu$  and  $\Lambda$  can be related to the more familiar Young's modulus  $E$  and the Poisson ratio  $\nu$  by  $\mu = \frac{E}{2(1+\nu)}$  and  $\Lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}$ . Details on the effective stress tensor and fourth-order spatial tangent modulus for this particular law can be found in section A.2.1. For the permeability law

we chose

$$\kappa_0(\chi) = \kappa_0 \mathbf{I}. \quad (6.16)$$

### 6.2.1 3D unconfined compression stress relaxation

In this test, a cylindrical specimen of porous tissue is exposed to a prescribed displacement in the axial direction while left free to expand radially. The original experiment involved a specimen of articular cartilage being compressed via impervious smooth plates as shown in Figure 6.1a, note that the two plates are not explicitly modelled in the simulation, but are realised through displacement boundary conditions. After loading the tissue, the displacement is held constant while the tissue under the displacement relaxes in the radial direction due to interstitial fluid flow through the material and the frictionless plates. For the porous tissue, the outer radial boundary is permeable and free-draining, the upper and lower fluid boundaries are impermeable and have a no flux condition imposed. The outer radius and height of the cylinder is  $5mm$ , whereas the axial compression is  $0.01mm$ . The bottom of the tissue is constrained in the vertical direction. The fluid pressure was constrained to zero at the outer radial surface. The parameters used for the simulation can be found in Table 6.1. For the special case of a cylindrical geometry and assumptions regarding the direction of the fluid flow, Armstrong et al. [1984] found a closed-form analytical solution for the radial displacement on the porous medium in response to a step loading function. The analytical solution for the radial displacement to this unconfined compression test is

Parameter	Description	Value
$\kappa_0$	Dynamic permeability	$10^{-3} \text{ m}^3 \text{s kg}^{-1}$
$\nu$	Poisson ratio	0.15
$E$	Young's modulus	$1000 \text{ kg m}^{-1} \text{s}^{-2}$
$\Delta t$	Time step used in the simulation	4 s
$T$	Final time of the simulation	1000 s
$\delta$	Stabilization parameter	$10^{-3}$

Table 6.1: Parameters used for the unconfined compression test problem.

given by

$$\frac{u}{a}(a, t) = \epsilon_0 \left[ \nu + (1 - 2\nu)(1 - \nu) \sum_{n=1}^{\infty} \frac{\exp(-\alpha_n^2 \frac{H_A k t}{a^2})}{\alpha_n^2 (1 - \nu)^2 - (1 - \nu)} \right]. \quad (6.17)$$

Here  $\alpha_n$  are the solutions to the characteristic equation, given by  $J_1(x) - (1 - \nu)xJ_0(x)/(1 - 2\nu) = 0$ , where  $J_0$  and  $J_1$  are Bessel functions. We also have that  $\epsilon_0$  is the amplitude of the applied axial strain,  $a$  is the radius of the cylinder, and  $t_g$  is the characteristic time of diffusion given by  $t_g = a^2/Hk$ , where  $H = \lambda + 2\mu_s$  is the aggregate modulus of the elastic solid skeleton, and  $k$  is the permeability. The radial displacement predicted by our implementation (Figure 6.2) shows good agreement with the analytical solution provided by Armstrong et al. [1984], and yields a stable solution. The same test problem has also been used to verify other large deformation poroelastic software such as FEBio Maas et al. [2012]. The effect of the stabilization parameter on the numerical solution for this test problem has been investigated in [Berger et al., 2014a], and has shown to have a negligible effect, since the stabilization parameter can be chosen to be very small in 3D.

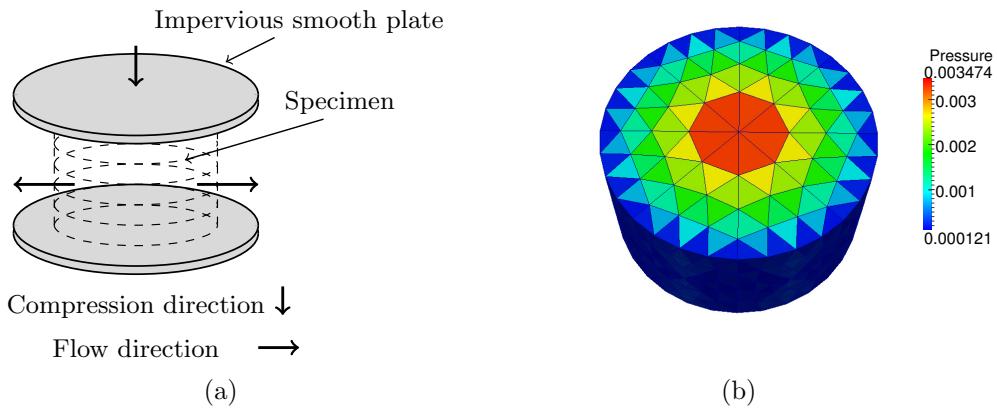


Figure 6.1: (a) Sketch of the test problem. The porous medium is being compressed via a smooth impervious plate. The frictionless plate permits the porous medium to expand radially, it then gradually relaxes as the fluid seeps out horizontally. (b) Pressure field solution at  $t = 200s$ , using a mesh with 3080 tetrahedra.

### 6.2.2 Terzaghi's problem

This a classic geomechanics example of an analytical poroelastic solution, and has been used to investigate finite element pressure oscillations, caused by overshooting of the numerical solution near the boundary Murad and Loula [1994], White and Borja [2008]. The domain consists of a porous column of unit height, bounded at the sides and bottom by rigid and impermeable walls. The top is free to drain ( $p_D = 0$ ) and has a downward traction force,  $p_0$ , applied to it. The boundary and initial conditions for this 1D problem

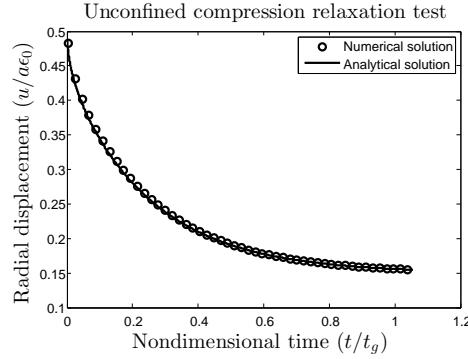


Figure 6.2: Radial expansion versus time calculated using the small deformation analytical solution, and using the proposed numerical method.

can be written as

$$t_N = -p_0, \quad p_D = 0 \quad \text{on } x = 0, \quad (6.18a)$$

$$u = 0, \quad z = 0, \quad \text{on } x = 1, \quad (6.18b)$$

$$u = 0, \quad z = 0, \quad p = 0 \text{ in } (0, 1). \quad (6.18c)$$

The analytical pressure solution, in non-dimensional form is given by

$$p^* = \sum_n^{\infty} \frac{2}{\pi(n + 1/2)} \sin(\pi(n + 1/2)) \exp^{-\pi(n+1/2)(\lambda+2\mu)\kappa t}. \quad (6.19)$$

When the poroelastic medium is subjected to the sudden loading, the saturating fluid undergoes an overpressurization. Subsequently this overpressure progressively vanishes, owing to the diffusion process of the fluid towards the boundary at the top of the column, which remains drained. For a detailed explanation and derivation of this solution see [Wriggers, 2008, section

5.2.2]. We discretized the column using 60 hexahedra elements and solved the problem using the proposed stabilized low-order finite element method and a higher-order inf-sup stable finite element method that uses a piecewise linear pressure approximation. The simulation results of the pressure for the two methods, taken at  $t = 0.01\text{s}$  and  $t = 1\text{s}$  are shown in Figure 6.3. The material parameters used for the simulation can be found in Table 6.2. At  $t = 0.01\text{s}$  the piecewise linear (continuous) approximation suffers from overshooting due to the boundary layer solution (Figure 6.3a). The proposed method, which uses a piecewise constant pressure approximation does not suffer from this problem, and captures the pressure boundary layer solution reliably (Figure 6.3b). At  $t = 1\text{s}$  the boundary layer has diffused away and both the piecewise linear (Figure 6.3c) and piecewise constant (Figure 6.3d) approximation yield satisfactory results.

Parameter	Description	Value
$\kappa_0$	Dynamic permeability	$10^{-5} \text{ m}^3 \text{ s kg}^{-1}$
$\nu$	Poisson ratio	0.25
$E$	Young's modulus	$100 \text{ kg m}^{-1} \text{ s}^{-2}$
$\Delta t$	Time step used in the simulation	0.01 s
$T$	Final time of the simulation	1 s
$\delta$	Stabilization parameter	$2 \times 10^{-5}$

Table 6.2: Parameters used for Terzaghi's problem.

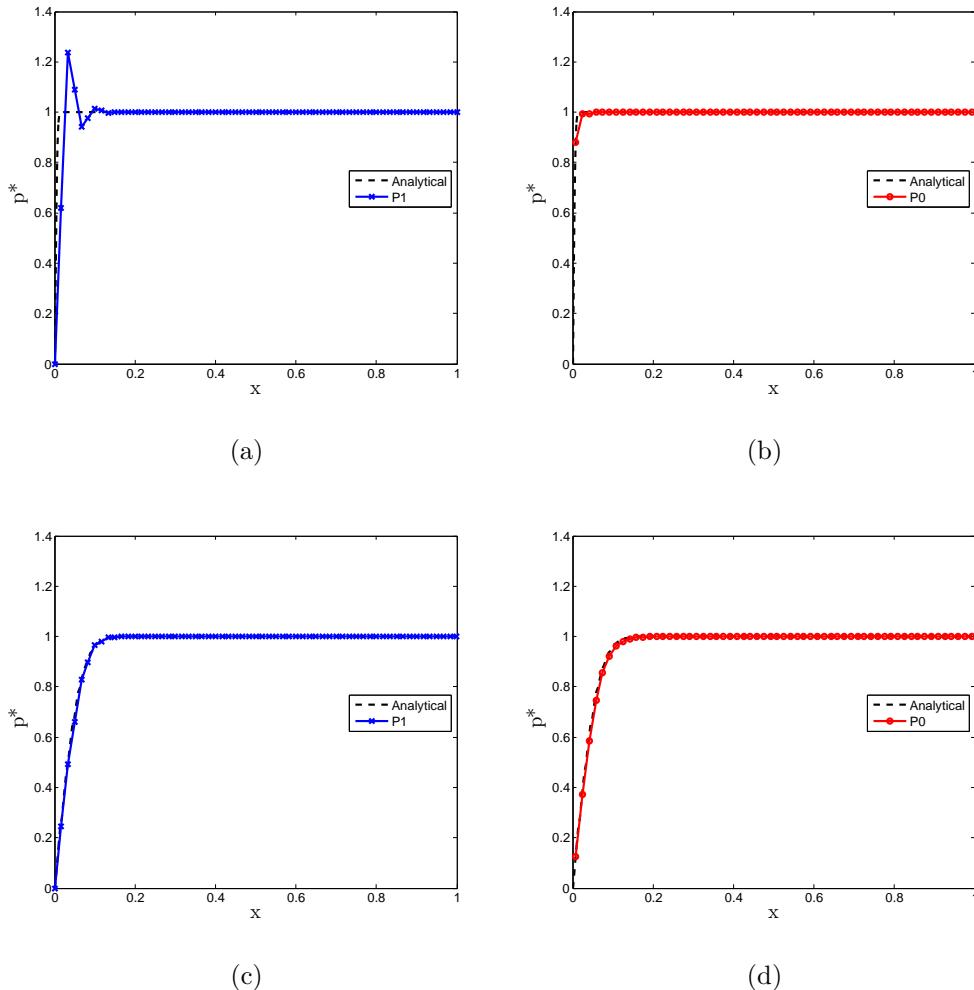


Figure 6.3: The normalized pressure solution at  $t = 0.01s$  using a piecewise linear (a) and piecewise constant (b) approximation, and at  $t = 1s$  using a piecewise linear (c) and piecewise constant (d) approximation.

### 6.2.3 Swelling test

To illustrate the behavior of the poroelastic model under large deformations we present a swelling test problem similar to the one tested in Chapelle et al. [2010]. This test problem also highlights the methods ability to reliably capture jumps in the pressure solution due to changes in material parameters.

**Setup:** In this swelling test, no external force is applied on the skeleton but a fluid pressure gradient is imposed between two opposite faces (at  $X = 0$  and  $X = 1$ ). For the skeleton, normal displacements are zero for the planes  $X = 0$ ,  $Y = 0$  and  $Z = 0$ . For the fluid, there are neither sinks nor sources. The pressure  $p_D$  on the inlet face  $X = 0$  increases from zero to a limit value of 10kPa ( $p_D = 10^4(1 - \exp(-t^2/0.25))$  Pa). On the outlet face ( $X = 1$  in the initial configuration), pressure is zero,  $p_D = 0$ . A null flux condition is applied for the fluid velocity on the four other faces ( $Y = 0, 1$ ,  $Z = 0, 1$ ). In addition to this we reduce the permeability of a small section of the cube by a factor of 500, a diagram highlighting the affected area ( $X < 0.5, Y > 0.5, Z < 0.5$ ) is shown in Figure 6.4a.

**Simulation results:** Fluid enters the medium from the inlet face. The velocity increases as the inlet pressure rises and the cube swells like a sponge undergoing large deformation as shown in Figure (6.4b). When the inlet pressure has reached its limit value, a steady state takes place in most of the tissue such that the elastic forces are in equilibrium with the pressure increase due the added fluid. The pressure decreases roughly linearly with  $x$ , the increase in volume also follows a similar pattern . This is shown by the

plots of  $p$  in Figure (6.5a) and  $J$  in Figure (6.5b) over time for three points at  $(0, 0, 1)$ ,  $(0.5, 0, 1)$  and  $(1, 0, 1)$ , in the reference configuration, across the cube. The position of these points are show as a red, blue and green balls respectively in Figure 6.4a. This close to linear relationship between the pressure and the volume increase can be explained by the the Neo-Hookean material law that has been used in this swelling test. We have also recorded the pressure and volume change at the point  $(0, 1, 0)$  (black ball in Figure 6.4a) where the permeability has been reduced by a factor of 500. Because of the lower permeability this part of the cube takes longer to swell and reach its equilibrium state. Due to the sudden change in permeability within the cube there is a jump in the pressure field as can be seen in Figure 6.4b. The proposed method is able to capture this jump due to the piecewise constant elements that are used to approximate the pressure. Due to the assigned no-flux boundaries and the area of reduced permeability, the fluid mainly flows along the x-direction and around the area of reduced permeability, see Figure 6.4b. The parameters chosen for this test problem are shown in Table 6.3.

Parameter	Value
$\kappa_0$	$10^{-5} \text{ m}^3 \text{s kg}^{-1}$
$\nu$	0.3
$E$	$8000 \text{ kg m}^{-1} \text{s}^{-2}$
$\Delta t$	0.02 s
$T$	20 s
$\delta$	$10^{-4}$

Table 6.3: Parameters used for the swelling test problem.

