# Analytical and Numerical Aspects of Porous Media Flow

Dissertation zur Erlangung des Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)
am Fachbereich Mathematik und Informatik
der Freien Universität Berlin

von

Laura Caroline Blank

Berlin, Germany
2019


1. Gutachter: Dr. habil. Alfonso Caiazzo
*Freie Universität Berlin* und
*Weierstraß-Institut für Angewandte Analysis und Stochastik, Berlin*

2. Gutachter: Prof. Dr. Malte Braack
*Christian-Albrechts-Universität zu Kiel*

Datum der Disputation: 14.02.2020

## Acknowledgements

First and foremost I want to express my deep gratitude to my supervisor, Dr. habil. Alfonso Caiazzo. I am very glad that he decided to accompany, encourage, and guide me as his first PhD student through the research in the last years. I would like to thank him for all the time he spent with me in numerous inspiring discussions which extended in particular my scientific horizon. The joy and enthusiasm he has for the research was always contagious and motivational for me.

Furthermore, I owe a sincere thank you to Prof. Dr. Volker John. His constant support was indeed essential for the success of my PhD project.

I am also pleased to thank all my colleagues in the research group 'Numerical Mathematics and Scientific Computing' at the WIAS for the pleasant atmosphere and sharing insightful suggestions with me. Many thanks to my office mates, Dr. Patricio Farrell, Dr. Clemens Bartsch, and Dr. Zahra Lakdawala for the relaxed tea times and interesting blackboard sessions we had. Special thanks also to Dr. Ulrich Wilbrandt who shared in particular his programming expertise with me and to Dr. Christian Merdon and Priv.-Doz. Dr. Alexander Linke for many valuable discussions.

Moreover I want to acknowledge my collaboration partners Prof. Dr. Franz Chouly, Prof. Dr. Alexei Lozinski, Prof. Dr. Joaquin Mura, and Dr. Ernesto Meneses Rioseco. I enjoyed the numerous fruitful visits, phone calls, and mails.

Of special importance for me was the encouragement of my beloved ones. My parents Birgitt and Abdel Gadir, my sister Jillian, my grandmother Else, as well as my very close friends Vivienne and Anne, supported me morally and emotionally throughout writing this thesis and my life in general. Most of all I appreciate the unconditional love, the continuous support, and the patience of my adorable fiancé Max. He watched my back, helped me through the difficult stages of this PhD, and brightens my everyday life.

There were many more friends and colleagues who have contributed in various ways but have not been mentioned personally. Thank You.

# Abstract

The Brinkman equations model fluid flow through porous media and are particularly interesting in regimes where viscous shear effects cannot be neglected. Two model parameters in the momentum balance function as weights for the terms related to inter-particle friction and bulk resistance. If these are not in balance, then standard finite element methods might suffer from instabilities or error estimates might deteriorate. In particular the limit case, where the Brinkman problem reduces to a Darcy problem, demands for special attention.

This thesis proposes a low-order finite element method which is uniformly stable with respect to the flow regimes captured by the Brinkman model, including the Darcy limit. To that end, linear equal-order approximations are combined with a pressure stabilization technique, a grad-div stabilization, and a penalty-free non-symmetric Nitsche method. The combination of these ingredients allows to develop a robust method, which is proven to be well-posed for the whole family of problems in two spatial dimensions, even if any Brinkman parameter vanishes. An a priori error analysis reveals optimal convergence in the considered norm.

A convergence study based on problems with known analytic solutions confirms the robust first order convergence for reasonable ranges of numerical (stabilization) parameters. Further, numerical investigations that partly extend the theoretical framework are considered, revealing strengths and weaknesses of the approach.

An application motivated by the optimization of geothermal energy production completes the thesis. Here, the proposed method is included in a multi-physics discrete model, appropriate to describe the thermo-hydraulics in hot, sedimentary, essentially horizontal aquifers. An immersed boundary method is adopted in order to allow a flexible, automatic optimization without re-generating the computational mesh. Utilizing the developed computational framework, the optimized multi-well arrangements with respect to the net energy gain are presented and discussed for different geothermal and hydrogeological setups. The results show that taking into account heterogeneous permeability structures and variable aquifer temperatures might drastically affect the optimal configuration of the wells.

# Contents

# 1 Introduction

## 1.1 Motivation

The study of the dynamics of fluids which are located in porous materials is relevant in a great number of industrial processes including the fields of civil engineering (e.g., thermal insulation of building material), petroleum production (e.g., fuel cells), and industrial filter design (e.g., open foams). Further it is also the fundament for the description of numerous natural phenomena in disciplines as for example hydro-geology (e.g., groundwater flow) and biomedical engineering (e.g., tissue perfusion and noninvasive diagnosis). In that respect, questions related to the design of products towards desired filter properties or to understanding the effects due to mid- and long-term processes are raised and demand for detailed investigations. Depending on the context, appropriate experiments might be highly inefficient in terms of financial costs, human resources, and time. In these situations the power of numerical simulations becomes evident.

Although enhanced applications often involve further phenomena, the omnipresent module is the fluid flow. Given a detailed description of the porous matrix on the microscopic scale, a reasonable modeling approach is to consider the (Navier–)Stokes equations in the pore space with no-slip boundary conditions at the solid walls of the matrix. In general and in particular for natural porous media, the required analytic description of the pore geometry of a porous medium is often either not available or highly complex such that another model approach is needed. The key idea is to consider the flow on a larger scale instead, which does not account for the microscopic details of the porous structure. A family of macroscopic porous media flow models based on the idea to interpret a porous medium as an averaged continuum is the result.

In this respect, the simplest system of partial differential equations is based on Darcy's law, which states a linear relation between the pressure gradient and the filter velocity and hence accounts only for the viscous damping force due to the porous mass. Tied to the simplicity of the Darcy model is its limitation. In regimes where viscous (shear) stresses become relevant, the model fails to appropriately prescribe the fluid flow, which is particularly delicate with respect to high permeabilities and boundary effects. Moreover, the differentiation between non- and Darcian flow regimes is not trivial. In order to influence the range of validity, appropriate additional terms can be included in the model.

In this spirit, the Brinkman model was proposed, which takes into account viscous stresses via an additional parameter-dependent macroscopic velocity term. To be precise, with respect to the velocity, the differential operator in the Brinkman momentum equation is a weighted sum of the differential operators in the Stokes and Darcy momentum equations. Consequently, the Brinkman model

captures the whole range of equations between these two limit cases, controlled by the values of two physical coefficients.

From the mathematical point of view, the two limit problems (Stokes and Darcy) correspond to different functional settings in terms of the regularity of the function spaces and the boundary conditions, that yield well-posed weak formulations. This flexibility of the Brinkman model might also affect the numerical discretization.

One of such approaches and one of the most popular discretization methods, especially in computational fluid dynamics (CFD), is the finite element method. Here the well-posedness of mixed finite element problems relies on the compatibility of the finite element spaces which is generally speaking one of the main difficulties in the design of finite element methods. For the Brinkman problem the situation is even more complicated since the finite element spaces have to be compatible for two different problems, a Stokes problem and a Darcy problem. Additionally, an alternative to the typically strong imposition of essential boundary conditions in the ansatz space has to be utilized in order to enable a smoother parameter-dependent transition inherent to the problem. Altogether, such a method would allow to write and use code, with little regard to whether the Stokes or the Darcy regime is appropriate. Furthermore, approaches which are computationally cheap and require a low implementational extra effort, based on elementary finite element spaces, are very attractive in practice.

The development and analysis of a single finite element method that fulfills this complex of objectives, i.e, which in particular can handle any choice of parameter values in the Brinkman problem, is challenging. In fact it necessitates a combination of non-standard approaches and an extension of the existing theory.

The availability of such a method would contribute to research areas of high impact as for example those which tackle urgent questions related to the continuously increasing global energy consumption and the climate change. It is known that the traditional usage of fossil fuels as energy sources is temporally limited and that burning processes emit greenhouse gases, provoking global warming. In particular the impressive worldwide civil engagement, known as "Fridays for Future", has recently heated up the debate and insists on concrete short- to mid-term political plans in order to diminish global warming. To that end an energy transition has to be pushed further forward, based on the idea, that renewable energy has to become the main resource of energy. In this context, the extraction of energy in form of heat from geothermal reservoirs has gained attention and geothermal plants are progressively implemented. Such reservoirs might possess geological characteristics which influence the thermo-hydraulics based on high as well as low permeabilities, such that a flexible porous media flow model as the Brinkman model becomes an interesting ingredient. The increasing number of geothermal installations is linked to an increasing concentration, such that optimized well positioning taking into account heterogeneities, interference, and sustainability becomes an important issue that has to be addressed for a strategic positioning of new wells.

## 1.2 Main Contributions

A robust, unconditionally stable, low-order finite element method for the Brinkman problem is proposed. It is designed to behave well in both the Stokes limit and the Darcy limit considered in its dual form. While the Darcy problem requires weak continuity of the normal velocity, for the Stokes problem both, the normal and the tangential components have to be (weakly) continuous. Two main difficulties arise from this difference in the discrete setting as well - the appropriate choice of the finite element spaces suitable for both limits and a proper treatment of essential boundary conditions.

Firstly, since both problems can be cast in the theory of (generalized) saddle point problems, equal-order interpolation combined with a residual-based pressure stabilization technique is employed. Here different variants can be considered and we will use the non-symmetric Galerkin least squares (GLS) method also called Douglas–Wang method which was originally introduced as an unconditionally stable formulation of the Stokes problem. Focusing on computationally efficient approaches, continuous piecewise linear polynomials are chosen. Further a grad-div stabilization is used to maintain stability in the Darcy limit.

Secondly, the Nitsche method is employed to weakly impose the essential boundary conditions. This approach enables to formulate the Brinkman problem at the discrete level using the same finite element spaces, and independent of the values of the physical parameters. Here, the penalty-free non-symmetric Nitsche method, that was recently proven to be well-posed for incompressible elasticity problems, will be used.

The resulting finite element method for the Brinkman problem will be analyzed in detail (in 2D) concerning stability and a priori error estimation with the following results:

(a) Well-posedness for any choice of parameters including the limit cases,
(b) Optimal convergence of order one in a mesh-dependent norm including the limit cases.

For the validity of the well-posedness proof in the Darcy limit, the penalty-free non-symmetric Nitsche method is extended by an additional so-called corner stabilization. The latter concerns only the dofs at the corners of the domain (in 2D) whose two adjacent boundary edges are both associated with weakly imposed essential boundary conditions due to Nitsche. In that turn, the analysis proving the existence and uniqueness of a solution and convergence of the penalty-free non-symmetric Nitsche method in combination with a non-symmetric GLS method, a grad-div stabilization, and a corner stabilization can be considered valid for the Stokes and the Darcy problems separately. Therefore, the existing theory is extended.

Based on the theoretical results, extensive parameter studies which confirm the proven robust convergence of the method are performed and the impact of different parameter choices on the condition number of the finite element matrix is numerically assessed. Moreover, different numerical tests based on benchmarks available in literature are discussed in order to assess the potential

competitiveness of the proposed method. To better understand its advantages and limitations, these partly go beyond the theory.

Finally, a thermo-hydraulic model for the simulation of geothermal reservoir developments is proposed. Its main features are a sequential coupling of essentially horizontal Brinkman flow and heat transfer, and an immersed boundary method to model the wells. The latter approach decouples the mesh generation process from the application of (derivative-free) optimization algorithms. The model is used to optimize the placement (with respect to the produced net energy) of multi-well arrangements in heterogeneous aquifers using data from the Upper Jurassic formation in the greater Munich area. The simulations reveal that the production and injection flow rates, the heterogeneous permeability structure, and the aquifer temperature field influence the optimal placement. Through an optimization with temporal and spatial constraints, the expected net energy gain can be significantly influenced.