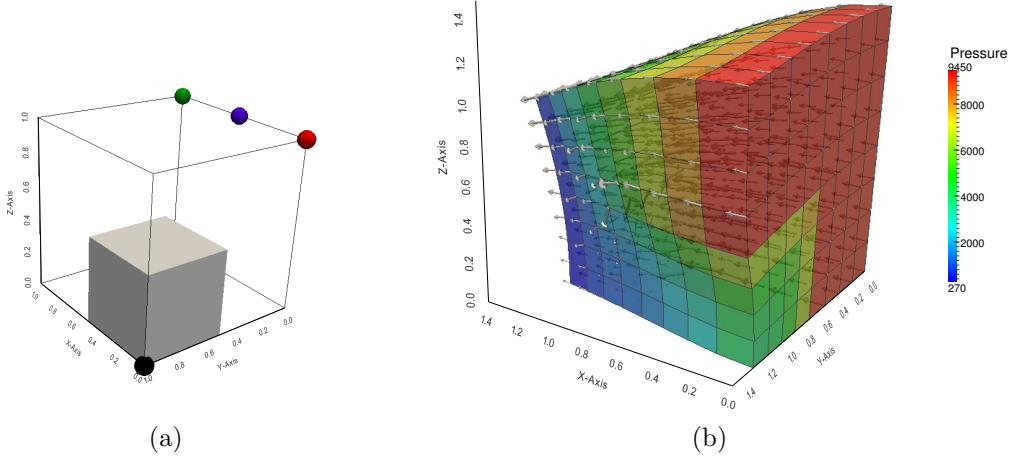


Figure 6.4: (a) Initial simulation setup. The grey cube represents the area of reduced permeability. The colored balls highlight the position of the points used for tracking the pressure and volume change during the simulation, shown in Figures 6.5a and 6.5b. (b) The deformed cube after 1s. The pressure solution is plotted and the jumps in pressure at the interface between the high and low permeability regions can clearly be seen. The arrows illustrate the fluid-flux profile.

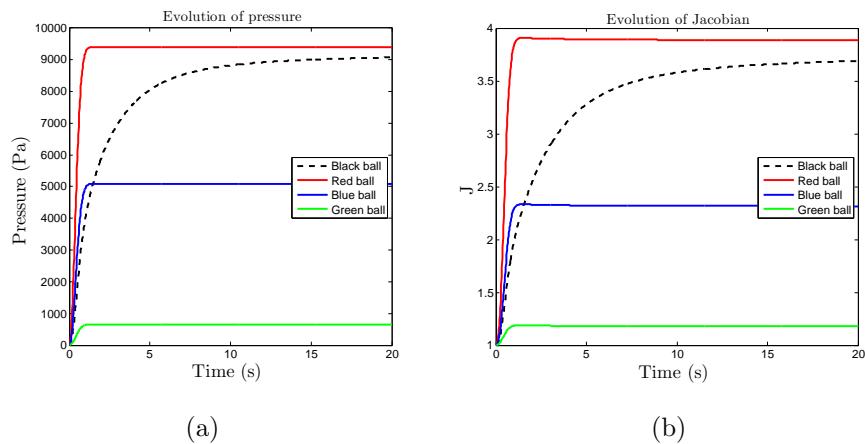


Figure 6.5: Pressure (a) and volume change,  $J$ , (b) are plotted against time for four points,  $(0, 0, 1)$  (red),  $(0.5, 0, 1)$  (blue),  $(1, 0, 1)$  (green), and  $(1, 0, 1)$  (black) in the reference configuration. The position of these balls is also shown in Figure 6.4a.

## 6.3 Conclusion

Stabilized low-order methods can offer significant computational advantages over higher order approaches. In particular, one can employ meshes with fewer degrees of freedom, fewer Gauss points, and simpler data structures. The additional stabilization terms can also improve the convergence properties of iterative solvers. These factors become crucial when considering large-scale, coupled, three-dimensional problems. There has also been a need for a method that is able to overcome both pressure oscillations due to the mixed finite element formulation not satisfying the LBB (inf-sup) condition and due to steep pressure gradients in the solution.

The main contribution of this paper has been to extend the local pressure jump stabilization method [Burman and Hansbo, 2007], already applied to three-field linear poroelasticity [Berger et al., 2014a] to the large deformation case. Thus, the proposed scheme is built on an existing scheme, for which rigorous theoretical results about the stability and optimal convergence have been proven, and numerical experiments have confirmed its ability to overcome spurious pressure oscillations. Due to the discontinuous pressure approximation, sharp pressure gradients due to changes in material coefficients or boundary layer solutions can be captured reliably, circumventing the need for severe mesh refinement. Also, the addition of the stabilization term introduces minimal additional computational work, can be assembled locally on each element using standard element information, and leads to a sym-

metric addition to the original system matrix, thus preserving any existing symmetry. As the numerical examples have demonstrated, the stabilization scheme is robust and leads to high-quality solutions.

# **Chapter 7**

## **A poroelastic model coupled to a fluid network with applications in lung modelling**

### **7.1 Model assumptions**

We will now give a review of the main modelling assumptions and how they might affect the model's ability to predict deformation and ventilation within healthy and diseased lungs.

### 7.1.1 Approximating lung parenchyma using a poroelastic medium

**Averaging over the tissue:** One of the major assumptions is that we can apply averaging techniques to approximate the lung parenchyma using a poroelastic continuum description. This makes our model computationally tractable and allows us to use the well studied theory of poroelasticity to couple the air with the tissue. In publications such as Owen and Lewis [2001] homogenisation techniques have already been applied to approximate the lung parenchyma as a poroelastic continuum model. The use of averaging methods to obtain a continuum model can be further supported by looking at the different length scales and structures of the tissue. For the microscopic length scale denoted by  $l$  of the parenchyma we will use the diameter of an alveoli that can be approximated from Figure (2.1c) to be 0.02 cm. The macroscopic length scale  $L$  can be taken to be the diameter of a segment which measures around 4 cm of tissue. So the ratio of the different length scales is small i.e  $\epsilon := \frac{l}{L} \approx 0.005 \ll 1$ . This along with the assumption that the structure of an acini is porous (see Figure 2.1b) and periodic supports the use of averaging techniques over the tissue to obtain a continuum description in the form of a poroelastic medium. To further simplify the poroelastic equations we assume that the poroelastic continuum can be described by a solid phase (blood and tissue) and a fluid phase (air), where both phases are assumed to be incompressible. The interaction between the fluid pressure

and the deformation of the solid skeleton is assumed to obey the effective stress principle. Note that by averaging over the tissue we are not able to model individual alveoli and instead have to rely on macroscopic parameters such as the permeability and elasticity coefficients. The diseases we are interested in modelling usually affect whole regions of alveoli (lung tissue), thus, by changing the macro-scale parameters over the affected area of tissue we are still able to model changes in the tissue due to disease.

**Ignoring blood flow:** Apart from the collagen, fibers and air the other major component in the lung is blood. From Figure 2.1c it can be seen that the volume taken up by collagen and elastin fibers is similar to the volume occupied by the capillaries filled with blood. In fact, the space not occupied by air is about 7% of the parenchyma's volume and is made up of 50% capillary blood and 50% of collagen and elastin fibers Weichert [2011]. Also the density of blood is similar to the density of tissue and much larger than that of air ( $1060 \text{ kg m}^{-3} \gg 1.18 \text{ kg m}^{-3}$ ). Due to the large accelerations present in blood, the inertia and gravity forces of blood acting on the surrounding tissue could therefore be of importance when predicting deformation and ventilation in the lung. Since the capillaries are constantly filled with blood and the density of blood is similar to that of alveolar tissue we will make the assumption that the blood is simply part of the tissue (solid phase) and thus ignore accelerations and any redistribution of blood during breathing. It is still unclear what effect this assumption will have on the accuracy of

the model and this should be taken into account when evaluating results. However due to the modular framework of the poroelastic theory it should be possible to include the blood as a separate phase in a future version of the model.

**Assuming incompressibility of the solid and the fluid:** Blood and tissue are often assumed to be incompressible. Under physiological conditions, air is also often assumed to be incompressible Ismail et al. [2013].

**Ignoring solid inertia forces:** The maximum accelerations of the solid skeleton can be assumed to be found near the diaphragm, since this is where the largest range of motion occurs. We assume that one breathing cycle takes around 3 seconds, the distance moved by the diaphragm during one breath is around 5 cm and the motion can be described by a simple sinusoidal function. Then simple calculations yield that the maximum accelerations have an approximate magnitude of  $0.02\text{ ms}^{-2}$ . Compared to the acceleration of gravity this is negligible, and it is therefore reasonable to ignore the inertia forces in the tissue.

**Ignoring fluid inertia forces:** The fluid's Reynolds number in the lower airways has been estimated to be around 1 to 0.01 Pedley et al. [1970], see Table 2.1. Due to this relatively low Reynolds we will choose to ignore fluid inertia forces. Tal also showed in their whole lung model that inertia forces

can be neglected during quiet breathing.

**Ignoring viscous forces in the fluid:** The magnitude of the viscous stress in the fluid  $\|\nabla\sigma_{visc}\|$  can be evaluated according to Coussy [2004] as  $\|\nabla\sigma_{visc}\| = \frac{\phi\mu^f}{lL} \|\mathbf{v}^f - \mathbf{v}^s\|$ , where  $l$  and  $L$  are the microscopic and macroscopic characteristic length scales, respectively. This can be balanced with the order of magnitude of  $\left\| \frac{\phi^2}{\kappa} (\mathbf{v}^f - \mathbf{v}^s) \right\|$  that describes the viscous resistance opposed by the shear stress to the fluid flow from the drag at the internal walls of the porous network. To see whether the viscous forces play an important role we take the ratio between these two forces

$$\|\nabla\sigma_{visc}\| / \left\| \frac{\phi^2}{\kappa} (\mathbf{v}^f - \mathbf{v}^s) \right\| = \frac{\mu_f \kappa}{\phi l L}.$$

Using parameter estimates for healthy human tissue found in Table 7.1, and the previous estimates of the length scales, we get a non-dimensional value for (7.1.1) of about  $0.002 \ll 1$ , indicating that viscous forces are insignificant. However, if we were to consider diseased states such as emphysema, where large areas of lung tissue completely break down leaving big holes, see Figure 7.1, it could be argued that the permeability here increases by large amounts that would also cause the above estimate to increase, and suggest that in this case viscous forces could well play an important role, making it important to include them in our model. For simplicity we will choose to ignore this viscous stress for now but may well include it in a future version

of the model by replacing the Dracy flow model with a Brinkman flow, or even a Stokes model for big holes where the homogenisation assumption of a porous structure is not valid.

**Constitutive laws:** We use an isotropic hyperelastic strain energy law to model the tissue. However for diseased cases such as emphysema (Figure 7.1) or fibrosis this isotropic assumption may become invalid. For simplicity we also assume that the permeability of the tissue is isotropic. Again the assumption of an isotropic permeability may well become invalid during disease. With the increasing availability of detailed imaging data on the structure of the lung along with further modelling work on constitutive laws for lung tissue, anisotropic laws could be used in the future.

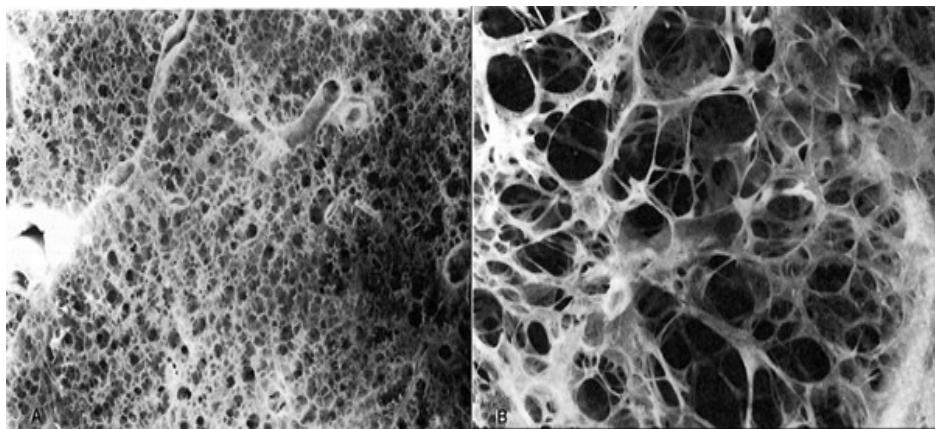


Figure 7.1: Left, a cross section of healthy parenchyma. Right, a cross section of diseased (emphysemic) lung parenchyma, with big holes appearing. Images are reproduced from G. Snell, ctsnet.org.

### 7.1.2 Approximating the airways using a fluid network model

In order to make the coupled model computationally feasible we assume that a simple laminar flow model can describe the air flow in the airways. We will make the common **Poisuelle flow assumption**. This flow assumption is also made in Swan et al. [2012] where the air flow in a whole airway tree, from trachea down to the final bronchioles was assumed to be governed by Poiseuille flow. The Poisuelle flow assumption requires flow to be fully developed and laminar which is often true for long pipes and fluid flowing at low Reynolds numbers. This may also be true for the smaller airways but is certainly false for the larger upper airways where high Reynolds number flows can be measured, see Table 2.1. Such a model will therefore not be able to capture the high Reynolds number flows and turbulent effects that are known to exists in the upper airways. However diseases affecting the airway tree can still be modelled effectively by changing resistance (airway radius) parameters in the network flow model.

## 7.2 A mathematical lung model

We will now recall the quasi-static fully incompressible poroelasticity equations, used for modelling the lung parenchyma, a simple fluid network model to model the airway tree, and outline how these two parts are coupled together.

### 7.2.1 A poroelastic model for lung parenchyma

Before presenting the final balance equations of poroelasticity, we give a short review of the kinematics required to model a poroelastic medium undergoing large deformations. A more detailed description and full derivation of the poroelastic equations can be found in Berger et al. [2014b] and Coussy [2004].