## 1.3 Outline

**Chapter 2** begins with an introduction to important concepts and terminology in the context of porous media flow. Following the historical line, we will discuss the models which are highly relevant in this thesis: the Stokes model, the Darcy model, and finally the Brinkman model, including their derivations, limitations, and extensions.

Based on this knowledge, **Chapter 3** starts with the presentation of elementary definitions and relevant results from functional theory, in particular concerning the weak uniqueness and existence theory. Appropriate functional settings for the (non-dimensionalized) Brinkman boundary value problem in its weak form will be identified for the limit cases separately. This will reveal similarities and differences, where the latter in particular manifest in different appropriate functional spaces and boundary conditions. The result are weak formulations of the Stokes respectively Darcy problems which will be proven to admit a unique solution using the solution theory for saddle point problems.

**Chapter 4** is dedicated to the derivation of appropriate ingredients of a robust finite element method for the Brinkman problem. After a condensed introduction to the finite element method and a presentation of theoretical results, standard Galerkin methods for the limit cases and their missing potential to build simple, unified approaches for the Brinkman problem are discussed.

Keeping the focus on low-complexity methods, stability issues of low-order Lagrangian finite elements are figured out. Among the possibilities to cure the incompatibility of certain low-order finite element pairs, pressure stabilizations that consist of adding mesh-dependent weighted element-wise residuals of the momentum equation to the discrete formulation will be considered. Targeting continuous equal-order approximations, the non-symmetric GLS method, which is known for its unconditional stability (for the Stokes problem), will be presented.

In the last part of this chapter, we will introduce approaches to weakly impose essential boundary conditions in the finite element context as opposed to the

common strong imposition, which refers to their explicit incorporation in the ansatz space. A presentation of the Nitsche method and its variants, which are based on modifications of the discrete formulation, will lead the reader finally to the penalty-free non-symmetric Nitsche method.

An unconditionally stable and robust equal-order finite element formulation for the Brinkman problem in (2D) is proposed in **Chapter 5**. The key ingredients, as developed in the foregoing chapter, are the stabilized low-order equal-order finite element pair and a grad-div stabilization, combined with the penalty-free non-symmetric Nitsche method and a corner stabilization. A weighted, mesh-dependent norm, which reduces to appropriate norms in the Stokes and Darcy limits and measures the velocity respectively its normal component at the Nitsche boundary, is considered in the finite element analysis. The presented finite element formulation is then proven to yield a well-posed problem in particular in both of the respective limits, the Stokes limit and the Darcy limit. The grad-div and corner stabilizations, are necessary for the presented proof in the Darcy limit only. Here, the corner stabilization emanates from the inability of the continuous pressure space to represent the possibly discontinuous normal velocity along corners of the computational domain and therefore stabilizes jumps of the normal velocity across corners of the Nitsche boundary. Finally, optimal a priori error estimates are derived, keeping track of the physical and numerical parameters.

**Chapter 6** provides numerical results for the method proposed in Chapter 5. Convergence studies confirming the theoretical outcomes, in particular the robustness, are performed. Also the impact of the stabilization parameters on the condition number of the coefficient matrix is tested here. Furthermore, benchmarks whose analytical solution is not available are used to assess the sensitivity with respect to the GLS stabilization parameter and also to investigate the qualitative competitiveness of the proposed method with other stabilized methods and with multi-domain approaches (Stokes–Darcy coupling with interface conditions).

**Chapter 7** is devoted to an application of current interest, geothermal energy supply. The central goal is to simulate the flow field and the evolution of the cold water front from injection towards production wells in order to estimate the net energy gain by the operation of multi-well arrangements - utilizing solely free software components and libraries. Therefore, the proposed finite element method from Chapter 5 is adopted to compute an advection field, used to determine the corresponding heat transfer in an essentially horizontal aquifer with geological properties similar to the Upper Jurassic formation in the Greater Munich area. The usage of an immersed boundary approach to model sources and sinks of mass increases the flexibility of the optimization without necessitating a high mesh resolution of well boundaries and thus remeshing during the optimization. A comparative numerical study considers hexagonal and lattice type multi-well arrangements with varying heterogeneous permeability structures, production and injection flow rates, and background temperature fields, for an 80+ years operation. Hence, in particular the impact

of interference on the optimal well placement with respect to the net energy gain is included.

# 2 Background

The purpose of this introductory section is to provide, besides historical information, a solid mathematical and physical foundation, crucial for understanding the basic notions and the mathematical models that are used to describe porous media flow. Therefore, we will briefly discuss relevant mathematical operators and properties of fluids and porous media before going into detail with respect to the models of interest, namely the (Navier–)Stokes, Darcy, and Brinkman models.

## 2.1 Notation and Basic Properties of Differential Operators

Let us begin with the definition of first and second order differential operators.

**Definition 2.1.1** (Differential Operators)**.**
*Let $v\colon \boldsymbol{\Omega} \to \mathbb{R}$ be a scalar function and $\boldsymbol{u}\colon \boldsymbol{\Omega} \to \mathbb{R}^n$, $\boldsymbol{u} = (u_1, u_2, \dots, u_n)^T$, as well as $\tilde{\boldsymbol{u}}\colon \boldsymbol{\Omega} \to \mathbb{R}^n$, $\tilde{\boldsymbol{u}} = (\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n)^T$, be vector-valued functions, all three defined on $\boldsymbol{\Omega} \subset \mathbb{R}^n$, $n \in \mathbb{N}_{>0}$, and sufficiently smooth. Then, we define the following operators:*

1. **Laplace operator:**

$$\Delta v \in \mathbb{R}, \qquad \Delta v := \sum_{i=1}^{n} \frac{\partial^2 v}{\partial x_i^2}\,,$$

$$\Delta \boldsymbol{u} \in \mathbb{R}^n, \quad (\Delta \boldsymbol{u})_i := \sum_{j=1}^{n} \frac{\partial^2 u_i}{\partial x_j^2}\,, \quad i = 1, \dots, n,$$

2. **gradient:**

$$\nabla v \in \mathbb{R}^n, \qquad \nabla v := \left(\frac{\partial v}{\partial x_1}, \frac{\partial v}{\partial x_2}, \cdots, \frac{\partial v}{\partial x_n}\right)^T,$$

$$\nabla \boldsymbol{u} \in \mathbb{R}^{n\times n}, \quad (\nabla \boldsymbol{u})_{ij} := \frac{\partial u_i}{\partial x_j}\,, \quad i, j = 1, \dots, n,$$

3. **divergence for vector-valued functions:**

$$\nabla \cdot \boldsymbol{u} \in \mathbb{R}, \qquad \nabla \cdot \boldsymbol{u} := \sum_{i=1}^{n} \frac{\partial u_i}{\partial x_i}\,,$$

*4.* **divergence for tensor-valued functions:**

$$\nabla \cdot (\nabla \boldsymbol{u}) \in \mathbb{R}^n, \qquad (\nabla \cdot (\nabla \boldsymbol{u}))_i := \nabla \cdot (\nabla \boldsymbol{u})^i, \qquad i = 1, \ldots, n,$$

*($(\nabla \boldsymbol{u})^i$ being the $i$-th row of $\nabla \boldsymbol{u}$ as a column vector),*

*5.* **tensor-product:**

$$\nabla \boldsymbol{u} \colon \nabla \tilde{\boldsymbol{u}} \in \mathbb{R}, \qquad \nabla \boldsymbol{u} \colon \nabla \tilde{\boldsymbol{u}} := \sum_{i=1}^{n} \nabla u_i \cdot \nabla \tilde{u}_i = \sum_{i,j=1}^{n} \frac{\partial u_i}{\partial x_j} \frac{\partial \tilde{u}_i}{\partial x_j},$$

*6.* **symmetric part of the gradient:**

$$\mathbb{D}(\boldsymbol{u}) \in \mathbb{R}^{n \times n}, \qquad \mathbb{D}(\boldsymbol{u}) := \frac{\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^T}{2}.$$

**Lemma 2.1.2.** *For $v$ and $\boldsymbol{u}$ as in Definition 2.1.1, both sufficiently smooth and $n \in \{2, 3\}$, the following identities hold:*

$$\begin{aligned}
&(i) \quad \nabla \cdot \nabla v = \Delta v, \\
&\qquad \nabla \cdot \nabla \boldsymbol{u} = \Delta \boldsymbol{u}, \\
&(ii) \quad \nabla \cdot (v \boldsymbol{u}) = \nabla v \cdot \boldsymbol{u} + v \nabla \cdot \boldsymbol{u}, \\
&(iii) \quad \nabla \cdot (\nabla \boldsymbol{u})^T = \nabla (\nabla \cdot \boldsymbol{u}).
\end{aligned}$$

*Furthermore, if the divergence of $\boldsymbol{u}$ is constant (in particular $\nabla \cdot \boldsymbol{u} = 0$), then*

$$\begin{aligned}
&(iv) \quad \nabla \cdot (\nabla \boldsymbol{u})^T = \boldsymbol{0}, \\
&(v) \quad 2 \nabla \cdot \mathbb{D}(\boldsymbol{u}) = \nabla \cdot \nabla \boldsymbol{u}.
\end{aligned}$$

*Proof:* The proof is straightforward by using the symmetry of second order partial derivatives (for sufficiently smooth functions) and the definitions of the differential operators. In order to deduce $(iv)$ and $(v)$, $(iii)$ is used. □

**Remark 2.1.3.** *The term sufficiently smooth indicates that the used functions are smooth enough for the applied calculations and considerations. For the definition of, e.g., the Laplace operator in Definition 2.1.1, twice differentiable would be sufficient.*

## 2.2 Fluids and Porous Media

Porous media flow can be phenomenologically modeled, as flow through channels or as flow around obstacles, [Kav02, Sect. 2.1]. Thereby, a porous medium is formally a combination of a solid matrix (formation of particles) and voids, [NB06, Sect. 1]. In order to enable fluid flow through the porous medium, at least some of the voids have to be connected.

**Assumption I:** Throughout this thesis we will focus on porous media with a rigid solid matrix, i.e., neglecting the deformation of the solid structures.

In fact, there is a great variety of naturally occurring porous media, for which examples are beach sand, wood, luffa sponges, or even biological tissue as the human lung and liver. However, due to their valuable contribution to diverse fields, numerous synthetic porous media have been designed and ease our everyday life in form of wipes, diapers, various kinds of filters, or even titanium foam which is used as an implant in human bones, [MFK17].

While synthetic porous media often have a well-defined or even periodic structure, in nature, a porous medium is typically highly irregular on the microscopic scale (the so-called *pore scale*). The size and the shape of the solid matrix and the voids may vary strongly. Additionally, fractures that largely influence the fluid migration might alter the flow pattern significantly.

As a result, the physical quantities describing the flow through such porous media, like the fluid velocity and the pressure, mirror this high irregularity. Several samples of porous media consisting of the same substance, e.g., natural stones, typically have similar properties on a macroscopic level although they vary strongly on a micro-scale. The prediction of such a micro-scale behavior is not only an extremely complex task, which might be even impossible due to limited knowledge of the microscopic structure of the solid matrix. It is also unnecessary from the point of view of application where one is often actually interested in mesoscopic or macroscopic behaviors only, [DV+12, Ch. 2.1]. For a visualization of micro- and macro-scales see Figures 2.1 and 2.2.

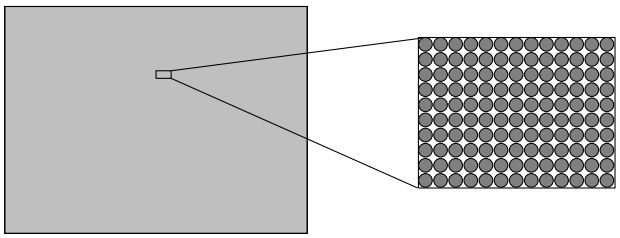


Figure 2.1: Sketch of a macro-scale (left) and a micro-scale (right) for an artificial homogeneous porous medium with solid parts in gray.