#### Kinematics

In this section we will briefly outline the theory required for modelling of the porous medium made up of a solid and a fluid phase. The kinematic quantities described here will be associated with the solid skeleton, since this also describes the motion of the fluid domain. Let the volume  $\Omega_0$  be the undeformed Lagrangian (material) reference configuration and let  $\mathbf{X} = \{X\mathbf{e}_1 + Y\mathbf{e}_2 + Z\mathbf{e}_3\}$  indicate the position of a solid particle in  $\Omega_0$  at  $t = 0$ , where  $X, Y$  and  $Z$  are the components of the position with respect to the standard orthonormal basis  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$  for  $\mathbb{R}^3$ . The position of a solid particle in the current Eulerian (spatial) configuration  $\Omega_t$  is given by  $\mathbf{x} = \{x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3\}$ , with  $\mathbf{x} = \boldsymbol{\chi}(\mathbf{X}, t)$ . The deformation map,  $\boldsymbol{\chi}(\mathbf{X}, t)$ , is a continuously differentiable, invertible mapping from  $\Omega_0$  to  $\Omega_t$ . The displacement field is given by

$$\mathbf{u}(\mathbf{X}, t) = \boldsymbol{\chi}(\mathbf{X}, t) - \mathbf{X}. \quad (7.1)$$

The deformation gradient tensor,

$$\mathbf{F} = \frac{\partial \boldsymbol{\chi}(\mathbf{X}, t)}{\partial \mathbf{X}} = \mathbf{I} + \frac{\partial \mathbf{u}(\mathbf{X}, t)}{\partial \mathbf{X}}, \quad (7.2)$$

maps a material line element in the reference configuration  $d\mathbf{X}$ , to a line element  $d\mathbf{x}$  in the current configuration, i.e.  $d\mathbf{x} = \mathbf{F}d\mathbf{X}$ . The symmetric right Cauchy-Green deformation tensor is given by  $\mathbf{C} = \mathbf{F}^T \mathbf{F}$ . The jacobian is defined as  $J = \det(\mathbf{F})$ , and represents the change in an infinitesimal small volume from the reference to the current configuration.

### Volume fractions

The porous medium is assumed to be saturated consisting of a solid part and a fluid part. The fluid part accounts for volume fractions  $\phi_0(\mathbf{X}, t = 0)$  and  $\phi(\mathbf{x}, t)$  of the total volume in the reference and the current (deformed) configuration. The fractions for the solid are thus given by  $1 - \phi_0$  and  $1 - \phi$  in the reference and the current configuration respectively,  $\phi$  is also often known as the porosity. In standard elasticity theory  $\rho$  denotes the density of the solid, and for an incompressible solid we have that  $\rho d\Omega_t = \rho_0 d\Omega_0$ . We now change the definition of  $\rho$  to be the density of the mixture in the current configuration which is given by

$$\rho = \rho^s(1 - \phi) + \rho^f\phi \quad \text{in } \Omega_t, \quad (7.3)$$

where  $\rho^s$  and  $\rho^f$  are the densities of the fluid and solid, respectively. We will also assume that both the solid and the fluid are incompressible, such that  $\rho^s = \rho_0^s$  and  $\rho^f = \rho_0^f$ . Due to mass conservation and the incompressibility of both the solid and the fluid phase we have

$$J = \frac{1 - \phi_0}{1 - \phi}. \quad (7.4)$$

Here  $J$  represents the change in volume of the solid skeleton. The solid skeleton includes the tissue and the voids occupied by the fluid (air).

### Poroelastic equations

Having made the assumptions in section 7.1 for the tissue we are left with the following poroelastic system of equations. We define the boundary  $\partial\Omega_t = \Gamma_d \cup \Gamma_t$  for the displacement and stress boundary conditions of the poroelastic mixture, and  $\partial\Omega_t = \Gamma_p \cup \Gamma_f$  for the pressure and flux boundary conditions of the fluid, with an outward pointing unit normal  $\mathbf{n}$ . The strong problem is

to find  $\chi(\mathbf{X}, t)$ ,  $\mathbf{z}(\mathbf{x}, t)$  and  $p(\mathbf{x}, t)$  such that

$$-\nabla \cdot (\boldsymbol{\sigma}_e - p\mathbf{I}) = \rho\mathbf{f} \quad \text{in } \Omega_t, \quad (7.5a)$$

$$\kappa^{-1}\mathbf{z} + \nabla p = \rho^f\mathbf{f} \quad \text{in } \Omega_t, \quad (7.5b)$$

$$\nabla \cdot (\mathbf{v}^s + \mathbf{z}) = g \quad \text{in } \Omega_t, \quad (7.5c)$$

$$\chi = \mathbf{X} + \mathbf{u}_D \quad \text{on } \Gamma_d, \quad (7.5d)$$

$$(\boldsymbol{\sigma}_e - p\mathbf{I})\mathbf{n} = \mathbf{t}_N \quad \text{on } \Gamma_t, \quad (7.5e)$$

$$\mathbf{z} \cdot \mathbf{n} = q_D \quad \text{on } \Gamma_f, \quad (7.5f)$$

$$p = p_D \quad \text{on } \Gamma_p, \quad (7.5g)$$

$$\chi = \mathbf{X} + \mathbf{u}^0, \quad \text{in } \Omega_0, \quad (7.5h)$$

The first equation (7.5a) is the balance of linear momentum for the poroelastic mixture. The elastic stress is denoted by  $\boldsymbol{\sigma}_e$  (defined in section 7.2.1);  $p$  denotes the fluid pressure, and  $\mathbf{f}$  is a body force. This equations is governed by the effective stress concept, which says that the deformations of the skeleton are governed by the so-called ‘effective stress’ or often known as ‘elastic stress’,  $\boldsymbol{\sigma}_e = \boldsymbol{\sigma} + p\mathbf{I}$ , where  $\boldsymbol{\sigma}$  denotes the total stress, which is the quantity we wish to evaluate in the momentum balance equation (7.5a). This principle states that the deformation of the porous solid is only driven by the excess of stress over the pore pressure Boer [2005], Coussy [2010]. Equation (7.5b) is the balance of linear momentum for the fluid, and is also commonly known as Dracy’s law. Here  $\mathbf{z}$  denotes the fluid flux and  $\kappa$  is the permeability tensor,

defined in section 7.2.1. The balance of mass is modelled by equation (7.5c). Here  $\mathbf{v}^s = \frac{\partial}{\partial t}\boldsymbol{\chi}(\mathbf{X}, t)$  denotes the velocity of the solid skeleton. Equations (7.5h)-(7.5h) are the associated boundary and initial conditions. A detailed derivation of all the above equations is given in Berger et al. [2014b].

### Constitutive laws

To close the model (7.5) we need to choose constitutive laws for the permeability and strain energy. The permeability tensor in the current configuration is given by

$$\boldsymbol{\kappa} = J^{-1} \mathbf{F} \boldsymbol{\kappa}_0(\boldsymbol{\chi}) \mathbf{F}^T, \quad (7.6)$$

where  $\boldsymbol{\kappa}_0(\boldsymbol{\chi})$  is the permeability in the reference configuration, which may be chosen to be some (nonlinear) function dependent on the deformation. We will use the same law that has already been proposed in Kowalczyk and Kleiber [1994] to model lung parenchyma,

$$\boldsymbol{\kappa} = \boldsymbol{\kappa}_0 \left( J \frac{\phi}{\phi_0} \right)^{2/3} \mathbf{I}. \quad (7.7)$$

The elastic (effective) stress tensor given by

$$\boldsymbol{\sigma}_e = \frac{1}{J} \mathbf{F} \cdot 2 \frac{\partial W(\boldsymbol{\chi})}{\partial \mathbf{C}} \cdot \mathbf{F}^T. \quad (7.8)$$

Here  $W(\boldsymbol{\chi})$  denotes a strain-energy law (hyperelastic Helmholtz energy functional) dependent on the deformation of the solid. Exponential strain energy

laws for lung parenchyma exist, for example the popular law by Fung [1975]. However little is known about how the constants in these laws should be interpreted and altered to model weakening of the tissue in an diseased state. Further, the constants in these laws are thought to have no physical meaning [Tawhai et al., 2009]. To make the interpretation of the elasticity constants and dynamics of the model as simple as possible we chose a Neo-Hookean law taken from Wriggers [2008, eqn. (3.119)],

$$W(\mathbf{C}) = \frac{\mu}{2}(\text{tr}(\mathbf{C}) - 3) + \frac{\Lambda}{4}(J^2 - 1) - (\mu + \frac{\Lambda}{2})\ln J. \quad (7.9)$$

The material parameters  $\mu$  and  $\Lambda$  can be related to the more familiar Young's modulus  $E$  and the Poisson ratio  $\nu$  by  $\mu = \frac{E}{2(1+\nu)}$  and  $\Lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}$ . The values of these constants for modeling lung tissue have been investigated De Wilde et al. [1981] and are shown in Table 7.1.

### 7.2.2 A network flow model for the airway tree

The flow rate  $Q_i$  through the  $i$ th pipe segment in the fluid network is given by the pressure-flow relationship

$$P_{i,1} - P_{i,2} = R_i Q_i, \quad (7.10)$$

where  $R_i = \frac{8l\mu_f}{\pi r^4}$  is the Poiseuille flow resistance of the pipe segment ( $r$  is the radius,  $l$  is the length of the pipe,  $\mu_f$  is the dynamic viscosity) and  $P_{i,1}$  and

$P_{i,2}$  are the pressures of the two nodes at each end of the pipe segment. We also have conservation of flow at branches such that

$$Q_i = \sum_{Q_{i,j} \in Q_i} Q_{i,j}, \quad (7.11)$$

where  $Q_{i,j}$  are the flow rates of the children branches of the  $i$ th flow segment. The outlet pressure of the fluid network is set using the boundary condition  $P_0 = \hat{P}$ .

### 7.2.3 Coupling the fluid network to the poroelastic model

We split our domain into  $N$  separate subdomains such that  $\Omega_t = \sum_i^N \Omega_t^i$ . The introduction of subdomains allows each endpoint of the fluid network to supply and remove fluid from the poroelastic medium at different spatial locations as is shown in Figure 7.2a. The  $i$ th subdomain  $\Omega_t^i$  is defined as the volume (set of elements in discretised space, see Figure 7.2b for a simple example) closest to the position of the  $i$ th inlet, denoted by  $\text{pos}(P_{di})$ .

$$\Omega_t^i := \{\mathbf{x} \in \Omega_t : \|\mathbf{x} - \text{pos}(P_{di})\| < \|\mathbf{x} - \text{pos}(P_{dj})\|, j = 1, 2, \dots, N, j \neq i\}. \quad (7.12)$$

For notational purposes we have added subscript  $di$  to the most distal branches that have no further conducting branches coming from them but instead enter a group of acinar units (approximated by the poroelastic model), this

would be at around generation 13-17.

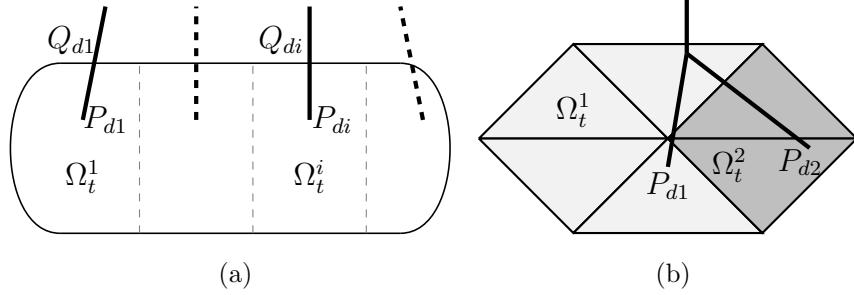


Figure 7.2: This figure shows how the domain can be split up into subdomains depending on the position of end points of the fluid network, in a continuous setting (a), and a simple example of a finite element discretization using piecewise constant pressure elements (b).

We couple the airway network by adding the flow contribution from each distal airway to the poroelastic domain as a source term in the poroelastic mass conservation equation (7.5c), such that

$$\nabla \cdot (\mathbf{v}^s + \mathbf{z}) = Q_{di} \quad \text{in } \Omega_t^i \quad (7.13)$$

We also couple the airway network to the poroelastic domain by setting the average pressure in the poroelastic domain within  $\Omega_t^i$  to be the same as the corresponding distal pressure node  $P_{di}$  of the flow segment  $Q_{di}$ .

$$\frac{1}{|\Omega_t^i|} \int_{\Omega_t^i} p = P_{di}, \quad (7.14)$$

where  $|\Omega_t^i|$  denotes the volume of the segment  $\Omega_t^i$ . Equation (7.14) enforces the assumption that the end pressure in a terminal bronchiole is the same as

the alveolar pressure in the surrounding tissue.

#### 7.2.4 Summary of the complete lung model

The governing equations of the whole fully coupled poroelastic-fluid-network model are as follows:

$$-\nabla \cdot (\boldsymbol{\sigma}_e - p\mathbf{I}) = \rho\mathbf{f} \quad \text{in } \Omega_t, \quad (7.15a)$$

$$\kappa^{-1}\mathbf{z} + \nabla p = \rho^f\mathbf{f} \quad \text{in } \Omega_t, \quad (7.15b)$$

$$\nabla \cdot (\mathbf{v}^s + \mathbf{z}) = Q_{di} \quad \text{in } \Omega_t^i, \quad (7.15c)$$

$$\boldsymbol{\chi} = \mathbf{X} + \mathbf{u}_D \quad \text{on } \Gamma_d, \quad (7.15d)$$

$$(\boldsymbol{\sigma}_e - p\mathbf{I})\mathbf{n} = \mathbf{t}_N \quad \text{on } \Gamma_t, \quad (7.15e)$$

$$\mathbf{z} \cdot \mathbf{n} = q_D \quad \text{on } \Gamma_f, \quad (7.15f)$$

$$p = p_D \quad \text{on } \Gamma_p, \quad (7.15g)$$

$$\boldsymbol{\chi}(0) = \mathbf{X} + \mathbf{u}^0, \quad \text{in } \Omega_0, \quad (7.15h)$$

$$P_0 = \hat{P}, \quad (7.15i)$$

$$P_{i,1} - P_{i,2} = R_i Q_i, \quad (7.15j)$$

$$Q_i = \sum_{Q_{i,j} \in Q_i} Q_{i,j}, \quad (7.15k)$$

$$\frac{1}{|\Omega_t^i|} \int_{\Omega_t^i} p = P_{di}. \quad (7.15l)$$

### 7.2.5 Implementation

Since the system of equations (7.15) is highly nonlinear, its solution requires a scheme such as Newton's method. In Berger et al. [2014b] a finite element scheme using Newton's method for the solution of the poroelastic equations valid in large deformations (7.5) has already been presented. In this work we adopt the same finite element scheme as presented in [Berger et al., 2014b] for solving the poroelastic equations and expand the linear system (discretized linearization) to include additional matrices required for solving the fluid network and its coupling to the poroelastic medium (equations (7.15c,7.15j,7.15k,7.15l)). This results in a monolithic coupling scheme which ensures good convergence even for problems with strong coupling interactions between the poroelastic medium and the fluid network (see section 7.4.2). Let  $\mathbf{u}_i^n$  denote the fully discrete solution of all the poroelastic and fluid network variables at the  $i$ th step within the Newton method at time  $t^n$ , and let  $\xi\mathbf{u}_{i+1}^n$  denote the corresponding solution increment vector used to update the current solution within the Newton method. Let  $\mathbf{K}$  denote the stiffness matrix (discretized linearization of the full lung model (7.15)), and  $\mathbf{R}$  the residual vector, see section 7.6.1 for details on how these are built. To solve the nonlinear poroelastic problem using Newton's method at a particular time step, we perform the following steps:

---

**Algorithm 2** Fully discrete Newton's method (for a particular time step  $t^n$ )

---

Initialize  $\mathbf{u}_0^n = \mathbf{u}^{n-1}$

Iteration loop  $i = 0, 1, \dots$  until convergence.