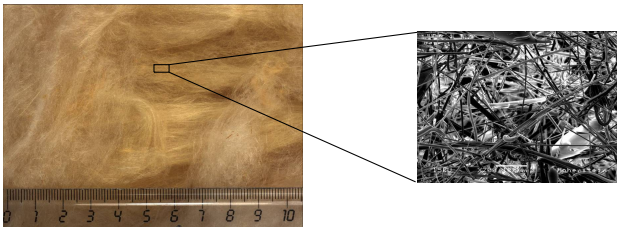


Figure 2.2: Macroscopic (left) and microscopic (right) pictures of mineral wool taken from [ILW09].

This leads to the idea to describe the flow on a mesoscopic- or macroscopic scale using space/volume averaging, which is in practice achieved considering velocity and pressure averages with respect to a so-called *representative elementary volume*.

In the rest of this Section 2.2, fundamental notions concerning porous media flow that underlie the discussions in this thesis are introduced. The basis therefore are the books [Hor97], [NB06], and [Bea72].

### 2.2.1 General Material Properties

In this subsection we will discuss two scale-dependent classifications of substances, *isotropic*, respectively *anisotropic*, and *homogeneous*, respectively *inhomogeneous*. These notions strongly depend on the respective context (scale) and are only meaningful from a mesoscopic or macroscopic (continuum) point of view, and not on, e.g., the atomic level, where all substances are naturally occurring to be anisotropic and inhomogeneous, [BS90, Ch. 9].

#### Homogeneity and Inhomogeneity

*A medium, $\boldsymbol{M} \subset \mathbb{R}^n$ is called homogeneous with respect to some (physical) property $\boldsymbol{P}$ if this property does not vary in space, otherwise $\boldsymbol{M}$ is called inhomogeneous or heterogeneous with respect to $\boldsymbol{P}$.*

Liquid, isothermal water is an example for a homogeneous fluid (with respect to density) and a coffee filter is a homogeneous (with respect to its ability to admit fluid flow) porous medium. A liquid that is heated with different temperatures at different points in space is likely to be inhomogeneous with respect to its resistance to flow (viscosity). Also a composition of two liquids that do not mix, e.g., oil and water, leading to two phases with different physical properties will be inhomogeneous with respect to them. In the literature, the term homogeneous fluid is also used to account for single-fluid phase flow, [MW37].

#### Isotropy and Anisotropy

*A medium $\boldsymbol{M} \subset \mathbb{R}^n$ is called isotropic with respect to some property $\boldsymbol{P}$, if this property does not depend on the direction, else it is called anisotropic.*

Many fluids are in general isotropic with respect to, e.g., density - an example for an exception are liquid crystals, [BS90, Ch. 9]. Concerning porous media, geological formations that serve as reservoirs and wood, whose material properties are different parallel and perpendicular to the grain, are typically anisotropic with respect to the resistance to fluid flow.

### 2.2.2 Properties of Porous Media

For a porous medium described on a macroscopic scale, it is necessary to reflect the microscopic properties in an adequate manner. This is the intention behind notions as *saturation*, *porosity*, and *permeability*. For some further information, see [Kav02, Ch. 2].

#### Saturation

*In the porous media flow context, the saturation $S_f$ is defined as the fraction of the interconnected void space that is occupied by a certain fluid called $f$:*

$$\mathcal{S}_f := \frac{\textit{volume occupied by the fluid } f}{\textit{volume of the void space}} \in [0,1].$$

For $\mathcal{S}_f = 1$ and $f$ being the only fluid in the porous medium, the flow is called *one-phase fully saturated* flow. In case of $m$ fluids exploring the interconnected void in the porous medium, the flow is referred to as an $m$-phase flow.

**Assumption II:** During this thesis, porous media with a totally interconnected void and one-phase fully saturated flow will be presupposed.

#### Porosity

*For a porous medium, the porosity is the dimensionless quantity which gives information about the relation between the solid parts and the voids. The porosity of a medium with a totally interconnected void is defined by*

$$\phi := \frac{\textit{volume of the void space}}{\textit{total volume}} \in [0,1],$$

*with*

$$\textit{total volume} := (\textit{volume of the void space}) \cup (\textit{volume of the porous mass}).$$

Hence it is a measure of the pore space, i.e, the potential fluid capacity of the porous medium, [MW37]. The extreme cases $\phi \in \{0,1\}$ represent a pure solid matrix and a pure void space, respectively.

#### Permeability

*The (intrinsic) permeability $\mathbb{K}$ $[\mathrm{m}^2]$ is a measure of the ability of a porous medium to let fluids pass through. In practice, the unit darcy [Kav02, Sect. 2.1] is used as well, where*

$$1\,\mathrm{darcy} = 9.87 \cdot 10^{-13}\,\mathrm{m}^2.$$

The notation $\mathbb{K}$ indicates that the permeability is in general modeled as a tensor, since the flow can evolve in any spatial direction differently, [Kav02, Ch. 2.1]. For

example natural stratified soils have a permeability which is typically higher in the horizontal direction than in the vertical direction. Formally, $\mathbb{K}$ is the inverse of the resistance to flow due to a porous mass - high values in the permeability tensor corresponds to low resistance to flow by the porous mass, and low values in the permeability tensor induce a high resistance, [SC11, p. 25]. According to [Kav02, Sect. 2.1], the permeability of a porous mass depends, amongst others, on the interstitial surface area and the path that fluid particles might take.

The permeability is a symmetric, positive definite, second-order tensor, such that the existence of its inverse, which will be part of the flow models, is guaranteed, see, e.g., [Ant13] for 3D. It further effectively simplifies for isotropic soils to a multiple of the identity tensor $\mathbb{I}$, i.e., $\mathbb{K} = k\mathbb{I}$, where $k$ is scalar, see [Lia65]. For possibilities to determine the (intrinsic) permeability of a porous medium, see [Ber14], [IL04], and [MW37, Ch. 2].

In fact, the porosity and the permeability of a porous medium are interrelated, although two media might have the same (effective) porosity while having different permeabilities based on different flow paths with different amounts of boundary and bulk effects exerted on the fluid. The precise relationship depends on the (microscopic) properties of the porous medium, as the shapes of the voids and their level of connectedness have an impact on $\mathbb{K}$, [MW37, Sect. 2.3]. Hence, explicit relations between permeability and porosity can only be attempted in concrete examples, depending on the medium. An example for packed grains is the Kozeny–Carman equation ([Cos06], [AC03], [Car97, (10)], [NB06, Ch. 1.4.2]), which states as

$$\mathbb{K} = \frac{c\, d^2\, \phi^3}{(1-\phi)^2},$$

where $c$ is some constant and $d$ is the average diameter of the grains.

In Table 2.1, examples for porous media together with typical porosity and permeability ranges are given.

Table 2.1: Examples for porous media, their porosities $\phi$, and permeabilities $\mathbb{K}$ in [$m^2$] collected from [NB06], [Bea72], [Kav02], and [GK10].

| **Medium** | **Porosity** | **Permeability** |
|---|---|---|
| Soil | $4.3 \cdot 10^{-1} - 5.4 \cdot 10^{-1}$ | $2.9 \cdot 10^{-13} - 1.4 \cdot 10^{-11}$ |
| Leather | $5.6 \cdot 10^{-1} - 5.9 \cdot 10^{-1}$ | $9.5 \cdot 10^{-14} - 1.2 \cdot 10^{-13}$ |
| Limestone | $4.0 \cdot 10^{-2} - 1.0 \cdot 10^{-1}$ | $2.0 \cdot 10^{-15} - 4.5 \cdot 10^{-14}$ |
| Sand | $3.7 \cdot 10^{-1} - 5.0 \cdot 10^{-1}$ | $2.0 \cdot 10^{-11} - 1.8 \cdot 10^{-10}$ |
| Sand stone | $8.0 \cdot 10^{-2} - 3.8 \cdot 10^{-1}$ | $5.0 \cdot 10^{-16} - 3.0 \cdot 10^{-12}$ |
| Clean gravel | / | $1.0 \cdot 10^{-7} - 1.0 \cdot 10^{-9}$ |
| Textile | / | $1.0 \cdot 10^{-10} - 1.0 \cdot 10^{-13}$ |

### 2.2.3 Properties of Fluids

The description of the physical behavior of fluids requires the knowledge of certain concepts that will be partly defined and discussed in the following.

#### Compressibility and Incompressibility

*A fluid is called compressible if its density (respectively volume) changes in the presence of a pressure variation. It is called incompressible if the volume respectively density is constant in case of pressure variation, see, e.g., [Kav02, Sect. 2.1].*

In reality, all fluids are compressible, i.e., their density is sensitive to pressure and thus temperature. This property is much more present for gases than for fluids. In order to simplify the models for fluid dynamics, in case of sufficiently small changes of pressure and temperature and corresponding small changes of density (and thus volume), the flow of a fluid often can be modeled as *incompressible flow*. For homogeneous fluids, incompressible flow is characterized by a constant density, [LL59, §10], [Kav02, Sect. 2.1].

#### Viscosity and Newtonian Fluids

*The viscosity of a fluid describes its resistance to deformation (flow) due to mechanical stress, i.e., shear and tensile/compressive stresses. Fluids that have no resistance to shear and tensile/compressive stresses are called inviscid otherwise they are called viscous. Ideal fluids are by definition inviscid (viscosity = 0).*

In [New87, Book II, Sect. 9], Newton hypothesized the relation which is nowadays referred to as *Newtons law of viscosity*. In modern language it states that for an incompressible, laminar flow of an isotropic, viscous fluid the shear stress is proportional to the shear rate (strain/deformation rate due to shearing), see [Bar94], [Gal10], [BSL06]. For an essentially one-dimensional (1D) flow, this linear stress-deformation relation is mathematically formulated by

$$\mathbb{V}_{yx} := \frac{F}{A} = \mu \frac{\partial u}{\partial y}, \tag{2.2.1}$$

where the proportionality constant $\mu$ is named *dynamic* (sometimes also *absolute*) *viscosity* [Pa s][1] and $\mathbb{V}_{yx}$ [Pa] is a component of viscous (shear) stress corresponding to the force $F$ in $x$ direction, exerted on a unit area $A$ perpendicular to the $y$-axis. Thus, (2.2.1) states that the shear stress is proportional to the velocity gradient in the direction perpendicular to the shearing direction, and the proportionality constant is the fluid-specific dynamic viscosity. For an illustration explaining the setting for (2.2.1), see Figure 2.3.

[1] Note that it holds [Pa] = [N/m²] = [kg/ms²].

Thought experiment:

Consider a viscous fluid volume that is located between two parallel plates of area $A$ with distance $h$. Exerting mechanical stress in form of a constant force $F$ on the top plate in the direction of the $x$-axis enforces its motion with constant velocity $u_x$ and therefore results in a deformation of the confined fluid volume. Thereby, firstly, the force is proportional to the area of the plates, and secondly the force is proportional to the shear rate. Altogether, this gives Newtons law of viscosity.

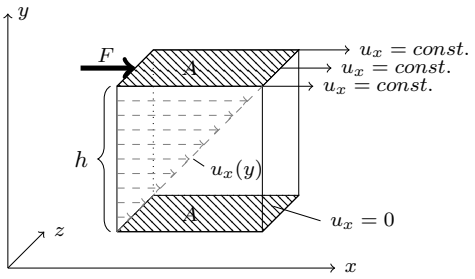



Figure 2.3: Illustration of the setting for Newton's law of viscosity for laminar, essentially 1D flow (in the sense that $\boldsymbol{u} = (u_x(y), 0, 0)$) – Couette flow.

The generalization of Newton's law of viscosity to three spatial dimensions and compressible flows can be written as [Joh16, Sect. 2.2]

$$\mathbb{V} = 2\mu\mathbb{D}(\boldsymbol{u}) + \left(\zeta - \frac{2}{3}\mu\right)(\nabla \cdot \boldsymbol{u})\,\mathbb{I}, \tag{2.2.2}$$

where $\zeta$ is called *dilatational viscosity*, *bulk viscosity*, or *second viscosity* [Pa s] and the tensor $\mathbb{D}(\boldsymbol{u})$ [1/s] (already defined in Definition 2.1.1) is in this context also referred to as the *deformation rate tensor* or *rate of strain tensor*. The dilatational viscosity $\zeta$ accounts for the stress due to density changes during the flow, see [BSL06, Ch. 1] and [Gal10, pp. 74]. The viscous stress tensor $\mathbb{V}$ represents the internal forces due to friction between particles, caused by a discrepancy in particle velocities. From that point of view, derivatives of the velocity are crucial ingredients of $\mathbb{V}$. However, the relationship (2.2.2) is by far not satisfied for all fluids.

*Those (isotropic) fluids which fulfill Newton's law of viscosity (equation* (2.2.1) *respectively* (2.2.2)*), i.e., whose viscosity is independent of shear stress and shear rate, are called Newtonian, [Ach90, p.* 26*], [Gal10, pp. 74].*

The viscosity of a Newtonian fluid depends only on the pressure (at least for liquids) and the temperature [Bar94]. Increasing the temperature of a fixed volume of fluid will lead to an increase of pressure based on the reduction of intermolecular distance. While changes in temperature might highly affect the fluid viscosity, pressure variations have a much weaker influence [Kne90].

Besides all gases and most of the simple liquids that obey Newton's law of viscosity, there are fluids which behave *non-Newtonian*, [Gal10, pp. 96]. For example dilatant/shear-thickening fluids as starch mixed with water and blood (see [Mez90]) belong to the class of non-Newtonian fluids.

*The kinematic viscosity* $\nu$ $[\mathrm{m}^2/\mathrm{s}]$*, sometimes also referred to as momentum diffusivity, [Gal10, p. 79], is defined as the ratio of dynamic viscosity* $\mu$ *and fluid density* $\rho$ $[\mathrm{kg}/\mathrm{m}^3]$*, namely*

$$\nu := \frac{\mu}{\rho}. \tag{2.2.3}$$

In Tables 2.2 and 2.3, the approximate dynamic and kinematic viscosities for some common fluids at constant temperature $T = +20°\mathrm{C}$ are presented.

Table 2.2: Dynamic viscosities $\mu$ in [Pa s] taken from [Mez90].

| **Fluid** | **Dynamic Viscosity** |
|---|---|
| Air | $1.8 \cdot 10^{-5}$ |
| Water | $1.0 \cdot 10^{-3}$ |
| Blood Plasma | $1.7 \cdot 10^{-3}$ |
| Olive Oils | $1.0 \cdot 10^{-1}$ |
| Honey | $1.0 \cdot 10^{+1}$ |

Table 2.3: Kinematic viscosities $\nu$ in $[\mathrm{m}^2/\mathrm{s}]$ taken from [LL59].

| **Fluid** | **Kinematic Viscosity** |
|---|---|
| Air | $1.5 \cdot 10^{-5}$ |
| Water | $1.0 \cdot 10^{-6}$ |
| Alcohol | $2.2 \cdot 10^{-6}$ |
| Glycerine | $6.8 \cdot 10^{-4}$ |
| Mercury | $1.2 \cdot 10^{-7}$ |

### Effective Viscosity

Moreover, there is also a notion of *effective viscosity* $\mu_{\text{eff}}$ [Pa s] in the context of porous media flows, sometimes (e.g., in [Ang99]) referred to as Brinkman viscosity.[2] This parameter can be introduced to describe the viscosity of suspensions of fluids, which consist at least of two phases on the micro-scale, but are effectively considered as single phase and Newtonian on the macro-scale. Therefore, the effective viscosity shall take into account the physics of all constituents. At the same time, the velocity and stress components in (2.2.2) have to be replaced by their volume averaged versions (abusing the same notation), for more details see [BSL06, §1.6].

There are various proposals for the definition of the effective viscosity. One of the first ideas goes back to Einstein [Ein06, §3][3] where he investigated the viscosity of dilute suspensions of rigid spherical particles (where the interparticle distance is much larger than the particle size) and proposed the relation

$$\mu_{\text{eff}} = \mu \left(1 + \frac{5}{2}\left(\frac{\text{volume of the rigid spheres}}{\text{total volume}}\right)\right) = \mu \left(1 + \frac{5}{2}(1-\phi)\right), \tag{2.2.4}$$

which yields $\mu_{\text{eff}} \in [\mu,\, 3.5\mu]$. For alternative definitions of $\mu_{\text{eff}}$ see [BSL06, §1.6]. We will come back to this topic in Subsection 2.3.3.

[2] This naming is due to its appearance in the Brinkman model which will play a key role in this thesis and will be introduced in detail later.

[3] In the original version in [Ein06] there was a mistake which is not shown here, see [BSL06, §1.6].

**Assumption III:** Whenever incompressible flow is considered in this thesis, the respective fluid is supposed to be homogeneous, such that one can infer constant density. Moreover, only Newtonian fluids, isothermal flows, and isotropic porous media (thus effectively scalar permeabilities) will be considered unless otherwise stated.

## 2.3 Boundary Value Problems for Porous Media Flow

The purpose of this section is to introduce three models describing fluid flow, that - considered as initial and/or boundary value problems - can be solved under certain conditions, at least approximately, using numerical methods. Therefore we will state, derive, and discuss the systems of partial differential equations, which define the Stokes, Darcy, and Brinkman problems.

**Assumption IV:** From now on we assume that $\boldsymbol{\Omega} \subset \mathbb{R}^n$, $n = 2, 3$, is a bounded domain (an open and connected set). The boundary $\Gamma := \partial\boldsymbol{\Omega}$ is assumed to be sufficiently regular (e.g., *Lipschitz boundary*) such that a unit outer normal vector $\boldsymbol{n}$ can be defined almost everywhere on $\Gamma$.

For an illustration of a flow domain, see Figure 2.4. Bold faced letters will be used in this dissertation to indicate vectorial quantities.

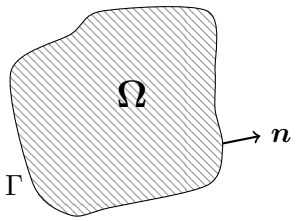



Figure 2.4: Illustration of a possible domain $\boldsymbol{\Omega}$, its boundary $\Gamma$, and corresponding normal vector $\boldsymbol{n}$.

### 2.3.1 The Incompressible Stokes Equations

For the steady, incompressible flow of a Newtonian fluid in a domain $\boldsymbol{\Omega}$, for which the convection effects can be omitted (creeping flow), the (isothermal) fluid motion, described by the state $(\boldsymbol{u}, p)$ is completely determined by the system of partial differential equations

$$-2\nabla \cdot (\mu \mathbb{D}(\boldsymbol{u})) + \nabla p = \boldsymbol{f} \quad \text{in } \boldsymbol{\Omega}, \tag{2.3.1a}$$

$$\nabla \cdot \boldsymbol{u} = 0 \quad \text{in } \boldsymbol{\Omega}, \tag{2.3.1b}$$

equipped with appropriate boundary conditions. Here $\boldsymbol{u}(\boldsymbol{x}) : \boldsymbol{\Omega} \to \mathbb{R}^n$ is a velocity field [m/s], $p(\boldsymbol{x}) : \boldsymbol{\Omega} \to \mathbb{R}$ denotes the pressure [Pa], $\boldsymbol{f}(\boldsymbol{x}) : \boldsymbol{\Omega} \to \mathbb{R}^n$ contains external forces [Pa/m], and $\mu$ is the dynamic viscosity. The first equation (2.3.1a) is called *momentum equation* or *momentum balance* and (2.3.1b) is

named *continuity equation* or *incompressibility condition*. The system (2.3.1) is a simplification of the Navier–Stokes equations (as we will see) and is referred to as the (incompressible) Stokes equations.

If the fluid is homogeneous with respect to viscosity, i.e., the viscosity is constant in space, the momentum balance (2.3.1a) simplifies to

$$-\mu\Delta\boldsymbol{u} + \nabla p = \boldsymbol{f} \quad \text{in } \boldsymbol{\Omega},$$

which will be argued later in this section.

### Derivation – A Short Introduction to Fluid Dynamics

Fluid dynamics is a sub-discipline of fluid mechanics and addresses the description of the motion of fluids, i.e., liquids and gases, due to the presence of forces. In order to better understand the (Navier–)Stokes model, a derivation of the system of equations from the point of view of theoretical physics, based on conservation of mass and momentum principles, will be presented in the following paragraphs. For more information see [Boe12, Sect. 3.4.2], [LL59], [Ach90], [Bat99], [ESW14], and [Joh16].

In (continuum) fluid dynamics one focuses on a scale reflecting the behavior of so-called particles, which are assumed to always contain numerous molecules. The basic functions that describe the motion of a fluid in a domain $\boldsymbol{\Omega} \subset \mathbb{R}^n$ for the time interval $[0, T]$ are the time-dependent fluid velocity and pressure for which we use the same notation as in the steady case, i.e., $\boldsymbol{u}(\boldsymbol{x}, t) : \boldsymbol{\Omega} \times [0, T] \to \mathbb{R}^n$ and $p(\boldsymbol{x}, t) : \boldsymbol{\Omega} \times (0, T] \to \mathbb{R}$, and the fluid density $\rho(\boldsymbol{x}, t) : \boldsymbol{\Omega} \times (0, T] \to \mathbb{R}$ with unit $[\mathrm{kg}/\mathrm{m}^3]$.

### I) The Conservation of Mass - The Equation of Continuity

The equation of continuity is one of the fundamental equations of fluid motion and represents the basic principle of conservation of mass: The amount of fluid flowing out of a volume $\boldsymbol{V}_0$ (flux of mass $\rho\boldsymbol{u}$ across the boundary $\partial\boldsymbol{V}_0$),

$$\int_{\partial\boldsymbol{V}_0} \rho\,\boldsymbol{u} \cdot \boldsymbol{n}\,\mathrm{d}\boldsymbol{s},$$

has to equal the decrease in the mass of fluid in the volume $\boldsymbol{V}_0$ per unit time (rate of change of mass in $\boldsymbol{V}_0$),

$$-\frac{\partial}{\partial t}\underbrace{\int_{\boldsymbol{V}_0} \rho\ \mathrm{d}\boldsymbol{x}}_{\text{mass}}.$$

Assuming sufficient smoothness of the boundary $\partial\boldsymbol{V}_0$ and the constituting functions, the Gaussian theorem (also divergence theorem) can be applied to obtain

$$0 = \frac{\partial}{\partial t}\int_{\boldsymbol{V}_0} \rho\ \mathrm{d}\boldsymbol{x} + \int_{\partial\boldsymbol{V}_0} \rho\,\boldsymbol{u} \cdot \boldsymbol{n}\ \mathrm{d}\boldsymbol{s} = \int_{\boldsymbol{V}_0} \left(\frac{\partial\rho}{\partial t} + \nabla \cdot (\rho\,\boldsymbol{u})\right)\mathrm{d}\boldsymbol{x},$$

where spatial integration and temporal differentiation were interchanged. Since this equation has to hold for arbitrary volumes $\boldsymbol{V}_0$, the integrand has to vanish. Hence, the **equation of continuity** is given by

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{u}) = 0. \tag{2.3.3}$$

In the special case of a homogeneous fluid, whose density does neither depend on time nor on space and is in particular not zero, i.e., there is no (noticeable) compression or expansion of the fluid throughout its motion, the flow is *incompressible* (see Subsection 2.2.3). Thus, the continuity equation (2.3.3) simplifies to

$$\nabla \cdot \boldsymbol{u} = 0, \tag{2.3.4}$$

which is then also referred to as **incompressibility constraint** or *divergence constraint*. A velocity field $\boldsymbol{u}$ which satisfies equation (2.3.4) is called *divergence-free*. For a divergence-free velocity field it holds (applying the Gaussian theorem):

$$0 = \int_{\boldsymbol{\Omega}} \nabla \cdot \boldsymbol{u} \ \mathrm{d}\boldsymbol{x} = \int_{\Gamma} \boldsymbol{u} \cdot \boldsymbol{n} \ \mathrm{d}\boldsymbol{s},$$

which reflects mass conservation in the sense that the inflow and the outflow with respect to the domain in question are in balance.