1. Assemble  $\mathbf{R}(\mathbf{u}_i^n, \mathbf{u}^{n-1})$  and  $\mathbf{K}(\mathbf{u}_i^n)$ .
2. Compute the solution increment from the linear system:  
$$\mathbf{K}(\mathbf{u}_i^n) \xi \mathbf{u}_{i+1}^n = -\mathbf{R}(\mathbf{u}_i^n, \mathbf{u}^{n-1}).$$
3. Compute the new solution:  $\mathbf{u}_{i+1}^n = \mathbf{u}_i^n + \xi \mathbf{u}_{i+1}^n$ .
4. Update the mesh,  $\Omega_t = \mathbf{X} + \mathbf{u}_i^n$ .
5. Test for convergence:  $\|\mathbf{R}(\mathbf{u}_i^n, \mathbf{u}^{n-1})\| > \text{TOL} \rightarrow$  Set  $i = i + 1$  go to 1,

$$\|\mathbf{R}(\mathbf{u}_i^n, \mathbf{u}^{n-1})\| \leq \text{TOL} \rightarrow$$

Stop and move to the next time step.

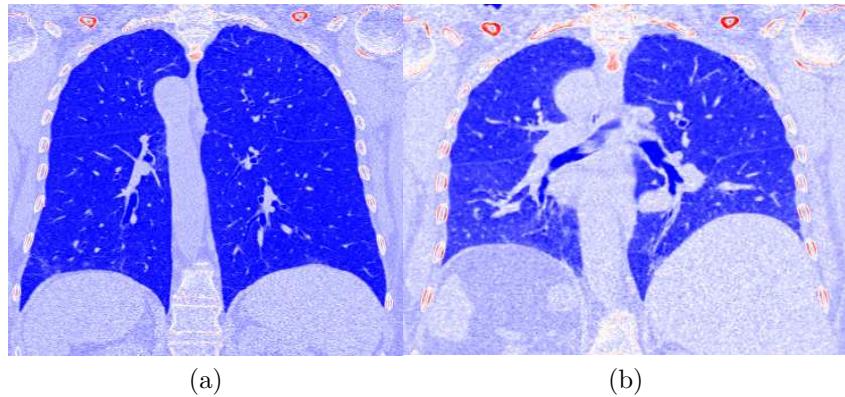
---

For the implementation we use the C++ library libmesh [Kirk et al., 2006], and the multi-frontal direct solver mumps [Amestoy et al., 2000] to solve the resulting linear system. We set the relative tolerance to be  $\text{TOL} = 10^{-4}$ .

## 7.3 Model generation

### 7.3.1 Mesh generation

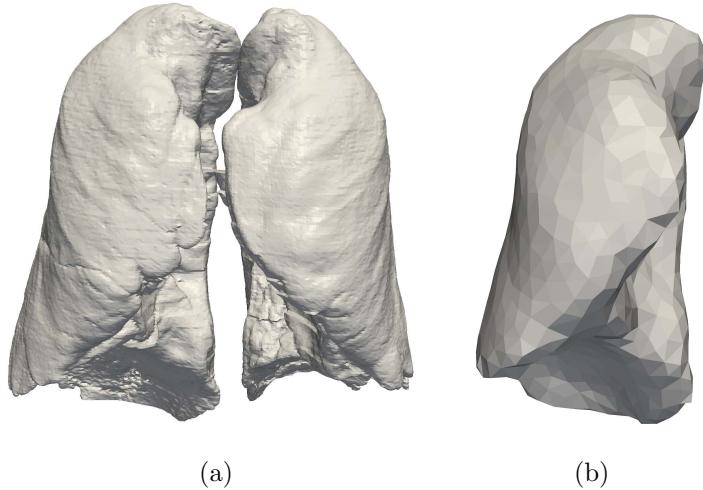
We derive a whole organ lung model, of the right lung, from a high-resolution CT image taken at total lung capacity (TLC) and functional residual capacity (FRC), shown in Figure 7.3. The bulk lung is first segmented from the CT data using standard image processing techniques, see Figure 7.4a. We then use the open-source image processing toolbox iso2mesh [Fang and Boas, 2009] to generate a Tetrahedral mesh containing 38369 elements as shown in Figure 7.4b. The conducting airways are also segmented from the CT data taken at TLC level (Figure 7.5a), and a centerline (Figure 7.5b) with radial information is calculated. To approximate the remaining airways up to generation 13-14 we use a volume filling airway generation algorithm to generate a mesh of the airway tree containing 13696 nodes (Figure 7.5c) [Oxford ref]. The complete mesh containing both the lung and the airway tree is shown in Figure 7.5d.



(a)

(b)

Figure 7.3: (a) Sagittal CT slice at TLC, (b) sagittal CT slice at FRC.



(a)

(b)

Figure 7.4: (a) Segmentation of the left and right lung at TLC, (b) tetrahedral mesh of the right lung at TLC.

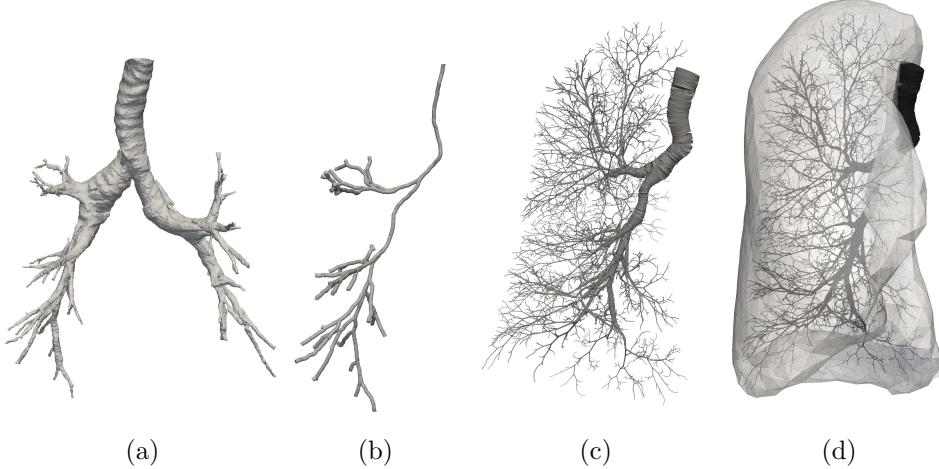


Figure 7.5: (a) Segmented upper airways, (b) centerline of the upper airways in the right lung, (c) generated airway tree model with radial information, (d) combined lung and airway tree mesh.

### 7.3.2 Reference state, boundary conditions and initial conditions

In this work, we uniformly deflate the lung from FRC, which is known to typically have a mean elastic recoil of about  $0.49 \times 10^3$  Pa [West, 2008], to a configuration, the reference state, in which the internal stresses and strains are zero. The geometry of the reference state is then used as the initial configuration of the lung model. The lung model is then inflated from the reference state to a pre-stressed FRC configuration. From there we simulate tidal breathing. A similar approach has also been used in Lee et al. [1983].

We also register the expiratory (FRC) segmentation to the segmentation at TLC using a simple rigid registration procedure. This yields a rough

estimate of the deformation field for the lung surface from expiration to inspiration. To simulate tidal breathing we assume a sinusoidal breathing cycle and expand the lobe surface from FRC to 40% of TLC (normal inspiratory volume).

$$\mathbf{u}_D(t) = 0.4 \left( 1 + \sin \left( \frac{t\pi}{2} + \frac{3\pi}{2} \right) \right) \mathbf{u}_{D,TLC}. \quad (7.16)$$

Here  $\mathbf{u}_{D,TLC}$  is the deformation of the lung surface from FRC to TLC, obtained using the registration procedure. This results in a physiologically realistic tidal volume of 0.59 litres. We simulate breathing for a total of 8 seconds (2 breathing cycles) resulting in a breathing frequency of 15 breaths per minute. Due to the incompressibility of the poroelastic tissue this also determines the total volume of air inspired/expired and the flowrate at the trachea, see Figure 7.7a and 7.7b respectively. In future this method should be upgraded to include a non-linear registration procedure to give a more accurate description of the lung surface deformation.

For the fluid boundary condition we have that the whole lung is sealed so that no fluid can escape through the lung surface, with  $\mathbf{z} \cdot \mathbf{n} = 0$  along the whole boundary. For the airway network boundary condition we set the outlet pressure of the airway network to zero atmospheric pressure,  $P_0 = 0$ .

### 7.3.3 Simulation parameters

Several parameters for lung tissue elasticity and poroelasticity have been proposed [De Wilde et al., 1981, Lande and Mitzner, 2006, Owen and Lewis,

2001, Werner et al., 2009, Zhang et al., 2004]. There is no consensus in the values in the literature. In this pilot study we have chosen the parameters shown in Table 7.1.

Parameter	Value	Reference
$\phi_0$	0.99	Lande and Mitzner [2006]
$\kappa_0$	$10^{-5} \text{ m}^3 \text{ s kg}^{-1}$	Lande and Mitzner [2006]
$E$	$0.73 \times 10^3 \text{ Pa}$	De Wilde et al. [1981]
$\nu$	0.3	De Wilde et al. [1981]
$\mu_f$	$1.92 \times 10^{-5} \text{ kg m}^{-1} \text{ s}^{-1}$	Swan et al. [2012]
$T$	8s	-
$\Delta t$	0.2s	-
$\delta$	$10^{-5}$	-

Table 7.1: Simulation parameters.

Here  $\delta$  is a stabilization parameter required by the finite element scheme, see Berger et al. [2014b,c] for details. Also note that we have chosen to not include gravity in our model. This is because we will not be comparing against imaging data and do not wish to additionally complicate the interpretation of simulation results.

## 7.4 Model exploration

We simulate tidal breathing in the right lung on a realistic geometry extracted from CT data. The simulations are driven by applying deformation boundary conditions on the lung surface, obtained from image registration.

We explore the behavior of the coupled model under homogeneous tissue conditions and after changing material parameters in the airway tree and the tissue. In particular, we investigate the effect of airway constriction and tissue weakening on the ventilation, stresses and pressure distribution.

In the subsequent analysis the total and elastic stress is calculated as  $\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}$ , where  $\lambda_1, \lambda_2, \lambda_3$  are the three eigenvalues of the stress tensor, respectively. We also use the relative Jacobian, denoted by  $J_V$ , as a measure for ventilation, which is calculated to be the volume ratio between the current state and FRC, and is a direct measure of tissue expansion. Due to the quasi-static nature of the equations differences between the results during the second breath and subsequent breaths were negligible. Therefore only results from the second breath,  $t = 4s$  to  $t = 8s$  are presented. As part of the tissue we will take a sagittal slice as shown in Figure 7.6a. All subsequent figures that do not show time courses are taken at  $t = 5.8s$ , just before reaching peak inhalation for the second time during the simulation.

#### 7.4.1 Normal breathing

To simulate tidal breathing we apply the boundary conditions and simulation parameters previously discussed in sections 7.3.2 and 7.3.3, respectively. Figure 7.7 details the lung tidal volume, flow rate and pressure drop information obtained from simulations of tidal breathing. Due to the incompressibility of the poroelastic medium and the steady state nature of the airway network, the lung tidal volume (Figure 7.7a) and flow rate (Figure 7.7b) follow a sinu-

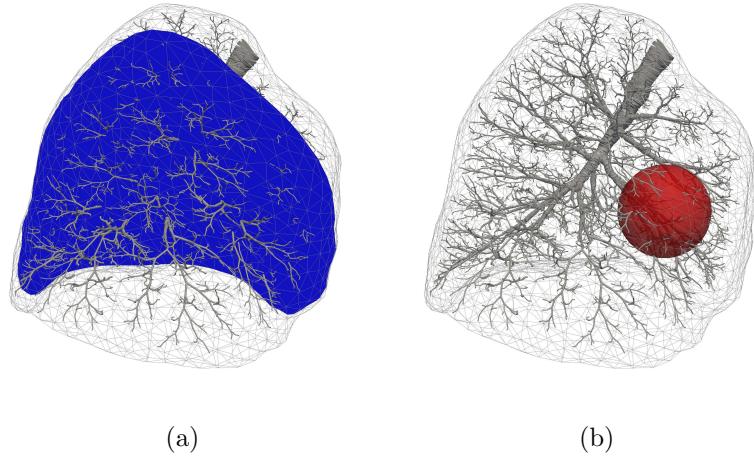


Figure 7.6: The blue sagittal slice indicates the position of subsequent slices used for the data analysis of the tissue (a). The red ball represents the structurally modified region, used to prescribe airway constriction and tissue weakening (b).

soidal pattern that matches the form of the deformation boundary condition prescribed by equation (7.16). The mean pressure drop of the airways, is shown in Figure 7.7c, and agrees with previous simulation studies on full airway trees [Ismail et al., 2013, Swan et al., 2012].

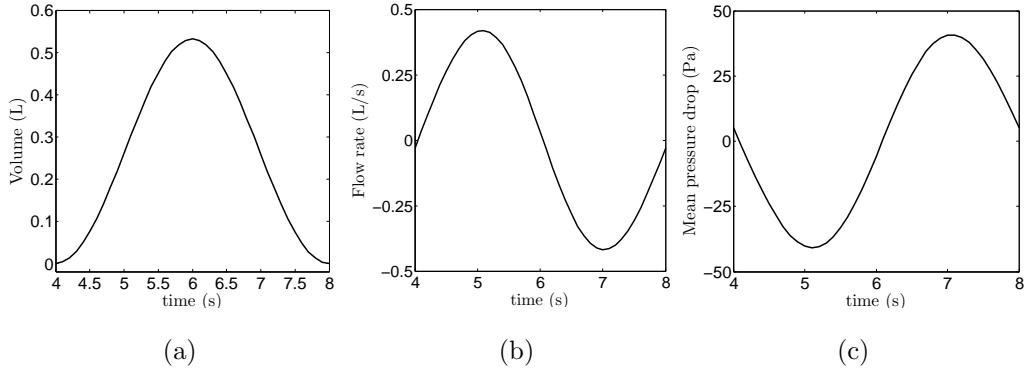


Figure 7.7: Simulated natural tidal breathing: (a) lung tidal volume (volume increase from FRC), (b) flow rate at the inlet, (c) mean pressure drop from the inlet to the most distal branches.