Note that the conclusion that incompressibility necessitates constant density is only meaningful for fluids which are homogeneous respectively uniform with respect to density (which is the case by Assumption III). For example non-homogeneous fluids as the mixture of different particles or some molecules and ions, might reveal a non-constant density even if their flow is incompressible.

### II) The Conservation of (Linear) Momentum - Motion of Ideal and Viscous Fluids

The description of fluid motion, where thermal conductivity and viscosity are negligible, is given by the fluid dynamics of ideal fluids. In this situation, the well known *momentum balance in the Euler equations* describes the conservation of (linear) momentum based on Newton's second law of motion. For constant-mass systems it reads:

$$\boldsymbol{F} = m\,\boldsymbol{a},$$

where $\boldsymbol{F}$ is the net/total force [$\mathrm{kg\,m/s^2}$], $m$ is the mass [kg], and $\boldsymbol{a}$ is the acceleration [$\mathrm{m/s^2}$]. The net force $\boldsymbol{F}$ considered per unit volume can be decomposed into internal forces per unit volume, $\boldsymbol{f}_{\mathrm{int}}$ [$\mathrm{Pa/m}$], and external forces per unit volume, $\boldsymbol{f}_{\mathrm{ext}}$ [$\mathrm{Pa/m}$]. Note that energy dissipation due to internal friction (viscosity) and heat exchange (thermal conductivity) is not taken into account here, since we are describing the motion of ideal fluids. Therefore, the internal force field is determined by the pressure only. Since the pressure exerted on a fluid volume

$\boldsymbol{V}_0$ is given by the pressure $p$ applied to the boundary $\partial\boldsymbol{V}_0$ of the fluid volume directed into the volume $\boldsymbol{V}_0$, $\boldsymbol{f}_{\text{int}}$ is determined by

$$\int_{\boldsymbol{V}_0} \boldsymbol{f}_{\text{int}} = \int_{\partial\boldsymbol{V}_0} p\,(-\boldsymbol{n}) = -\int_{\boldsymbol{V}_0} \nabla p, \tag{2.3.5}$$

where integration by parts was applied. Proceeding similar as in the derivation of the continuity equation (see [Joh16, Sect. 2.2]) and assuming constant density (i.e., incompressible flow and a homogeneous fluid) one obtains

$$\underbrace{\rho}_{\substack{\text{mass per}\\ \text{unit volume}}} \underbrace{\left[\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u}\cdot\nabla)\,\boldsymbol{u}\right]}_{\text{acceleration}} = \underbrace{-\nabla p}_{\boldsymbol{f}_{\text{int}}} + \boldsymbol{f}_{\text{ext}}, \tag{2.3.6}$$

describing the motion of an ideal fluid subject to an external force per unit volume $\boldsymbol{f}_{\text{ext}}$, e.g., $\boldsymbol{f}_{\text{ext}} = \rho\boldsymbol{g}$ with $\boldsymbol{g}$ [m/s²] being a gravitational field.

In case of a Newtonian fluid whose viscosity is not negligible, the motion of the fluid is additionally influenced by internal friction. This leads to the *momentum balance in the Navier–Stokes equations* as an extension of the Euler equation (2.3.6), known as the Cauchy momentum equation

$$\rho\left[\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u}\cdot\nabla)\,\boldsymbol{u}\right] - \nabla\cdot\mathbb{S} = \boldsymbol{f}_{\text{ext}}. \tag{2.3.7}$$

Here, $\mathbb{S}$ is a second-order tensor called *total stress tensor* or *Cauchy stress tensor* whose diagonal entries refer to normal stresses and whose off-diagonal entries represent shear stresses. A model for the total stress tensor $\mathbb{S}$ is given by

$$\mathbb{S} := \mathbb{V} - p\mathbb{I},$$

where $\mathbb{V}$ is the *viscous stress tensor* [Pa] and $\mathbb{I}$ is the second-order identity tensor. The tensor $\mathbb{V}$ models the internal forces due to friction between particles and $p\mathbb{I}$ represents the internal forces acting on the surface of any fluid volume $\boldsymbol{V}_0$ (in accordance with equation (2.3.5)). As observed in Subsection 2.2.2, for Newtonian fluids, $\mathbb{V}$ can be modeled by equation (2.2.2). Since the total stress tensor $\mathbb{S}$ is symmetric (see, e.g., [Joh16, Rem. 2.9]), and $p\mathbb{I}$ is symmetric anyway, $\mathbb{V}$ has to be symmetric as well. Inserting the resulting representation of $\mathbb{S}$ into (2.3.7) gives

$$\rho\left[\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u}\cdot\nabla)\,\boldsymbol{u}\right] \underbrace{- \nabla\cdot\left[2\mu\mathbb{D}\,(\boldsymbol{u}) + \left(\zeta - \frac{2}{3}\mu\right)(\nabla\cdot\boldsymbol{u})\,\mathbb{I}\right]}_{=\nabla\cdot\mathbb{V}} + \underbrace{\nabla p}_{=\nabla\cdot(p\mathbb{I})} = \boldsymbol{f}_{\text{ext}},$$

which is named **(compressible) Navier–Stokes momentum equation**. Apparently the Euler equation (2.3.6) can be derived from the (compressible) Navier–Stokes momentum equation by neglecting the term $\nabla\cdot\mathbb{V}$.

In the special case of incompressible flow of a homogeneous fluid, we can use the continuity equation (2.3.4) to simplify the representation of the viscous stress tensor (2.2.2) to

$$\mathbb{V} = 2\mu\mathbb{D}\left(\boldsymbol{u}\right),$$

such that the compressible Navier–Stokes momentum equation simplifies to the **incompressible Navier–Stokes momentum equation**

$$\rho\left[\frac{\partial \boldsymbol{u}}{\partial t} + \left(\boldsymbol{u}\cdot\nabla\right)\boldsymbol{u}\right] - \nabla\cdot\left(2\mu\mathbb{D}\left(\boldsymbol{u}\right)\right) + \nabla p = \boldsymbol{f}_{\text{ext}}.$$

In the literature one often encounters another form, where the divergence of the deformation rate tensor is formally replaced by the Laplacian of the velocity. This transformation is justified in the case of an incompressible fluid flow with constant fluid viscosity. Using Lemma 2.1.2, it holds:

$$\nabla\cdot\left(2\mu\mathbb{D}\left(\boldsymbol{u}\right)\right) = \mu\left(\nabla\cdot\nabla\boldsymbol{u} + \nabla\left(\nabla\cdot\boldsymbol{u}\right)\right) = \mu\Delta\boldsymbol{u}. \tag{2.3.8}$$

This yields the **incompressible Navier–Stokes momentum equation in Laplacian form**

$$\rho\left[\frac{\partial \boldsymbol{u}}{\partial t} + \left(\boldsymbol{u}\cdot\nabla\right)\boldsymbol{u}\right] - \mu\Delta\boldsymbol{u} + \nabla p = \boldsymbol{f}_{\text{ext}}.$$

The incompressible Navier–Stokes momentum equation for homogeneous (constant density and viscosity) fluids can be written in a non-dimensional form by introducing a representative/characteristic velocity $\overline{u}$ [m/s] of the fluid and a representative/characteristic length $l$ [m] of the domain, see, e.g., [Joh16, Sect. 2.3]. Thereby, the dimensionless quantity

$$Re := \frac{\rho\,\overline{u}\,l}{\mu} \overset{(2.2.3)}{=} \frac{\overline{u}\,l}{\nu},$$

called *Reynolds number*, appears and reflects certain characteristics of fluid flows. The quantity $Re$ describes the ratio of inertial to viscous effects: Taking the order of magnitudes of the non-linear term $\left(\boldsymbol{u}\cdot\nabla\right)\boldsymbol{u}$, which is $\frac{\overline{u}^2}{l}$, and the viscous term $\frac{\mu}{\rho}\Delta\boldsymbol{u}$ respectively $\frac{1}{\rho}\nabla\cdot\left(\mu\mathbb{D}\left(\boldsymbol{u}\right)\right)$, which is $\frac{\mu\overline{u}}{\rho l^2}$, we immediately get the Reynolds number as their ratio. Hence, a very small Reynolds number ($Re \ll 1$) indicates that the viscous forces dominate the inertial effects (convective term) which is characteristic for a laminar flow, whereas a large Reynolds number ($Re \gg 1$) indicates that inertial effects dominate the viscous forces which is characteristic for turbulent flow. Thus, for high Reynolds numbers, almost inviscid flow is assumed and the Euler equation (2.3.6) can be used instead of the Navier–Stokes equation. In contrast to that, for steady flow, i.e., $\frac{\partial \boldsymbol{u}}{\partial t} = \boldsymbol{0}$, a low Reynolds number suggests to neglect the inertial effects represented by the non-linear convective term, i.e., to simplify the incompressible Navier–Stokes momentum equation to the then called **Stokes momentum equation**

$$-\nabla\cdot\left(2\mu\mathbb{D}(\boldsymbol{u})\right) + \nabla p = \boldsymbol{f}_{\text{ext}},$$

which is sometimes also referred to as the *creeping flow equation*. For constant viscosity, the incompressibility property $\nabla \cdot \boldsymbol{u} = 0$ can be used (as in equation (2.3.8)) to derive the **Stokes momentum equation in Laplacian form**

$$-\mu \Delta \boldsymbol{u} + \nabla p = \boldsymbol{f}_{\text{ext}}.$$

Usually, $\boldsymbol{f}_{\text{ext}}$ is abbreviated by $\boldsymbol{f}$ which will be done from here on as well.

### 2.3.2 The Darcy Equations

The Darcy equations build a system of partial differential equations that describes the laminar, incompressible, fully saturated flow of a Newtonian fluid through a porous medium at a macroscopic scale (which is large with respect to the pore scale):

$$\begin{aligned} \nabla p + \underbrace{\mu \mathbb{K}^{-1} \boldsymbol{u}}_{\substack{\text{Darcy} \\ \text{resistance}}} &= \boldsymbol{f} \quad \text{in } \boldsymbol{\Omega}, \\ \nabla \cdot \boldsymbol{u} &= 0 \quad \text{in } \boldsymbol{\Omega}. \end{aligned}$$

Here the velocity field $\boldsymbol{u}(\boldsymbol{x}) : \boldsymbol{\Omega} \to \mathbb{R}^n$ [m/s] and the pressure field $p(\boldsymbol{x}) : \boldsymbol{\Omega} \to \mathbb{R}$ [Pa] shall be considered as average velocity – also called *Darcy velocity* – and average pressure, respectively, $\boldsymbol{f}(\boldsymbol{x}) : \boldsymbol{\Omega} \to \mathbb{R}^n$ is an external force field [Pa/m], and the inverse of the permeability $\mathbb{K}^{-1}$ is sometimes denoted as resistance tensor, see [MPPM12].
Equipped with appropriate boundary conditions, this becomes a boundary value problem.

#### Experimental Derivation

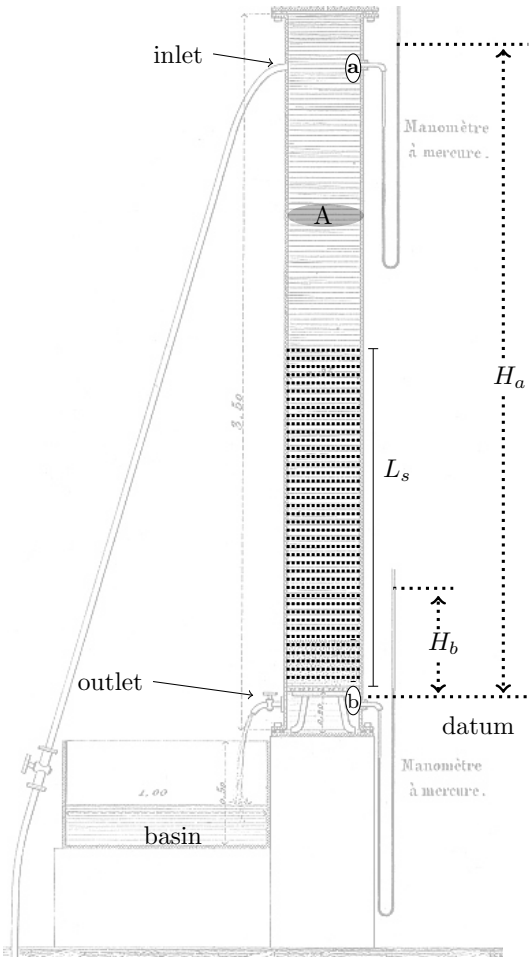
Henry Darcy derived the law named after him experimentally (empirically) as he was trying to find a way to supply large parts of the city of Dijon with fresh water via a system of fountains. In the Appendix D of his book [Dar56], published in 1856, the focus is on the description of the filtration rate (i.e., unidirectional flow) of water through sand beds.

**Remark 2.3.1** (On Manometers)**.**
*A manometer is an instrument which enables the measurement of pressure in a fluid. Darcy used a U-pipe manometer, partly filled with mercury, whose functionality is as follows. If a pressure is applied to one of the sides of the manometer, the mercury will transform due to the pursuit of equilibrium. The height difference of mercury together with one known pressure (e.g., atmospheric pressure) has to equal the unknown pressure (for negligible velocity heads). For more information on the functionality of manometers see [You+10].*

With a construction as shown in Figure 2.5, Darcy investigated the *volumetric flow rate* $Q$ [m³/s].

Original Experimental Setup:

A vertical, closed, cylindrical column of height 2.5 m with diameter 0.35 m is split into two parts by a horizontal filter at height 0.2 m.[a] The upper chamber is filled with a certain amount of siliceous sand resulting in a sand bed of thickness $L_s$ (with originally approximately 38% void space) and is fully saturated with water. The lower chamber contains only water since the filter on its top is permeable for water but keeps the sand outside. The upper chamber receives water from a water supply pipe where the flow can be moderated by a tab. The lower chamber can be opened by a tap and the outflowing water is kept and measured in a basin. Two mercury U-tube manometers are installed for measuring the water pressures at the two ends of the cylindrical column.




In each of Darcy's experiments, a certain pressure was established via the tap at the upper chamber and the resulting amount of water that had passed the filter (kept in the basin) during a certain time was noted. All pressures were measured relatively with respect to a datum located at the lower face of the filter.

[a] The measures in the experimental description (used also here) and those in the corresponding figure do not agree in [Dar56].

Figure 2.5: Experimental set-up based on [Dar56].

To that end he measured the volume of water, $V_{H^2O}$ [$m^3$], leaving the column through the outlet per time $t$ [s], in dependence of the thickness $L_s$ [m] of the sand bed and the difference of the *hydraulic/piezometric head*s $\Delta H := H_a - H_b$ [m]. The latter quantity can be considered as energy per unit weight and thus as a measure of water pressure in terms of height, [Bea72, Ch. 5]. The result of his experiments were the following two relations:

(I) The volumetric flow rate $Q$ is proportional to the difference between the two hydraulic heads of the water at the inlet tube and the outlet tube:

$$Q \propto \Delta \mathrm{H}. \tag{2.3.10}$$

(II) The volumetric flow rate $Q$ is inversely proportional to the thickness $L_s$ of the sand bed (in the direction of applied hydraulic head difference, i.e., the length of the filter path):

$$Q \propto \frac{1}{L_s}. \tag{2.3.11}$$

To summarize, Darcy observed the proportionality of the volumetric flow rate $Q$ and the dimensionless quantity $\frac{\Delta \mathrm{H}}{L_s}$ interpreted as *hydraulic head gradient*, which is representing a potential difference that causes water flow in this setting. Moreover:

(III) The volumetric flow rate $Q$ is proportional to the cross-sectional area $A$ [$\mathrm{m}^2$] of the sand bed (perpendicular to the direction of flow):

$$Q \propto A. \tag{2.3.12}$$

The combination of these three dependencies, (2.3.10)-(2.3.12), yields the equation

$$Q = K\, A\, \frac{\Delta \mathrm{H}}{L_s}, \tag{2.3.13}$$

where $K$ is a proportionality constant characteristic for the sand bed (the porous medium) and the water (the fluid).

The volumetric flow rate $Q$, i.e., the volume of filtered water per time $t$ can be determined using $V_{H^2O} = A \cdot L_{H^2O}$, with $L_{H^2O}$ being the vertical height of the filtered water volume in terms of the cylindrical apparatus, such that we can deduce

$$Q := \frac{V_{H^2O}}{t} = A\, \frac{L_{H^2O}}{t},$$

where the second factor has the units of velocity and is defined by

$$u_D := \frac{Q}{A} \tag{2.3.14}$$

and referred to as the *volumetric flux*, in the literature often denoted by *Darcy* velocity, *filter velocity*, or *specific discharge*. Note that this so-called Darcy velocity is not the exact velocity in a point-wise sense, but a macroscopic (volume-averaged) quantity. The microscopic velocity inside the sand grains and in dead end voids is actually zero, whereas the microscopic average velocity in the interconnected voids is in general larger than the Darcy velocity [Car97], [Kav02, Ch. 2.1]. The latter can be calculated using the porosity $\phi$ via the Dupuit–Forchheimer relationship

$$u_D^{void} := \frac{u_D}{\phi},$$

for a totally interconnected void and $\phi \neq 0$, see [NB06, Sect. 1.3], [Bea72, Ch. 5] and [Dis04, Ch. 1.1].

Inserting equation (2.3.14) into equation (2.3.13) yields

$$u_D = K \frac{\Delta \mathrm{H}}{L_s}. \tag{2.3.15}$$

By equation (2.3.15), the proportionality constant $K$ is the ratio of the Darcy velocity and the hydraulic gradient, and is referred to as *hydraulic conductivity* having the SI unit $[\mathrm{m/s}]$. On the one hand the flow rate $Q$ depends on the resistance that the porous medium exerts on any fluid while passing through, which is determined by the permeability $\mathbb{K}$ $\left[\mathrm{m}^2\right]$. On the other hand, the dynamic viscosity $\mu$ $[\mathrm{kg/ms}]$ and the density $\rho$ $[\mathrm{kg/m}^3]$ of the fluid have a direct impact on $Q$. These quantities are related by

$$\mathbb{K} = K \frac{\mu}{\rho g}, \tag{2.3.16}$$

where $g$ $[\mathrm{m/s}^2]$ is the acceleration due to gravity, see [Bea72, (5.5.1)] or [MW37, p. 72]. Based on equations (2.3.15) and (2.3.16), the permeability can be defined as the volume of a viscous fluid being filtered through a cross-sectional porous area of the column per time due to a hydraulic gradient.

The hydraulic/piezometric head is given as the sum of pressure head and elevation head:

$$H := \underbrace{\frac{p}{\rho g}}_{\text{pressure head}} + \underbrace{z}_{\text{elevation head}}. \tag{2.3.17}$$

Inserting (2.3.16) and (2.3.17) into equation (2.3.15) leads to

$$u_D = \frac{\mathbb{K}}{\mu} \left( \frac{p_a - p_b}{L_s} + \frac{\rho g \left( z_a - z_b \right)}{L_s} \right).$$

For $z_a \approx z_b$, the second summand might be neglected, such that equation (2.3.15) reduces to

$$u_D = \frac{\mathbb{K}}{\mu} \frac{\Delta p}{L_s},$$

with $\Delta p := p_a - p_b$. Otherwise, if non-horizontal flow is considered, the term corresponding to the elevation head does not vanish but could be absorbed in simple cases in a vector field $\boldsymbol{f}$. From now on we will skip the elevation head term (as it is often the case in the literature) keeping in mind that for non-horizontal flow, $p$ might not just represent the pressure but a scaled hydraulic head.

Up to now - following Darcy - we have considered only one-dimensional flow. In fact, Darcy's law has been generalized to arbitrary homogeneous, viscous fluids and to three spatial dimensions by [MW37, Ch. 2 and Sect. 3.3]. In order to obtain a three-dimensional velocity corresponding to the coordinate system $(x, y, z)$, the direction of the velocity has to be taken into account. As water

flows from higher heads to lower heads, the quantity $u_D$ has to be interpreted here as the velocity contribution in $(-z)$-direction, meaning that the actually three–dimensional Darcy velocity is given by

$$\boldsymbol{u}_D = \begin{pmatrix} 0 \\ 0 \\ -u_D \end{pmatrix}.$$

Formally, the generalization of Darcy's law to higher dimensions and in differential form reads

$$\boldsymbol{u}_D = \frac{\mathbb{K}}{\mu}\left(-\nabla p + \boldsymbol{f}\right),$$

where $\boldsymbol{f}$ represents external forces that were not considered in Darcy's original experiments. Since the gradient points in the direction of steepest increase, the minus indicates that, in absence of external forces, the direction of flow is that of decreasing pressure. Note that all available information on the possibly complicated porous structure is 'captured' in the permeability tensor $\mathbb{K}$.

**Remark 2.3.2.** *For $\boldsymbol{f} = \boldsymbol{0}$, the Darcy equation states that if $\mu\mathbb{K}^{-1}$ is scalar (isotropic porous medium), then $-\nabla p$ and $\boldsymbol{u}_D$ are linearly dependent (parallel) vectors, and if $\mu\mathbb{K}^{-1}$ is of (second-order) tensorial character (anisotropic porous medium), then $\boldsymbol{u}_D$ is the result of a linear transformation applied to $-\nabla p$.*

In absence of external forces, the resistance to flow of a motion that obeys the Darcy equations is totally determined by the bulk stress resistance. This is a consequence of the experimental setup of Figure 2.5 in which the internal surface area (interstitial surface) was strongly dominating the area of the boundary surfaces in terms of size.

Subsequently, the index $D$ of the Darcy velocity will usually be omitted.