Figure 7.8 shows the change in elastic recoil (total stress) with volume throughout the breathing cycle for three different breathing rates. This curve is also known as a dynamic pressure-volume (PV) curve. The increase of hysteresis in the PV curve and its shift to the right as the breathing rate increases agrees with findings in the literature [Harris, 2005, Rittner and Döring, 2005]. In our model, this shift and widening of the curve can be attributed to the resistance in the airway tree which causes a larger and more heterogeneous pressure drop and flow distribution within the airways at increased flow rates. For tissue regions coupled to airways with high resistance this results in delayed, out of phase, alveolar recruitment (tissue expansion). This delayed filling of tissue during inspiration and emptying during expiration causes delayed pressure gradients that give rise to the hysteresis seen in Figure 7.8. In the literature, hysteresis associated with dynamic pressure volume (PV)

curves is mostly hypothesized to be caused by flow-dependent resistances, pendelluft effects, chest wall rearrangement, and recruitment and derecruitment of lung units [Albaiceta et al., 2008, Harris, 2005, Ranieri et al., 1994].

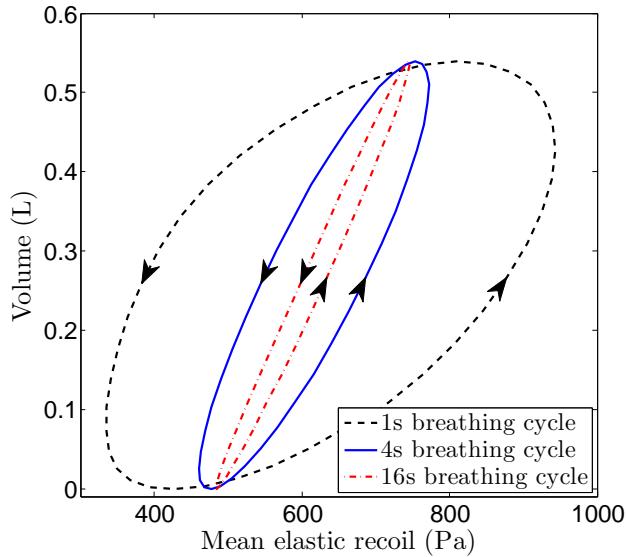


Figure 7.8: Pressure-volume curve: mean elastic recoil (total stress) against lung tidal volume for three different breathing rates.

For the subsequent analysis we have calculated the pathway resistance (Poiseuille flow resistance) from the inlet (right bronchi) to each terminal airway. The pathway resistance for the whole tree is shown in Figure 7.9a. In Figure 7.9b we show the pathway resistance of the terminal airways mapped onto the tissue. In order to quantify the contribution of airway resistance to tissue expansion (ventilation), measured by  $J_V$ , the correlations between pathway resistance and  $J_V$  were calculated in Figure 7.10a. There is a clear correlation between pathway resistance and ventilation ( $J_V$ ). This is ex-

pected since the elastic coefficients have been parameterized homogeneously. The Pearson correlation coefficients is  $-0.55$ , with a p-value  $< 0.0001$ . Figure 7.10b also shows a strong correlation between the pathway resistance and pressure in the poroelastic tissue. Here the Pearson correlation coefficients is also  $-0.55$ , with a p-value  $< 0.0001$ . Also note that for regions that are coupled to terminal branches with a low pathway resistance, positive pressures are possible to enable a pressure gradient that pushes fluid from these well ventilated regions to neighbouring less ventilated regions (collateral ventilation).

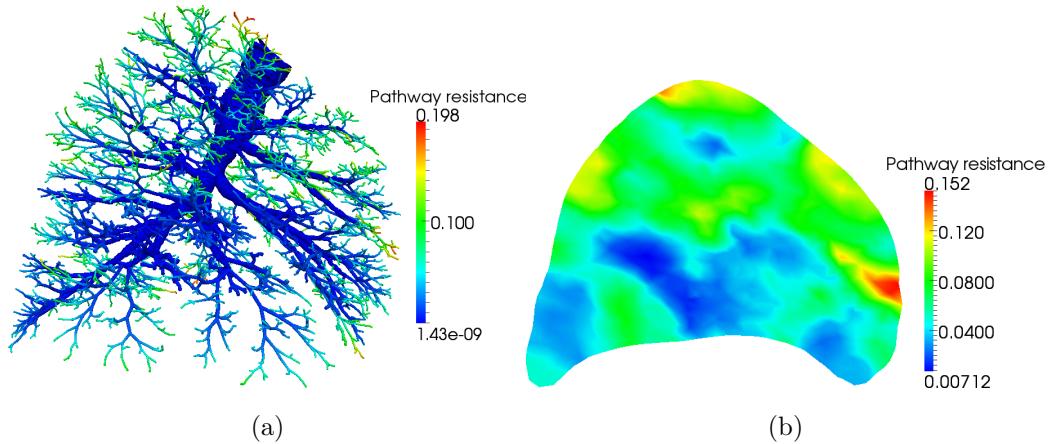


Figure 7.9: Pathway resistance from the inlet to the terminal branches in the airway tree (a) and pathway resistance mapped onto a slice of tissue (b). The deformation of both the tree and the tissue in this figure correspond to the reference configuration.

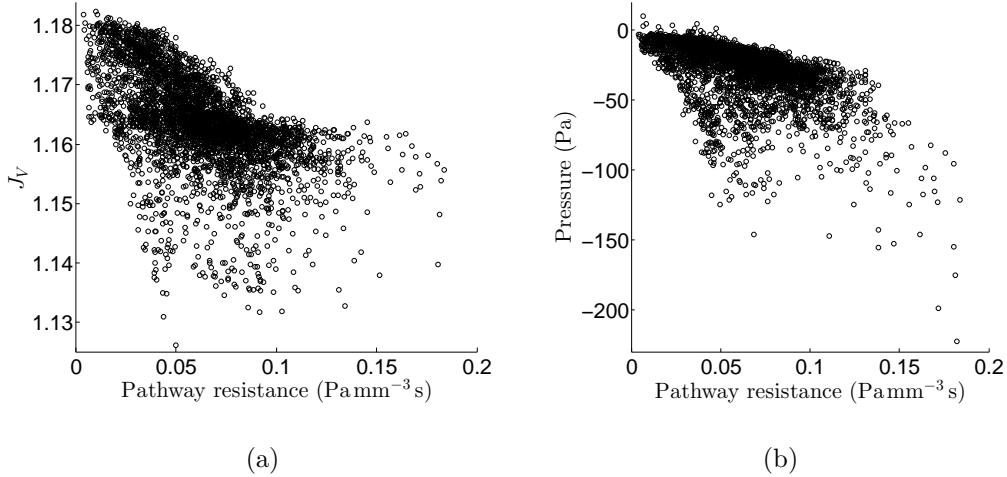


Figure 7.10: (a) Correlation between tissue expansion (ventilation) and resistance of the pathways from the inlet to the terminal branch. (b) Correlation between pressure in the poroelastic medium (alveolar pressure) and pathway resistance.

Figure 7.11a shows the distribution of pressure in the airway tree. As previously discussed the pressure in the airway tree correlates with airway tree resistance, shown in Figure 7.9b. The pressure inside the poroelastic tissue is shown in Figure 7.11c. Again the pressure distribution correlates well with the pathway resistance in the airways, shown in Figure 7.9b. Figure 7.11b shows the pressure on the lung surface. The patchy pressure field is well approximated by the piecewise constant pressure elements employed by the finite element method used to solve the poroelastic equations (see Berger et al. [2014b]). Figure 7.11d shows the distribution of tissue expansion. Although the distribution is heterogeneous the variations are quite small, since the elastic coefficients have been parameterized homogeneously.

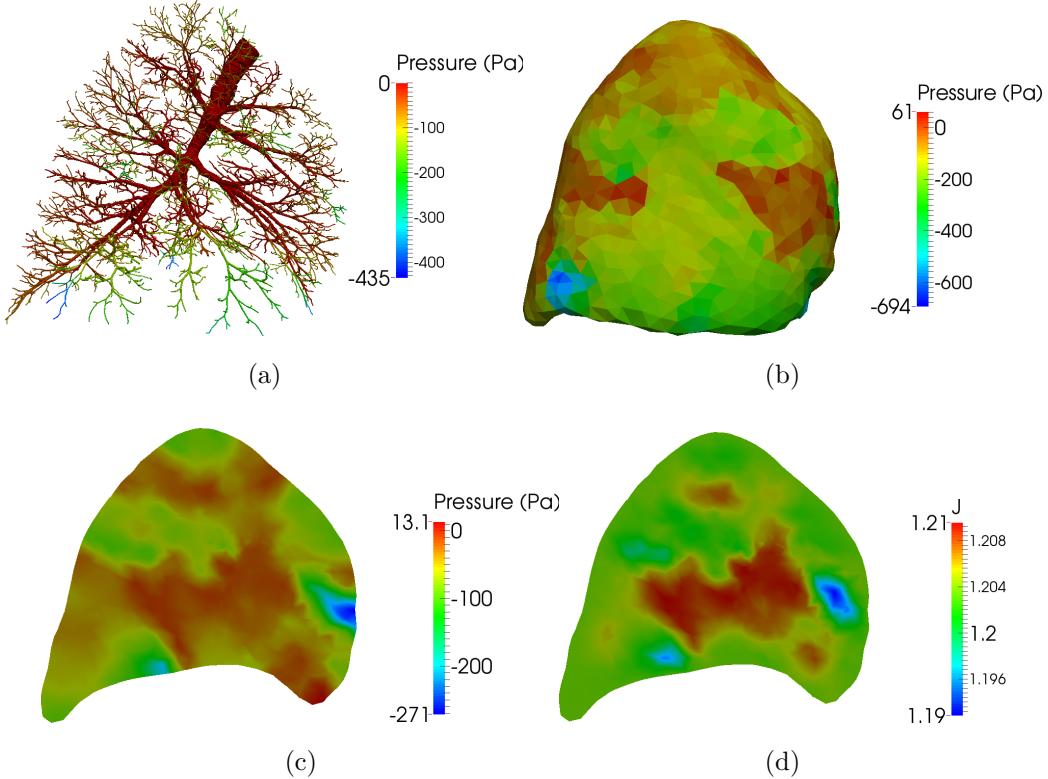


Figure 7.11: Pressure in the airway tree (a). Sagital slices showing pressure in the tissue using a linear interpolation (c), and tissue expansion from FRC (d). Pressure on the lung surface using piece wise constant finite elements (b).

#### 7.4.2 Breathing with airway constriction

In addition to the previously applied boundary conditions and simulation parameters, we now alter the simulation setup to simulate localize constriction of the airways. We implement this by reducing the radius of the lower airways (with radius less than 4mm) within a ball near the right middle lobe, represented by the red ball in Figure 7.6b. We reduce the radius of the afore-

mentioned lower airways by 0%, 40%, 50%, 60% and 65%. This corresponds to a mean pathway resistance within the ball of 0.0501, 0.101, 0.188, 0.399 and 1.889 Pa mm<sup>-3</sup>s, respectively. We will refer to this region as the structurally modified region. Figure 7.12 shows the changes in variables of physiological interest within the ball as the pathway resistance increases. The amount of tissue expansion during inspiration decreases as the airways become constricted (airway radius decreases and pathway resistance increases), shown in Figure 7.12a. This is due to the reduced amount of flow in these airways. Also the standard deviation increases because the pathway resistance of each branch increases by a different amount, depending on its original length and radius. Long and narrow branches will be affected most by the constriction. In Figure 7.12b the pressure decreases with increasing pathway resistance, since a larger pressure gradient is needed to force the fluid down the constricted branches. Figure 7.12c shows the elastic stress in the tissue decrease as pathway resistance increases. This is due to the decrease in tissue deformation, highlighted in Figure 7.12a.

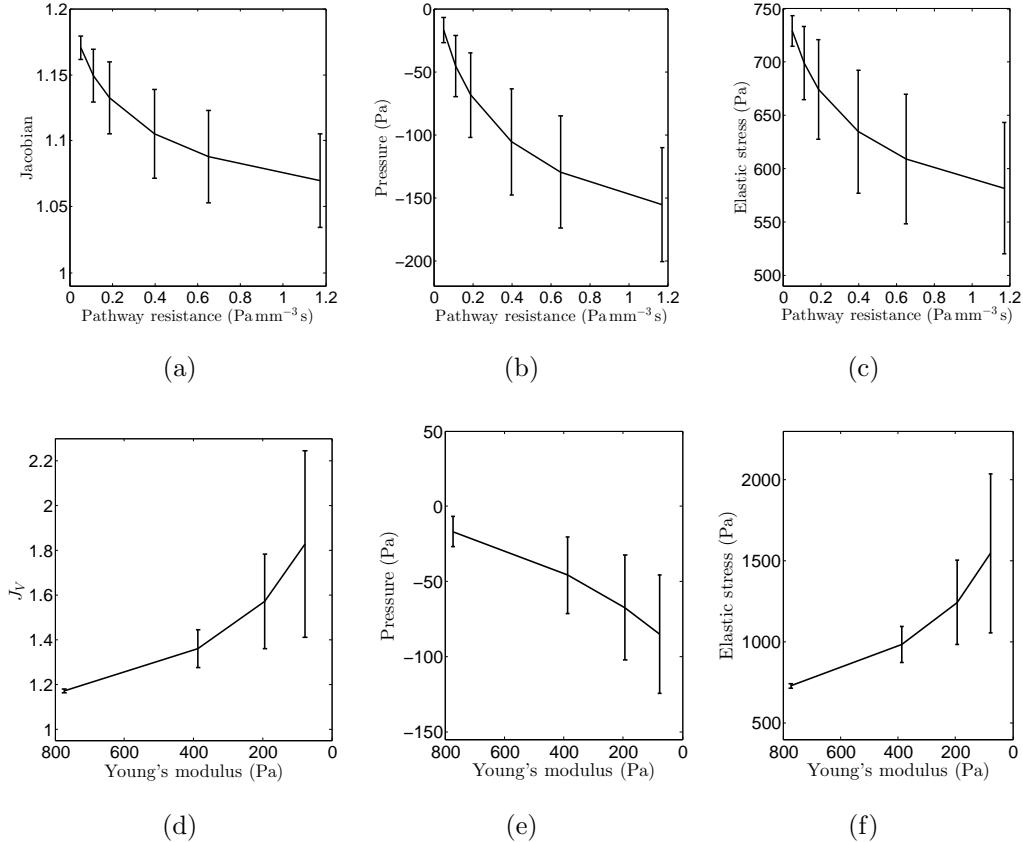


Figure 7.12: Mean and standard deviations of the relative Jacobian from FRC (a), pressure in the tissue (b) and elastic stress (c) are plotted against increasing pathway resistance and Young's modulus (d,e,f) within the structurally modified region.