**Remark 2.3.3** (Formal Derivations of the Darcy Equations)**.**
*Apart from experimental confirmations, there is also a natural interest in theoretical derivations of Darcy's law. Let us mention two approaches based on the microscopic description of the flow using the Stokes equations in the pore space with no-slip boundary conditions at the porous structure in the interior.*

- *Volume/Space Averaging:*
  *More than a century had passed after Darcy's experimental reasoning, when a mathematical derivation of Darcy's law was given in [Whi86]. Using the technique of (representative) volume/space averaging, Whitaker derived the Darcy equations from the incompressible Stokes equations* (2.3.1)*. The basic idea here is to consider the microscopic equations with no-slip boundary conditions in the interconnected void, take its average over a representative elementary volume (a volume that represents the local average properties), and derive a problem for the averaged velocity and the averaged pressure, which has in this case revealed Darcy's law. For more details see also [Kav02, Sect. 2.7].*

- *(Periodic) Homogenization:*
  *The central idea of homogenization is to "upscale" differential equations in order to overcome a large discrepancy between (length) scales of interest, e.g., a micro/pore scale and a macro/problem scale, respectively to smoothen rapidly oscillating coefficients. To that end, multi-scale expansions (involving the fluids behavior on all scales of interest) are inserted into the microscopic model, with the attempt to let the pore/micro-scale vanish (tend to zero), [EFM00], [Mik00]. In more detail, instead of one pair* $(\boldsymbol{u}, p)$*, one considers a family of pairs* $(\boldsymbol{u}^\varepsilon, p^\varepsilon)$ *with* $\varepsilon > 0$ *being a spatial scale parameter (the pore length scale divided by the system length scale), [Kav02, Sect. 2.8]. Then one determines the limit*

$$(\boldsymbol{u}, p) = \lim_{\varepsilon \to 0} (\boldsymbol{u}^\varepsilon, p^\varepsilon)\,, \tag{2.3.18}$$

  *which is considered as the result of the "upscaling" procedure, since* $\varepsilon \to 0$ *corresponds to the transition from micro-scales to macro-scales. Hence, with the homogenization technique, one aims to find differential equations whose solution is the limit* (2.3.18)*. Assuming a periodic structure of the medium, the steady, incompressible Stokes equations for a viscous fluid in a porous medium with no-slip boundary condition at the solid obstacles yield Darcy's law as a homogenized (macro-scale) model, for details see [Hor97, Sects. 1, 3, 4], [All91a], [All91b], [All90b], [All90b], and [AL06]. The method of periodic homogenization yields also the proof of the symmetry and the tensorial character of the permeability* $\mathbb{K}$*, see [Kav02, Sect. 2.8].*

### Fields of Application and Limitations of the Darcy Equations

The Darcy equations, also known as the groundwater flow equations, are one of the bases of hydro-geology and therefore widely used in soil mechanics. Further applications include the modeling of biological and biomechanical systems, [KV03].

Whereas Darcy's law describes the viscous, laminar flow at low velocity correctly, the discrepancy between experimental results and theoretical results obtained using Darcy's law increases for increasing velocity, [IS99, Sect. 1.1.1], [Kav02, Ch. 2.1]. This means, that for higher permeabilities, Darcy's law might fail to capture the correct flow behavior. The range of validity is typically discussed in terms of the (dimensionless) Reynolds number for porous media

$$Re_p := \frac{|\overline{\boldsymbol{u}_D}|d}{\nu}, \tag{2.3.19}$$

where, $|\overline{\boldsymbol{u}_D}|$ is a characteristic velocity (e.g., the average or a suitable norm), $\nu$ is the kinematic fluid viscosity, and $d$ is a characteristic length, e.g., related to the grain-size distribution (mean grain diameter) or chosen equal to $\sqrt{\mathbb{K}}$, see [ZG06], [IS99, Sect. 1.1.1], [Kav02], [Bea72], and [Gre81]. Similar to the scenarios discussed for the Reynolds number corresponding to the Navier–Stokes

equations in Subsection 2.3.1, the viscous forces dominate the inertial forces for small values of $Re_p$, while the opposite is indicated by large values of $Re_p$. In the majority of cases, the upper bound on the Reynolds number for the validity of the Darcy equation is estimated somewhere between one and ten, see, e.g., [Bea72], [ST14], [Suk$^+$13], [Ham94], [NB06, Sect. 1.5.2.], and [Vaf00, Ch. 7.2.1.1]. For an overview on related literature, see [ST14]. For alternative possibilities to identify non-Darcy flow see, e.g., [ZG06].

Furthermore, the resistance of solid walls (no-slip boundary condition) has a serious impact on the flow behavior in many practical applications, at least in a neighborhood of the respective boundary, but is incompatible with the Darcy equations, [Nis$^+$18]. This issue will be discussed in detail in Chapter 3.

### The (Navier–)Stokes–Darcy Coupling

The restricted validity of the Darcy equations to low Reynolds number flows leads to the consequence that fluid flow in mixed domains corresponding to multi-physics problems cannot be treated by the Darcy equation only. This includes, e.g., domains containing a totally or almost pure flow region (no bulk resistance or a very high permeability) next to a porous medium with moderate or low permeability. Realistic examples for such a setting are riverbeds, or more involved, karst aquifers that consist of a porous medium and a network of conduits build by cracks and fissures, whose detailed description is particularly important for the transport of fluid, [CGW10], having in mind, e.g., environmental problems as the groundwater contamination through rivers. In general, transport through fractured (in geology sometimes called *vuggy*) porous media or the simulation of perfusion in biological tissue pose interesting and demanding challenges in current multi-physics research, [GKR13].

With the models discussed so far, an intuitive idea to tackle such problems is to couple the Darcy equations with the (incompressible Navier–)Stokes equations via interface conditions. Indeed this is a widely considered treatment studied in, e.g., [Wil19], [LSY02], [VY09], [DZ09], [MS17], [Dis04], [DQ09], [CWW16], [Dis05], [DMQ02] , [JM96], [Ang11], [Kav02, Ch. 2.11], [NN74], in particular for applications related to hemodynamics [DZ11], oil filters [IL04], and fuel cells [Ehr$^+$08].

The coupling approach starts with decomposing the domain into a pure fluid flow subdomain (Stokes region) and a porous region (Darcy region)[4], such that each subdomain has a sufficiently smooth boundary (usually Lipschitz continuous). This results in a heterogeneous domain decomposition problem, i.e., the differential operators of the models used in the subdomains differ. The hypersurface[5] that separates the subdomains is referred to as *interface*. For modeling the inter-subdomain connection, i.e., the transport of information between the subdomains, appropriate coupling conditions have to be identified and applied across and along the interface. The continuity of the normal velocity

[4] In general, one might consider more than two subdomains.
[5] This is meaningful only for non-overlapping decompositions.

across the interface (originating from the conservation of mass equations) and the balance of the normal components of the stress forces acting on the interface are well accepted, see, e.g., [CJW14], [BN13], or [LSY02]. In order to close the system concerning the Stokes region, an additional interface condition is necessary, which describes the tangential component of the velocity in the pure fluid region at the interface. In fact there is a migration of fluid tangent to the interface within the porous region (near the interface), i.e., the viscous shear effect in the Stokes region penetrates into the Darcy region creating a so-called boundary layer, see experimental results in [BJ67]. Although the tangential velocity of the non-obstructed (Newtonian) fluid at the permeable interface (called *slip velocity*) can be considerably larger than the Darcy velocity within the porous mass, a no-slip condition on the tangential Stokes velocity is not adequate. In order to account for this phenomenon, a more satisfactory relation is given by the Beavers–Joseph condition, which is a semi-empirically derived slip-flow interface condition introduced and experimentally backed (for steady-state, laminar, unidirectional, interface-parallel channel flow) in [BJ67] and [BSM70]. The Beavers–Joseph interface condition is widely used for (Navier–)Stokes–Darcy couplings and replaces the (incorrect) classical interface condition which sets the tangential Stokes velocity at the interface equal to zero ([JT64]). The new condition is able to take into account steep gradients of the velocity in a small interface layer by introducing a (macroscopic) velocity jump. The basis therefore is the assumption of proportionality between the tangential component of the stress that the Stokes fluid exerts on the interface and the jump in the tangential velocity across the interface (discontinuity in the tangential velocity). Formally, the condition can be expressed as

$$\boldsymbol{t}_i \cdot (\mathbb{S}_S \cdot \boldsymbol{n}) = \alpha_{BJ} \frac{\mu}{\sqrt{\mathbb{K}}} (\boldsymbol{u}_S - \boldsymbol{u}_D) \cdot \boldsymbol{t}_i \quad \text{on } \Gamma_I, \forall i, \tag{2.3.20}$$

see, e.g., [EFL12], [Jon73], [GKR13], [SAD94], [CGW10], [Cao$^+$10], [Aur09a], [NB06], and [BJ67]. Here $\boldsymbol{n}$ is the outward pointing unit normal with respect to the Stokes region, $\mathbb{S}_S \cdot \boldsymbol{n}$ is the total stress exerted by the Stokes fluid on the interface, $\boldsymbol{u}_S$ is the Stokes velocity, $\boldsymbol{u}_D$ is the Darcy velocity, $\{\boldsymbol{t}_i\}_{i=1}^{n-1}$ is an orthonormal basis of the tangential space on the interface $\Gamma_I$, $\mathbb{K}$ is the scalar permeability of the porous region, and $\alpha_{BJ}$ is called (dimensionless) *Beavers–Joseph constant* or *slip coefficient*. From the physical point of view, the Darcy velocity $\boldsymbol{u}_D$ is evaluated at some small distance from the interface, such that the transition between the Darcy and the Stokes flow takes place in a thin (microscopic) layer inside the Darcy region which is not resolved by the Darcy equations. The constant $\alpha_{BJ}$ has to be heuristically chosen and was shown to depend in particular on the direction of flow at the interface and the Reynolds number, [NB06, Sect. 1.6]. The general form (2.3.20) is often used, but note that due to the orthogonality $\boldsymbol{t}_i^T \boldsymbol{n} = 0$, $\forall i$, the part of the stress that is related to the pressure $p$ is actually not present, such that only the viscous stress contributes to the Beavers–Joseph condition, i.e., for incompressible flow we actually have $\boldsymbol{t}_i \cdot (\mathbb{S}_S \cdot \boldsymbol{n}) = \boldsymbol{t}_i \cdot (2\mu \mathbb{D}(\boldsymbol{u}) \cdot \boldsymbol{n})$ (in accordance with the original form in [BJ67]). For a visualization of the original experimental setup, for that

the 1D Beavers-Joseph condition

$$\frac{\partial u_S}{\partial y} = \frac{\alpha_{BJ}}{\sqrt{\mathbb{K}}} \left( u_S - u_D \right)$$

was originally derived, see Figure 2.6.

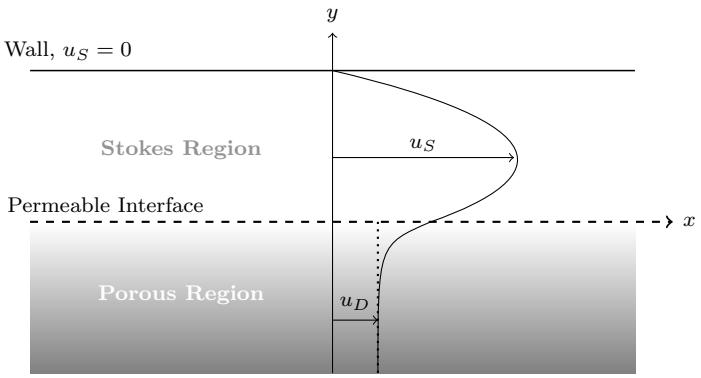



Figure 2.6: Velocity profile for unidirectional flow along a porous-fluid interface; solid line refers to the realistic velocity profile; dotted line in the porous region corresponds to the Darcy velocity according to the use of the Beavers–Joseph interface condition.

Although the condition (2.3.20) might describe the flow in the pure fluid region well (for appropriate values of $\alpha_{BJ}$), it fails describing the flow inside the porous region near the interface appropriately, [NN74]. The boundary layer in the porous (Darcy) region cannot be described due to the absence of a viscous shear term.