The simulations results shown in Figure 7.13 were performed with 65% airway constriction in the lower airways, applied within the structurally modified region. The volume conserving property of the method is illustrated in Figure 7.13b where the tissue surrounding the constricted area is expanding to compensate for the reduction of tissue expansion due to the constriction

within the structurally modified region. Figure 7.13a shows an increase in pressure near the boundary of this region. This facilitates a pressure gradient that allows for fluid to flow into the constricted region, as is shown in Figure 7.13d. As previously mentioned, the elastic stress is lower in the constricted region. The largest elastic stress appears around the constricted region, where the tissue is made to expand more.

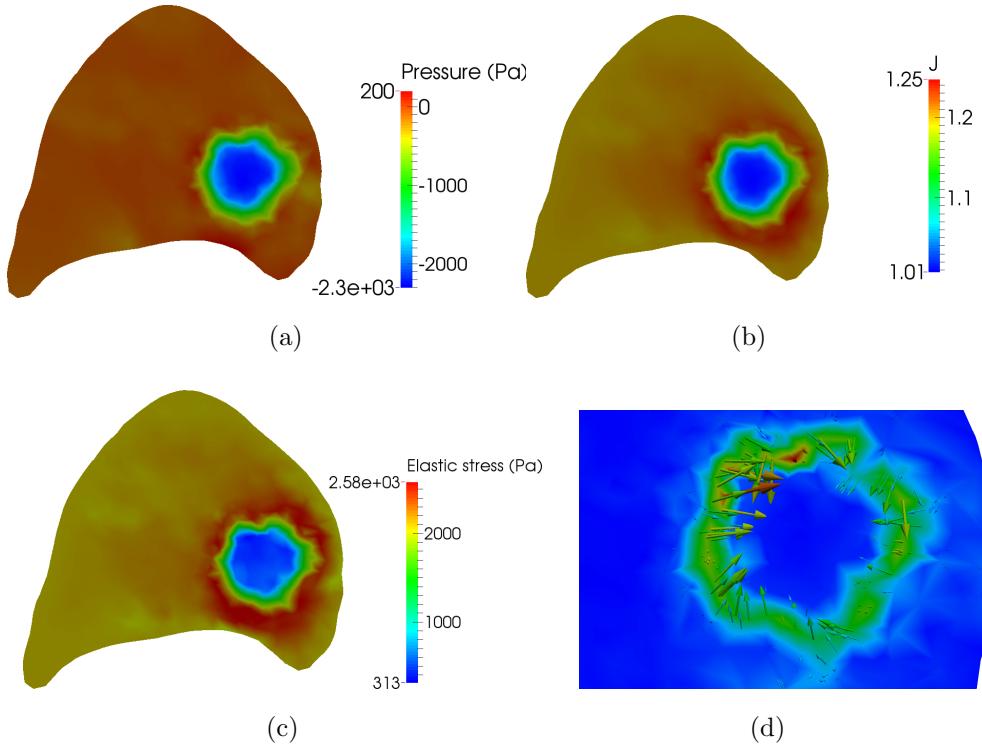


Figure 7.13: Sagittal slices showing the pressure (a), relative Jacobian (b), elastic stress (c) and fluid flux around the constricted region (d). The arrows in (d) indicate the direction of the flow and are scaled according to the magnitude of the flow.

### 7.4.3 Breathing with tissue weakening

We now alter the simulation setup to simulate localized weakening of the tissue. We implement this by reducing the Young's modulus of the tissue within the structurally modified region represented by a red ball in Figure 7.6b. We reduce the Young's modulus by 0%, 50%, 75% and 90%. This corresponds to a modified Young's modulus of 823, 411.5, 205.75 and 82.3 Pa, respectively. The amount of tissue expansion during inspiration increases as the tissue becomes weaker, shown in Figure 7.12d. This results in more fluid flowing along the branches into the weakened region, thus causing an increase in the magnitude of pressures in the tissue to drive the flow, shown in Figure 7.12e. The reduction in the Young's modulus within the modified region causes the calculated elastic stress to decrease, since the marginal increase in strain is outweighed by the reduction in the strain energy, see Figure 7.12f. Also, due to the large amount of tissue expansion within the structurally modified region, the tissue immediately surrounding this region expands the least. This can be seen in Figure 7.14, with a 90% reduction in the Young's modulus. This is again due to mass conservation enforced by equation (7.15c).

## 7.5 Discussion

We have presented a mathematical model of the lung that tightly couples tissue deformation with ventilation using a poroelastic model coupled to a

fluid network model. We have highlighted the assumptions necessary to arrive at such a model, and outlined its limitations. In comparison with previous ventilation models, the proposed model is volume conserving by modelling the tissue as a continuum, able to model collateral ventilation, and driven by deformation boundary conditions extracted from imaging data. In simulations of normal breathing, the model is able to produce physiologically realistic global measurements and dynamics. In simulations with altered airway resistance and tissue stiffness, the model illustrates the importance of a fully coupled model and the interdependence of the tissue and airway mechanics.

### **7.5.1 Contributors of airway resistance and tissue mechanics to lung function**

We have found that there is correlation between airway resistance and ventilation, see Figure 7.10a. Also due to the airway resistance, hysteresis effects appear during breathing (Figure 7.8) and result in complex ventilation distribution, caused by delayed filling and emptying of the tissue. Due to the Poisuelle law that governs the flow through the airways, small changes in airway radius can result in large changes in pathway resistance, which in turn can significantly affect the results of the coupled model. Thus, parametrizing the airways correctly is very important. However this is notoriously difficult since CT data is only available down to the 5-6th generation, and small er-

rors and biases in the segmentation, that get propagated by the airway generation algorithm, can have large influences in determining the simulation results. The tissue mechanics play an equally important role in determining the function of the lung model. This has been demonstrated in section 7.4.3.

### 7.5.2 Limitations and future work

In order to move towards a more realistic model of the lung breathing, many steps need to be taken. We will list the main limitations that exist in the airway tree, the poroelastic model, the boundary conditions and the geometry, and give indications on how these could be addressed in a future model.

**Airway tree limitations:**

1. The airway tree flow model currently implemented is the simplest possible flow model and makes the Poisuelle flow assumption for the whole tree. This could be improved by using a more sophisticated flow model for the airways, such as the 3D-0D model presented in Ismail et al. [2013].
2. The coupling of each terminal branch to the tissue currently assumes that there is no added resistance to air flowing from the terminal branch to each alveolar unit within the tissue. This could be

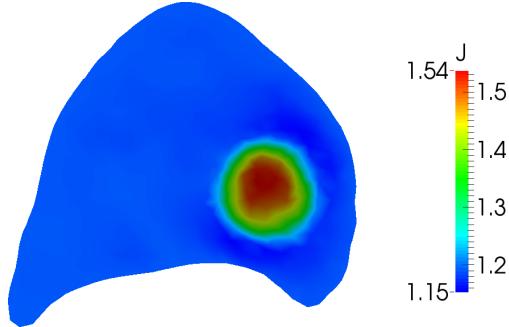


Figure 7.14: Slice showing the amount of tissue expansion ( $J_V$ ) from FRC during inspiration with localised tissue weakening.

improved by adding a resistive model, for example a model considering the volume of tissue that the terminal branch is feeding. This would also slightly increase the mean pressure drop of the lung model.

3. At the moment the airway tree is assumed to be static, and its configuration is not influenced by the deformation and stresses in the tissue. This could be improved by modelling the interaction of stresses and strains on the airway wall, opening up the airways during inspiration.

**Poroelastic tissue limitations:** 1. We have assumed a Neo-Hookean law for strain energy law to make the interpretation of the elasticity constants and dynamics of the model as simple as possible. However lung parenchyma is known to follow an exponential stress-strain relation, especially past tidal volume, where a law such as the one proposed by Fung [1975] might be more appropriate. Also little is known about the form of the strain-energy law during disease (e.g. fibrosis or emphysema). Similarly, for the permeability law little is known about its form for healthy or diseased tissue. Further experiments and modelling investigation would be needed to develop these.

2. Currently the tissue has been parameterized homogeneously. Density information from CT images could be used to parameterize the initial phase, and diffusion tensor MRI be used to parameterize the permeability.

**Boundary condition limitations:** The rigid registration should be updated to a more sophisticated non-linear registration algorithm (e.g. Yin et al. [2013], Heinrich et al. [2013]) that is able to account for the complicated deformation of the lung surface during breathing.

**Geometry limitations:** 1. To model the complete organ and give a more accurate pressure drop, both the right and left lung, and the trachea should be included. 2. The airway tree generated in this work goes down to generations 13-14. More generations could be added to result in a fuller tree. However this would also require a finer mesh to approximate the lung volume, to resolve the coupling between each terminal branch and a subregion of lung tissue. 3. Cavities in the lung parenchyma due to large airways are currently not accounted for, i.e. it is assumed that the volume occupied by the airways is zero. To improve on this, a mesh of the lung with the larger upper airways removed would need to be generated. This new mesh could also incorporate a model of the cartilage found in the upper airways. 4. Additional no-flux boundaries should be introduced to represent the well defined and thought to be impermeable boundaries, between fissures and lung segments.

### 7.5.3 Summary

The model presented in this paper appears to be a valid tool for solving the mechanical problem of tightly coupling lung deformation and ventilation during normal breathing and breathing with disease. The numerical simulations are shown to be able to reproduce global physiological realistic measurements. A fully nonlinear formulation permits the inclusion of various constitutive models, allowing investigation into different diseased states during various breathing conditions. A finite element method has been used to discretise the equations in a monolithic way to ensure convergence of the

nonlinear problem, even under strong fluid-network-poroelastic-medium coupling conditions. Due to the flexibility of the model, further improvements in its physiological accuracy are possible.

## 7.6 Appendix

### 7.6.1 Finite element matrices

For the fully-coupled large deformation poroelastic fluid network model we need to solve the linear system  $\mathbf{K}(\mathbf{u}_i^n)\xi\mathbf{u}_{i+1}^n = -\mathbf{R}(\mathbf{u}_i^n, \mathbf{u}^{n-1})$  at each Newton iteration. This can be expanded as

$$\begin{bmatrix} \mathbf{K}^e & 0 & \mathbf{B}^T & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{M} & \mathbf{B}^T & \mathbf{L}^T & 0 & 0 & 0 & 0 \\ -\mathbf{B} & -\Delta t\mathbf{B} & \mathbf{J} & 0 & 0 & 0 & 0 & -\Delta t\mathbf{G}^T \\ 0 & \mathbf{L} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{T}_{11} & \cdots & \cdots & \mathbf{T}_{14} \\ 0 & 0 & 0 & 0 & \vdots & & & \vdots \\ 0 & 0 & 0 & 0 & \mathbf{T}_{31} & \cdots & \cdots & \mathbf{T}_{34} \\ 0 & 0 & \mathbf{G} & 0 & 0 & -\mathbf{X} & 0 & 0 \end{bmatrix} \begin{bmatrix} \xi\mathbf{u}^n \\ \xi\mathbf{z}^n \\ \xi\mathbf{p}^n \\ \xi\lambda^n \\ \xi\mathbf{P}^n \\ \vdots \\ \xi\mathbf{P}_d^n \\ \xi\mathbf{Q}^n \\ \xi\mathbf{Q}_d^n \end{bmatrix} = - \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 - \Delta t\mathbf{G}^T\mathbf{Q}_d^n \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \mathbf{G}\mathbf{p}^n - \mathbf{X}\mathbf{P}_d^n \end{bmatrix},$$

where we have defined the following matrices:

$$\mathbf{K}^e = [\mathbf{a}_{kl}], \quad \mathbf{k}_{kl}^e = \int_{\Omega_t} \mathbf{B}_k^T \mathbf{D}(\mathbf{u}_i^n) \mathbf{B}_l + (\nabla \phi_k)^T \boldsymbol{\sigma}_e(\mathbf{u}_i^n) \nabla \phi_l \, dv,$$

$$\boldsymbol{M} = [\boldsymbol{m}_{kl}], \quad \boldsymbol{m}_{kl} = \int_{\Omega_t} \kappa^{-1}(\boldsymbol{u}_i^n) \boldsymbol{\phi}_k \cdot \boldsymbol{\phi}_l \, dv,$$

$$\boldsymbol{B} = [\boldsymbol{b}_{kl}], \quad \boldsymbol{b}_{kl} = - \int_{\Omega_t} \psi_k \nabla \cdot \boldsymbol{\phi}_l \, dv,$$

$$\boldsymbol{J} = [\boldsymbol{j}_{kl}], \quad \boldsymbol{j}_{kl} = \delta \sum_K \int_{\partial k \setminus \partial \Omega_t} h_{\partial K}[\psi_k][\psi_k] \, ds.$$

$$\boldsymbol{r}_1 = [\boldsymbol{r}_{1i}], \quad \boldsymbol{r}_{1i} = \int_{\Omega_t} (\boldsymbol{\sigma}_e(\boldsymbol{u}_i^n) - p_i^n \boldsymbol{I}) : \nabla \boldsymbol{\phi}_i - \rho(\boldsymbol{u}_i^n) \boldsymbol{\phi}_i \cdot \boldsymbol{f} \, dv - \int_{\Gamma_t} \boldsymbol{\phi}_i \cdot \boldsymbol{t}_N \, ds,$$

$$\boldsymbol{r}_2 = [\boldsymbol{r}_{2i}], \quad \boldsymbol{r}_{2i} = \int_{\Omega_t} \kappa^{-1}(\boldsymbol{u}_i^n) \boldsymbol{\phi}_i \cdot \boldsymbol{z}_i^n - p_i^n \nabla \cdot \boldsymbol{\phi}_i - \rho^f(\boldsymbol{u}_i^n) \boldsymbol{\phi}_i \cdot \boldsymbol{f} \, dv,$$

$$\boldsymbol{r}_3 = [\boldsymbol{r}_{3i}], \quad \boldsymbol{r}_{3i} = \int_{\Omega_t} \psi_i \nabla \cdot (\boldsymbol{u}_i^n - \boldsymbol{u}^{n-1}) + \Delta t \psi_i \nabla \cdot \boldsymbol{z}_i^n - \Delta t \psi_i g \, dv + \delta \sum_K \int_{\partial k \setminus \partial \Omega_t} h_{\partial K}[\psi_i] [p_i^n - p^{n-1}] \, ds$$

$$\mathbf{L} = [\mathbf{l}_{ij}], \quad \mathbf{l}_{ij} = \int_{\Omega} \epsilon_i \boldsymbol{\phi}_j \cdot \mathbf{n},$$

$$\mathbf{X} = [\mathbf{x}_{ij}], \quad \mathbf{x}_{ij} := \begin{cases} 1 & \text{if } ||Q_{di} - \text{cent}(E_j)|| < ||Q_{dk} - \text{cent}(E_j)||, \ k = 1, 2, \dots, N, \ k \neq i, \\ 0 & \text{otherwise,} \end{cases}$$

$$\mathbf{G} = [\mathbf{g}_{ij}], \quad \mathbf{g}_{ij} = \int_{\Omega} \mathbf{x}_{ij} \frac{\boldsymbol{\phi}_j}{|E_j|},$$

$\mathbf{T}$  represents the matrix entries required for the fluid network.

Here  $\boldsymbol{\phi}_k$  are vector valued linear basis functions such that the displacement vector at the  $i$ th iteration can be written as  $\boldsymbol{u}_i^n = \sum_{k=1}^{n_u} \boldsymbol{u}_{i,k}^n \boldsymbol{\phi}_k$ , with  $\sum_{k=1}^{n_u} \boldsymbol{u}_{i,k}^n \boldsymbol{\phi}_k \in \mathbf{W}_h^E$ . Similarly for the fluid flux vector we have  $\boldsymbol{z}_i^n = \sum_{k=1}^{n_z} \boldsymbol{z}_{i,k}^n \boldsymbol{\phi}_k$ , with  $\sum_{k=1}^{n_z} \boldsymbol{z}_{i,k}^n \boldsymbol{\phi}_k \in \mathbf{W}_h^D$ . The scalar valued constant basis functions  $\psi_i$  are used to approximate the pressure, such that  $\boldsymbol{p}_i^n = \sum_{k=1}^{n_p} p_{i,k}^n \psi_k$ ,

with  $\sum_{k=1}^{n_p} p_{i,k}^n \psi_k \in Q_h$ . Similarly,  $\boldsymbol{\epsilon}_k$  are scalar valued linear basis functions such that the Lagrangian multiplier vector at the  $i$ th iteration can be written as  $\boldsymbol{\lambda}_i^n = \sum_{k=1}^{n_\lambda} \boldsymbol{\lambda}_{i,k}^n \boldsymbol{\epsilon}_k$ , and  $cent(E_j)$  denotes the centroid of the  $j$ th element. Also to aid the assembly of the fourth order tensor we have adopted the matrix voigt notation. In particular  $\boldsymbol{D}$  is the matrix form of  $\mathbf{c}$ , and  $\boldsymbol{B}_k$  is the matrix version of  $\nabla^S \boldsymbol{\phi}_k$ , see (A.3) and (A.4) for details.

### 7.6.2 Matrix voigt notation

The spatial tangent modulus, fourth-order tensor, can be written as (see [Bonet and Wood, 1997, section 5.3.2] and [Holzapfel et al., 2000, section 6.6] )

$$\mathbf{c}_{ijkl} = \frac{1}{J} F_{iI} F_{jJ} F_{kK} F_{lL} \mathbb{C}_{IJKL}, \quad (7.17)$$

where  $\mathbb{C}$  is the associated tangent modulus tensor in the reference configuration, given by (in component form)

$$\mathbb{C}_{IJKL} = \frac{4\partial^2 W}{\partial C_{IJ}\partial C_{KL}} + pJ \frac{\partial C_{IJ}^{-1}}{\partial C_{KL}}. \quad (7.18)$$

To ease the implementation of the spatial tangent modulus we make use of matrix voigt notation. The matrix form of  $\mathbf{c}$  is given by  $\boldsymbol{D}$ , which can be

written as (see [Bonet and Wood, 1997, section 7.4.2])

$$\mathbf{D} = \frac{1}{2} \begin{pmatrix} 2\mathfrak{c}_{1111} & 2\mathfrak{c}_{1122} & 2\mathfrak{c}_{1133} & \mathfrak{c}_{1112} + \mathfrak{c}_{1121} & \mathfrak{c}_{1113} + \mathfrak{c}_{1131} & \mathfrak{c}_{1123} + \mathfrak{c}_{1132} \\ & 2\mathfrak{c}_{2222} & 2\mathfrak{c}_{2233} & \mathfrak{c}_{2212} + \mathfrak{c}_{2221} & \mathfrak{c}_{2213} + \mathfrak{c}_{2231} & \mathfrak{c}_{2223} + \mathfrak{c}_{2232} \\ & & 2\mathfrak{c}_{3333} & \mathfrak{c}_{3312} + \mathfrak{c}_{3321} & \mathfrak{c}_{3313} + \mathfrak{c}_{3331} & \mathfrak{c}_{3323} + \mathfrak{c}_{3332} \\ & & & \mathfrak{c}_{1212} + \mathfrak{c}_{1221} & \mathfrak{c}_{1213} + \mathfrak{c}_{1231} & \mathfrak{c}_{1223} + \mathfrak{c}_{1232} \\ & \text{sym.} & & & \mathfrak{c}_{1313} + \mathfrak{c}_{1331} & \mathfrak{c}_{1323} + \mathfrak{c}_{1332} \\ & & & & & \mathfrak{c}_{2323} + \mathfrak{c}_{2332} \end{pmatrix}. \quad (7.19)$$

We also make use of the following implementation friendly notation

$$\nabla^S \phi_k = \begin{bmatrix} \phi_{k,1} & 0 & 0 \\ 0 & \phi_{k,2} & 0 \\ 0 & 0 & \phi_{k,3} \\ \phi_{k,2} & \phi_{k,1} & 0 \\ 0 & \phi_{k,3} & \phi_{k,2} \\ \phi_{k,3} & 0 & \phi_{k,1} \end{bmatrix} = \mathbf{B}_k. \quad (7.20)$$

### 7.6.3 Neo-Hookean example

For the numerical examples we have used the following Neo-Hookean strain-energy law

$$W(\mathbf{C}) = \frac{\mu}{2}(I_1 - 3) + \frac{\Lambda}{4}(I_3 - 1) - (\mu + \frac{\Lambda}{2})\ln I_3^{\frac{1}{2}}. \quad (7.21)$$

Thus, the resulting effective stress tensor is given by

$$\boldsymbol{\sigma}_e = \frac{\Lambda}{2J}(J^2 - 1)\mathbf{I} + \frac{\mu}{J}(\mathbf{C}^T - \mathbf{I}), \quad (7.22)$$

and the spatial tangent modulus tensor is given as

$$\mathbf{c} = \mathbf{c}_e + p(\mathbf{I} \otimes \mathbf{I} - 2\mathcal{Z}) \quad (7.23)$$

where

$$\mathbf{c}_e = \Lambda J^2 \mathbf{I} \otimes \mathbf{I} + [2\mu - \Lambda(J^2 - 1)] \mathcal{B}, \quad (7.24)$$

and

$$\mathcal{B}_{ijkl} = \frac{1}{2}(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}), \quad \mathcal{Z}_{ijkl} = \delta_{ik}\delta_{jl}, \quad \mathbf{I} \otimes \mathbf{I} = \delta_{ij}\delta_{kl}. \quad (7.25)$$

See [Bonet and Wood, 1997, chapter 5] and [Wriggers, 2008, chapter 3] for details.

# Chapter 8

## Conclusion

# Appendix A

## A.1 Tangent modulus

The spatial tangent modulus, fourth-order tensor, can be written as (see [Bonet and Wood, 1997, section 5.3.2] and [Holzapfel et al., 2000, section 6.6] )

$$\mathbf{c}_{ijkl} = \frac{1}{J} F_{iI} F_{jJ} F_{kK} F_{lL} \mathbb{C}_{IJKL}, \quad (\text{A.1})$$

where  $\mathbb{C}$  is the associated tangent modulus tensor in the reference configuration, given by (in component form)

$$\mathbb{C}_{IJKL} = \frac{4\partial^2 W}{\partial C_{IJ}\partial C_{KL}} + pJ \frac{\partial C_{IJ}^{-1}}{\partial C_{KL}}. \quad (\text{A.2})$$

To ease the implementation of the spatial tangent modulus we make use of matrix voigt notation. The matrix form of  $\mathbf{c}$  is given by  $\mathbf{D}$ , which can be

written as (see [Bonet and Wood, 1997, section 7.4.2])

$$\mathbf{D} = \frac{1}{2} \begin{pmatrix} 2\mathfrak{c}_{1111} & 2\mathfrak{c}_{1122} & 2\mathfrak{c}_{1133} & \mathfrak{c}_{1112} + \mathfrak{c}_{1121} & \mathfrak{c}_{1113} + \mathfrak{c}_{1131} & \mathfrak{c}_{1123} + \mathfrak{c}_{1132} \\ & 2\mathfrak{c}_{2222} & 2\mathfrak{c}_{2233} & \mathfrak{c}_{2212} + \mathfrak{c}_{2221} & \mathfrak{c}_{2213} + \mathfrak{c}_{2231} & \mathfrak{c}_{2223} + \mathfrak{c}_{2232} \\ & & 2\mathfrak{c}_{3333} & \mathfrak{c}_{3312} + \mathfrak{c}_{3321} & \mathfrak{c}_{3313} + \mathfrak{c}_{3331} & \mathfrak{c}_{3323} + \mathfrak{c}_{3332} \\ & & & \mathfrak{c}_{1212} + \mathfrak{c}_{1221} & \mathfrak{c}_{1213} + \mathfrak{c}_{1231} & \mathfrak{c}_{1223} + \mathfrak{c}_{1232} \\ & \text{sym.} & & & \mathfrak{c}_{1313} + \mathfrak{c}_{1331} & \mathfrak{c}_{1323} + \mathfrak{c}_{1332} \\ & & & & & \mathfrak{c}_{2323} + \mathfrak{c}_{2332} \end{pmatrix}. \quad (\text{A.3})$$

## A.2 Matrix voigt notation

We also make use of the following implementation friendly notation

$$\nabla^S \boldsymbol{\phi}_k = \begin{bmatrix} \phi_{k,1} & 0 & 0 \\ 0 & \phi_{k,2} & 0 \\ 0 & 0 & \phi_{k,3} \\ \phi_{k,2} & \phi_{k,1} & 0 \\ 0 & \phi_{k,3} & \phi_{k,2} \\ \phi_{k,3} & 0 & \phi_{k,1} \end{bmatrix} = \mathbf{B}_k. \quad (\text{A.4})$$

### A.2.1 Neo-Hookean example

For the numerical examples shown in section 6.2 we have used the following Neo-Hookean strain-energy law

$$W(\mathbf{C}) = \frac{\mu}{2}(\text{tr}(\mathbf{C}) - 3) + \frac{\Lambda}{4}(J^2 - 1) - (\mu + \frac{\Lambda}{2})\ln J. \quad (\text{A.5})$$

Thus, the resulting effective stress tensor is given by

$$\boldsymbol{\sigma}_e = \frac{\Lambda}{2J}(J^2 - 1)\mathbf{I} + \frac{\mu}{J}(\mathbf{C}^T - \mathbf{I}), \quad (\text{A.6})$$

and the spatial tangent modulus tensor is given as

$$\mathfrak{c} = \mathfrak{c}_e + p(\mathbf{I} \otimes \mathbf{I} - 2\mathcal{Z}) \quad (\text{A.7})$$

where

$$\mathfrak{c}_e = \Lambda J^2 \mathbf{I} \otimes \mathbf{I} + [2\mu - \Lambda(J^2 - 1)] \mathcal{E}, \quad (\text{A.8})$$

and

$$\mathcal{E}_{ijkl} = \frac{1}{2}(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}), \quad \mathcal{Z}_{ijkl} = \delta_{ik}\delta_{jl}, \quad \mathbf{I} \otimes \mathbf{I} = \delta_{ij}\delta_{kl}. \quad (\text{A.9})$$

See [Bonet and Wood, 1997, chapter 5] and [Wriggers, 2008, chapter 3] for details.

# Bibliography

Porous elasticity: Lectures on the elasticity of porous materials as an application of the theory of mixtures.

Guillermo M Albaiceta, Lluis Blanch, and Umberto Lucangelo. Static pressure–volume curves of the respiratory system: were they just a passing fad? *Current opinion in critical care*, 14(1):80–86, 2008.

E.S. Almeida and R.L. Spilker. Finite element formulations for hyperelastic transversely isotropic biphasic soft tissues. *Computer Methods in Applied Mechanics and Engineering*, 151(3):513–538, 1998.

Patrick R Amestoy, Iain S Duff, and J L’Excellent. Multifrontal parallel distributed symmetric and unsymmetric solvers. *Computer methods in applied mechanics and engineering*, 184(2):501–520, 2000.

C.G Armstrong, W.M Lai, and V.C Mow. An analysis of the unconfined compression of articular cartilage. *Journal of Biomechanical Engineering*, 106:73–165, 1984.

Christopher Arthurs. *Effcient Simulation of Cardiac Electrical Propagation using Adaptive High-Order Finite Elements*. PhD thesis, University of Oxford, 2012.

Lyia Asner. *Effcient numerical methods for the solution of coupled multi-physics problems*. PhD thesis, University of Oxford, 2013.

Gerard A Ateshian, Steve Maas, and Jeffrey A Weiss. Finite element algorithm for frictionless contact of porous permeable media under finite deformation and sliding. *Journal of biomechanical engineering*, 132(6):061006, 2010.

Ivo Babuška. Error-bounds for finite element method. *Numerische Matematik*, 16(4):322–333, 1971.

S. Badia, A. Quaini, and A. Quarteroni. Coupling Biot and Navier-Stokes equations for modelling fluid-poroelastic media interaction. *Journal of Computational Physics*, 228(21):7986–8014, 2009.

Hélène Barucq, Monique Madaune-Tort, and Patrick Saint-Macary. Some existence-uniqueness results for a class of one-dimensional nonlinear Biot models. *Nonlinear Analysis: Theory, Methods & Applications*, 61(4):591–612, 2005.

J.H.T. Bates. *Lung mechanics: an inverse modeling approach*. Cambridge University Press, 2009.

Klaus-Jürgen Bathe, Ekkehard Ramm, and Edward L Wilson. Finite element formulations for large deformation dynamic analysis. *International Journal for Numerical Methods in Engineering*, 9(2):353–386, 1975.

Michele Benzi, Gene H Golub, and Jorg Liesen. Numerical solution of saddle point problems. *Acta numerica*, 14(1):1–137, 2005.

Lorenz Berger, Rafel Bordas, David Kay, and Simon Tavener. Stabilized low-order finite element approximation for linear three-field poroelasticity. *IMA Journal of Numerical Analysis (submitted)*, 2014a.

Lorenz Berger, Rafel Bordas, David Kay, and Simon Tavener. A stabilized nonlinear finite element method for three-field incompressible poroelasticity valid in large deformations. *IMA Journal of Numerical Analysis (submitted)*, 2014b.

Lorenz Berger, Rafel Bordas, David Kay, and Simon Tavener. Stabilized low-order finite element approximation for linear three-field poroelasticity. *IMA Journal of Numerical Analysis (submitted)*, 2014c.

Miguel Bernabeu. *An open source HPC-enabled model of cardiac defibrillation of the human heart*. PhD thesis, University of Oxford, 2011.

P.B. Bochev and C.R. Dohrmann. A computational study of stabilized, low-order C0 finite element approximations of Darcy equations. *Computational Mechanics*, 38(4):323–333, 2006.

Reint Boer. *Trends in continuum mechanics of porous media*, volume 18. Springer, 2005.

J. Bonet and R.D. Wood. *Nonlinear continuum mechanics for finite element analysis*. Cambridge University Press, 1997.

R.M. Bowen. Incompressible porous media models by use of the theory of mixtures. *International Journal of Engineering Science*, 18(9):1129–1148, 1980.

Susanne C Brenner and Larkin Ridgway Scott. *The mathematical theory of finite element methods*, volume 15. Springer, 2008.

Erik Burman and Peter Hansbo. A unified stabilized method for Stokes' and Darcy's equations. *Journal of Computational and Applied Mathematics*, 198(1):35–51, 2007.

D. Chapelle, J.F. Gerbeau, J. Sainte-Marie, and I.E Vignon-Clementel. A poroelastic model valid in large strains with applications to perfusion in cardiac modeling. *Computational Mechanics*, 46(1):91–101, 2010.

A.N. Cookson, J. Lee, C. Michler, R. Chabiniok, E. Hyde, D.A. Nordsletten, M. Sinclair, M. Siebes, and N.P. Smith. A novel porous mechanical framework for modelling the interaction between coronary perfusion and myocardial mechanics. *Journal of Biomechanics*, 45(5):850 – 855, 2012.

O. Coussy. *Poromechanics*. John Wiley & Sons Inc, 2004.

- O. Coussy. *Mechanics and physics of porous solids*. Wiley, 2010.
- O. Coussy, L. Dormieux, and E. Detournay. From mixture theory to biots approach for porous media. *International journal of solids and structures*, 35(34):4619–4635, 1998.
- R De Wilde, J Clement, JM Hellemans, Marc Decramer, Maurice Demedts, R Boving, and KP Van de Woestijne. Model of elasticity of the human lung. *J. appl. Physiol*, 51(2):254–261, 1981.
- Daniele Antonio Di Pietro and Alexandre Ern. *Mathematical aspects of discontinuous Galerkin methods*, volume 69. Springer, 2011.
- H.C. Elman, D.J. Silvester, and A.J. Wathen. *Finite elements and fast iterative solvers: with applications in incompressible fluid dynamics*. Oxford University Press, USA, 2005.
- Qianqian Fang and David A Boas. Tetrahedral mesh generation from volumetric binary and grayscale images. In *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on*, pages 1142–1145. IEEE, 2009.
- Xiaobing Feng and Yinnian He. Fully discrete finite element approximations of a polymer gel model. *SIAM Journal on Numerical Analysis*, 48(6):2186–2217, 2010.
- Massimiliano Ferronato, Nicola Castelletto, and Giuseppe Gambolati. A fully

coupled 3-D mixed finite element model of Biot consolidation. *Journal of Computational Physics*, 229(12):4813–4830, 2010.

Y.C. Fung. Stress, deformation, and atelectasis of the lung. *Circulation Research*, 37(4):481–496, 1975.

Craig J Galbán, Meilan K Han, Jennifer L Boes, Komal A Chughtai, Charles R Meyer, Timothy D Johnson, Stefanie Galbán, Alnawaz Rehemtulla, Ella A Kazerooni, Fernando J Martinez, et al. Computed tomography-based biomarker provides unique signature for diagnosis of copd phenotypes and disease progression. *Nature medicine*, 18(11):1711–1715, 2012.

F. Galbusera, H. Schmidt, J. Noailly, A. Malandrino, D. Lacroix, H.J. Wilke, and A. Shirazi-Adl. Comparison of four methods to simulate swelling in poroelastic finite element models of intervertebral discs. *Journal of the Mechanical Behavior of Biomedical Materials*, 4(7):1234 – 1241, 2011.

P.A. Galie, R.L. Spilker, and J.P. Stegemann. A linear, biphasic model incorporating a Brinkman term to describe the mechanics of cell-seeded collagen hydrogels. *Annals of Biomedical Engineering*, 39:2767–2779, 2011.

O. Gonzalez and A.M. Stuart. *A first course in continuum mechanics*, volume 42. Cambridge University Press, 2008.

J.B Haga, H Osnes, and H.P 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.

R Scott Harris. Pressure-volume curves of the respiratory system. *Respiratory care*, 50(1):78–99, 2005.

HP Heinrich, Mark Jenkinson, Michael Brady, and Julia A Schnabel. Mrf-based deformable registration and ventilation estimation of lung ct. *Medical Imaging, IEEE Transactions on*, 32(7):1239–1248, 2013.

M.H Holmes and V.C Mow. The nonlinear characteristics of soft gels and hydrated connective tissues in ultrafiltration. *Journal of Biomechanics*, 23(11):1145–1156, 1990.

G.A Holzapfel, C.A. Schulze-Bauer, and M. Stadler. Mechanics of angioplasty: Wall, balloon and stent. *Asme Applied Mechanics*, 242:141–156, 2000.

M Ismail, A Comerford, and WA Wall. Coupled and reduced dimensional modeling of respiratory mechanics during spontaneous breathing. *International journal for numerical methods in biomedical engineering*, 29(11):1285–1305, 2013.

A-RA Khaled and K Vafai. The role of porous media in modeling flow and heat transfer in biological tissues. *International Journal of Heat and Mass Transfer*, 46(26):4989–5003, 2003.

- J Kim, HA Tchelepi, and R Juanes. Stability and convergence of sequential methods for coupled flow and geomechanics: Fixed-stress and fixed-strain splits. *Computer Methods in Applied Mechanics and Engineering*, 200(13):1591–1606, 2011.
- B. S. Kirk, J. W. Peterson, R. H. Stogner, and G. F. Carey. libMesh: A C++ Library for Parallel Adaptive Mesh Refinement/Coarsening Simulations. *Engineering with Computers*, 22(3–4):237–254, 2006.
- P. Kowalczyk. Mechanical model of lung parenchyma as a two-phase porous medium. *Transport in Porous Media*, 11(3):281–295, 1993.
- P. Kowalczyk and M. Kleiber. Modelling and numerical analysis of stresses and strains in the human lung including tissue-gas interaction. *European Journal of Mechanics. A. Solids*, 13(3):367–393, 1994.
- WM Lai and VC Mow. Drag-induced compression of articular cartilage during a permeation experiment. *Biorheology*, 17(1-2):111, 1980.
- B. Lande and W. Mitzner. Analysis of lung parenchyma as a parametric porous medium. *Journal of Applied Physiology*, 101(3):926–933, 2006.
- Lawrence Berkeley National Laboratory. LBNL LungLab Tour. [http://imglib.lbl.gov/cgi-bin/ImgLib/displaytag/LUNG\\_STRUCTURE](http://imglib.lbl.gov/cgi-bin/ImgLib/displaytag/LUNG_STRUCTURE), 1995.
- GC Lee, NT Tseng, and YM Yuan. Finite element modeling of lungs including interlobar fissures and the heart cavity. *Journal of biomechanics*, 16(9):679–690, 1983.

Kritsada Leungchavaphongse. *Mathematical modelling of the liver microcirculation*. PhD thesis, Imperial College London, 2013.

M.E. Levenston, E.H. Frank, and A.J. Grodzinsky. Variationally derived 3-field finite element formulations for quasistatic poroelastic analysis of hydrated biological tissues. *Computer Methods in Applied Mechanics and Engineering*, 156(1-4):231–246, 1998.

Huanrong Li and Yukun Li. A discontinuous Galerkin finite element method for swelling model of polymer gels. *Journal of Mathematical Analysis and Applications*, 398(1):11–25, 2012.

X.G. Li, H. Holst, J. Ho, and S. Kleiven. Three dimensional poroelastic simulation of brain edema: Initial studies on intracranial pressure. In *World Congress on Medical Physics and Biomedical Engineering*, pages 1478–1481. Springer, 2010.

Konstantin Lipnikov. *Numerical methods for the Biot model in poroelasticity*. PhD thesis, University of Houston, 2002.

R. Liu. *Discontinuous Galerkin finite element solution for poromechanics*. PhD thesis, The University of Texas at Austin, 2004.

S Maas, D Rawlins, J Weiss, and G Ateshian. FEBio theory manual. *University of Utah*, 2010.

Steve A Maas, Benjamin J Ellis, Gerard A Ateshian, and Jeffrey A Weiss.

FEBio: finite elements for biomechanics. *Journal of biomechanical engineering*, 134(1):1–10, 2012.

Arif Masud and Thomas JR Hughes. A stabilized mixed finite element method for darcy flow. *Computer Methods in Applied Mechanics and Engineering*, 191(39):4341–4370, 2002.

V.C. Mow, S.C. Kuei, W.M. Lai, and C.G. Armstrong. Biphasic creep and stress relaxation of articular cartilage in compression: Theory and experiments. *Journal of Biomechanical Engineering*, 102:73–84, 1980.

Márcio A Murad and Abimael FD Loula. On stability and convergence of finite element approximations of Biot’s consolidation problem. *International journal for numerical methods in engineering*, 37(4):645–667, 1994.

M.R. Owen and M.A. Lewis. The mechanics of lung tissue under high-frequency ventilation. *SIAM Journal on Applied Mathematics*, pages 1731–1761, 2001.

T.J. Pedley, R.C. Schroter, and M.F. Sudlow. The prediction of pressure drop and variation of resistance within the human bronchial airways. *Respiration Physiology*, 9(3):387–405, 1970.

Phillip Joseph Phillips and Mary F Wheeler. A coupling of mixed and continuous Galerkin finite element methods for poroelasticity i: the continuous in time case. *Computational Geosciences*, 11(2):131–144, 2007a.

Phillip Joseph Phillips and Mary F Wheeler. A coupling of mixed and continuous Galerkin finite element methods for poroelasticity ii: the discrete-in-time case. *Computational Geosciences*, 11(2):145–158, 2007b.

Phillip Joseph Phillips and Mary F Wheeler. A coupling of mixed and discontinuous Galerkin finite-element methods for poroelasticity. *Computational Geosciences*, 12(4):417–435, 2008.

Phillip Joseph Phillips and Mary F Wheeler. Overcoming the problem of locking in linear elasticity and poroelasticity: an heuristic approach. *Computational Geosciences*, 13(1):5–12, 2009.

Alfio Quarteroni and Alberto Valli. *Numerical approximation of partial differential equations*, volume 23. Springer, 2008.

V MARCO Ranieri, Rocco Giuliani, Tommaso Fiore, Michele Dambrosio, and Joseph Milic-Emili. Volume-pressure curve of the respiratory system predicts effects of peep in ards:” occlusion” versus” constant flow” technique. *American journal of respiratory and critical care medicine*, 149(1):19–27, 1994.

Pierre-Arnaud Raviart and Jean-Marie Thomas. A mixed finite element method for 2-nd order elliptic problems. In *Mathematical aspects of finite element methods*, pages 292–315. Springer, 1977.

Frank Rittner and Martin Döring. *Curves and loops in mechanical ventilation*. Dräger Medical Incorporated, 2005.

- R.E Showalter. Diffusion in poro-elastic media. *Journal of mathematical analysis and applications*, 251(1):310–340, 2000.
- B. Suki and J.H.T. Bates. Lung tissue mechanics as an emergent phenomenon. *Journal of Applied Physiology*, 110(4):1111–1118, 2011.
- A.J. Swan, A.R. Clark, and M.H. Tawhai. A computational model of the topographic distribution of ventilation in healthy human lungs. *Journal of Theoretical Biology*, 300(0):222 – 231, 2012.
- Ben Tal. Simplified models for gas exchange in the human lungs.
- Merryn H Tawhai, Martyn P Nash, Ching-Long Lin, and Eric A Hoffman. Supine and prone differences in regional lung density and pleural pressure gradients in the human lung with constant shape. *Journal of Applied Physiology*, 107(3):912–920, 2009.
- M.H. Tawhai and C.L. Lin. Image-based modeling of lung structure and function. *Journal of Magnetic Resonance Imaging*, 32(6):1421–1431, 2010.
- K. Ün and R.L. Spilker. A penetration-based finite element method for hyperelastic 3D biphasic tissues in contact. Part II: finite element simulations. *Journal of Biomechanical Engineering*, 128:934, 2006.
- Jose G Venegas, Tilo Winkler, Guido Musch, Marcos F Vidal Melo, Dominick Layfield, Nora Tgavalekos, Alan J Fischman, Ronald J Callahan, Giacomo Bellani, and R Scott Harris. Self-organized patchiness in asthma as a prelude to catastrophic shifts. *Nature*, 434(7034):777–782, 2005.

R Verfürth. A posteriori error estimators for convection-diffusion equations.

*Numerische Mathematik*, 80(4):641–663, 1998.

L. Weichert. *Computational Modeling of Multi-Field and Multi-Scale Phenomena in Respiratory Mechanics*. PhD Thesis, Technischen Universität München, 2011.

René Werner, Jan Ehrhardt, Rainer Schmidt, and Heinz Handels. Patient-specific finite element modeling of respiratory lung motion using 4d ct image data. *Medical physics*, 36(5):1500–1511, 2009.

J.B. West. *Respiratory physiology: the essentials*. Lippincott Williams & Wilkins, 2008.

Mary F Wheeler and Xiuli Gai. Iteratively coupled mixed and Galerkin finite element methods for poro-elasticity. *Numerical Methods for Partial Differential Equations*, 23(4):785–797, 2007.

J.A. White and R.I. Borja. Stabilized low-order finite elements for coupled solid-deformation/fluid-diffusion and their application to fault zone transients. *Computer Methods in Applied Mechanics and Engineering*, 197(49-50):4353–4366, 2008.

Joshua A White and Ronaldo I Borja. Block-preconditioned Newton–Krylov solvers for fully coupled flow and geomechanics. *Computational Geosciences*, 15(4):647–659, 2011.

B. Wirth and I. Sobey. An axisymmetric and fully 3D poroelastic model for the evolution of hydrocephalus. *Mathematical Medicine and Biology*, 23(4):363–388, 2006.

P. Wriggers. *Nonlinear finite element methods*. Springer Verlag, 2008.

Son-Young Yi. A coupling of nonconforming and mixed finite element methods for Biot’s consolidation model. *Numerical Methods for Partial Differential Equations*, 2013.

Youbing Yin, Jiwoong Choi, Eric A Hoffman, Merryn H Tawhai, and Ching-Long Lin. A multiscale mdct image-based breathing lung model with time-varying regional ventilation. *Journal of computational physics*, 244:168–192, 2013.

Alexander Ženíšek. The existence and uniqueness theorem in Biot’s consolidation theory. *Applikace matematiky*, 29(3):194–211, 1984.

Tiezhi Zhang, Nigel P Orton, T Rockwell Mackie, and Bhudatt R Paliwal. Technical note: A novel boundary condition using contact elements for finite element based deformable image registration. *Medical physics*, 31(9):2412–2415, 2004.