A popular simplification called the Beavers–Joseph–Saffman condition [Saf71] is obtained by noting that $u_D = \mathcal{O}(\mathbb{K})$, such that it can be neglected compared to $u_S$. The resulting condition does not depend on the Darcy velocity anymore and can be considered as a boundary condition for the Stokes problem. In [CGW10] several coupling conditions for the Stokes–Darcy problem have been validated with respect to a reference solution obtained using a Stokes–Brinkman coupling (see the following Subsection 2.3.3). The Beavers–Joseph–Saffman condition was theoretically justified in [JM00] (asymptotically, for vanishing pore size) via the method of homogenization.

The appropriate coupling of pure and porous flow is still a topic of general interest and thus under research. For example in [LBW06], a transition zone near the interface inside the porous domain is introduced, where the Stokes equation is applied. This approach is concluded to yield a (slightly) better coincidence with the experimental values of Beavers and Joseph, compared with the Beavers–Joseph interface condition itself. For more information on compatibility issues concerning fluid-porous couplings see, e.g., [NB06].

#### Extensions of Darcy's law

Coupling equations modeling different phenomena is a possibility to describe complex multi-physics problems, for which a good knowledge of the setting under consideration is obligatory. However, the constraints on the applicability of the models restrict the toolbox. For non-Darcian regimes, e.g., those with rather high permeabilities and/or filter velocities, several generalizations of the Darcy equations have been proposed, taking into account additional phenomena as viscous (important at low velocities) or inertial effects (convection, important at high velocities), [Kav02, Ch. 2.5]. An overview on different generalizations of Darcy's law is given in [Ham94].

Among the most popular generalizations let us mention the Darcy–Forchheimer equation ([Vaf00, Ch. 3.5], [NB06, Sect. 1.5.2], [ZG06], [Kav02, Ch. 2.9])

$$-\nabla p = \mu \mathbb{K}^{-1} \boldsymbol{u} + \left( c_F \rho \mathbb{K}^{-\frac{1}{2}} |\boldsymbol{u}| \right) \boldsymbol{u},$$

where $c_F$ is called *form drag coefficient* and depends on the porous medium. The additional term here accounts for (microscopic) inertial effects based on the kinetic energy [ST14]. According to [GA94, p. 355], the incorporation of the Forchheimer drag term is recommended whenever the Reynolds number (equation (2.3.19)) is larger than one. For more information, in particular concerning the specification of model-appropriate fluid flow regimes, see, e.g., [Vaf00, Ch. 3.5] and references therein.

Moreover, the Darcy law may become inaccurate whenever shear induced momentum transfer due to boundary effects becomes important. This phenomenon is accounted for by the Brinkman model, another extension of the Darcy model, which will be discussed in detail in the next Subsection 2.3.3.

### 2.3.3 The Brinkman Equations

The Brinkman equations can be seen as a generalization of Darcy's equations given by

$$-\nabla \cdot (\mu_{\text{eff}} \mathbb{D}(\boldsymbol{u})) + \nabla p + \mu \mathbb{K}^{-1} \boldsymbol{u} = \boldsymbol{f} \quad \text{in } \boldsymbol{\Omega}, \tag{2.3.21a}$$

$$\nabla \cdot \boldsymbol{u} = 0 \quad \text{in } \boldsymbol{\Omega}. \tag{2.3.21b}$$

Here, $\boldsymbol{u}(\boldsymbol{x}) : \boldsymbol{\Omega} \to \mathbb{R}^n$ denotes an averaged velocity [m/s], $p(\boldsymbol{x}) : \boldsymbol{\Omega} \to \mathbb{R}$ is an average pressure [Pa], $\boldsymbol{f}(\boldsymbol{x}) : \boldsymbol{\Omega} \to \mathbb{R}^n$ an external force field [Pa/m], $\mathbb{K}$ the permeability, and $\mu_{\text{eff}}$ is the effective viscosity, see Remark 2.3.5. Depending on the values of the physical coefficients, $\mu_{\text{eff}}$ and $\mu \mathbb{K}^{-1}$, the Brinkman momentum and mass conservation equations define the whole range of intermediate models between the Stokes and the Darcy equations, and can be used to model, e.g., homogeneous, incompressible fluid flow through a porous medium with fractures, [PEQ09] (with $\mu_{\text{eff}} > 0$). Equipped with appropriate boundary conditions, the system (2.3.21) becomes a boundary value problem.

In case of a constant effective viscosity $\mu_{\text{eff}}$, the system (2.3.21) is equivalent to

$$-\mu_{\text{eff}}\Delta \boldsymbol{u} + \nabla p + \mu \mathbb{K}^{-1}\boldsymbol{u} = \boldsymbol{f} \quad \text{in } \boldsymbol{\Omega}, \tag{2.3.22a}$$

$$\nabla \cdot \boldsymbol{u} = 0 \quad \text{in } \boldsymbol{\Omega}, \tag{2.3.22b}$$

equipped with appropriate boundary conditions.

In the literature, the Brinkman equations (2.3.21) respectively (2.3.22) are also referred to as generalized Stokes equations [DZ09], a certain Oseen problem [Joh16], Stokes–Darcy problem [BH05], Stokes–Brinkman equations [Ing11], or even Darcy–Brinkman equations [LPN07].

**Remark 2.3.4** (On the Brinkman Coefficients)**.**
*From the point of view of physics, all the coefficients in the Brinkman problem are non-negative in the whole domain and* $\mathbb{K} \not\equiv 0$ *which physically means to allow flow through the porous medium. Moreover, for isothermal, single-phase flow of a homogeneous fluid, it is convenient to assume that the viscosity* $\mu$ *does not depend on the spatial variable.*

### Heuristic Derivation

Brinkman, [Bri49], observed the fact that Darcy's law does not account for viscous shearing stresses acting on a volume element of fluid, resulting in good approximations in case of small permeabilities only. His idea therefore was to develop a model which incorporates the Stokes drag force on a spherical particle moving through the plain fluid as well as the effect of neighboring spheres. In the case that the fluid flow involves for example shear flow in a neighborhood of the boundary of highly porous regions, velocity gradients are inevitable. Brinkman's objective was to find a formula which describes all phenomena such that in the case of high as well as low permeabilities $\mathbb{K}$, the validity of the model is guaranteed. Therefore he considered the forces exerted by slow, steady, incompressible flow of a Newtonian fluid through a swarm of spherical particles and proposed an equation describing the equilibrium between all the forces acting on a fluid volume, namely the pressure gradient, the damping force exerted by the porous structure and the divergence of the viscous stress tensor. The result is the Brinkman momentum equation[6]

$$\underbrace{\nabla p}_{\substack{\text{pressure gradient}}} = \underbrace{-\mu \mathbb{K}^{-1}\boldsymbol{u}}_{\substack{\text{damping force/}\\ \text{bulk resistance}}} + \underbrace{\mu_{\text{eff}}\Delta \boldsymbol{u}.}_{\substack{\text{divergence of the}\\ \text{viscous stress tensor}}} \tag{2.3.23}$$

Due to the presence of the viscous stress in the Brinkman equation, a flow in a porous medium can be described also in situations, where the velocity gradients cannot be neglected. Hence, in particular the impact of boundary effects on the

[6] Literature indicates that P. Debye proposed the equation (2.3.23) with $\mu_{\text{eff}} = \mu$ in the same year, see, e.g., [Ham94] and references therein.

velocity of the fluid is included.

Equation (2.3.23) can be interpreted as a superposition of bulk and boundary effects - of Darcy's and Stokes' law, respectively.
In case of small permeabilities, $\mathbb{K} \to 0$, (2.3.23) is an approximation of Darcy's law (*Darcy limit*), since then the *bulk resistance* (also called *damping force*) dominates the boundary effects based on external surface shear stresses (divergence of the viscous stress tensor), [Kav02, Ch. 2.6].
For the opposite asymptote, $\mathbb{K} \to \infty$, the bulk resistance vanishes such that only external surface shear stress is left. Hence, for high permeabilities $\mathbb{K}$, the Brinkman equation (2.3.23) approximates the momentum equation of the Stokes equations (*Stokes limit*).

**Remark 2.3.5** (On the Effective Viscosity)**.**
*Brinkman [Bri49] noticed a difference between $\mu_{\text{eff}}$ and the fluid viscosity $\mu$. It is apparent that the effective viscosity $\mu_{\text{eff}}$ takes into account the porous mass in some sense as well as the fluid viscosity, i.e., $\mu_{\text{eff}} = \mu_{\text{eff}}(\mu, \phi)$ (see [Kav02, Ch. 2.11.5, 2.11.6] and references therein). On the one hand, Brinkman suggested to use the Einstein formula for dilute (i.e., high porosity $\phi$) suspensions of spheres* (2.2.4) *for the approximation of $\mu_{\text{eff}}$. On the other hand, he mentioned the argument that a column packed by particles might yield the situation that the particles do either not contribute to, or even hinder the transport of momentum in the fluid. Thus, the effective viscosity might be smaller (which is excluded in* (2.2.4)*) or larger than the fluid viscosity, depending on the type of the porous medium (see [NB06, Sect. 1.5.3],[Vaf00, pp.110], [Aur09b], [Mut15, Anhang V], [Kop83], [Kav02, Ch. 2.11.5], [Lun72], and references therein for different possibilities to define $\mu_{\text{eff}}$). However, without further investigations, Brinkman decided to set $\mu_{\text{eff}} = \mu$ as many authors after him.*

*In the literature, there is no general consensus concerning the practical definition of the effective viscosity. For example, the choice*

$$\mu_{\text{eff}} = \frac{\mu}{\phi},$$

*which implies $\phi \neq 0$ and implies $\mu_{\text{eff}} \geqslant \mu$, can be found, e.g., in [Lun72], [LZT06], and [OMB70]. In other sources, the relation*

$$\frac{\mu_{\text{eff}}}{\mu} = \frac{\tau}{\phi},$$

*where $\tau = \tau(\phi)$ denotes the tortuosity of the medium (i.e., the resistance to flow diffusion due to local boundaries and local viscosity) was proposed for isotropic porous media, [KV03]. According to [Ang99] for a porous medium, the ratio $\frac{\mu_{\text{eff}}}{\mu}$ is in general contained in [0.1,2] and in [LPN07] a range of even up to ten (for high porosity media) is mentioned.*

*The authors of [LPN07] study appropriate choices of $\mu_{\text{eff}}$ concerning the effect of a solid boundary on the flow through cylindrical porous media, and report that the Brinkman equation with $\mu_{\text{eff}} \neq \mu$ does not describe the behavior at solid walls appropriately. Therefore, they propose to use a thin porous Brinkman layer with $\mu_{\text{eff}} = \mu$ near the solid boundary in order treat the no-slip boundary condition.*

*For a further discussion on the meaning of the effective viscosity and its relation to $\mu$, the reader is referred to [Pn+17] and references therein.*

**Remark 2.3.6** (Formal Derivations of the Brinkman Equations)**.**
*As in Remark 2.3.3, different possibilities to theoretically derive Brinkman's equations are noted.*

- *Statistical approach:*
  *Lundgren presented in [Lun72] a theoretical justification of the Brinkman equations based on the method of ensemble averaging. Probabilistic methods have been adopted as well in [Rub86] (assuming high porosities) and [Tam69].*
- *Local Volume Averaging:*
  *The Brinkman equations also have been derived via local space averaging, see [OTW95a] and [Vaf00, Ch. 3.5]. In [OTW95a], also some insight into defining the average velocity and the average pressure can be gained.*
- *(Periodic) Homogenization:*
  *The Brinkman equations can be asymptotically derived from the Navier–Stokes equations by using the technique of homogenization for high porosities, see [Hor97, pp. 66], [All90a], [All90b], [All91a], [All91b], and [GAB05].*
- *Time Discretization of the (Navier–)Stokes equations:*
  *The Brinkman model appears in time-stepping methods for Stokes and for high Reynolds number flows. For example, implicit time discretization of the incompressible unsteady Stokes equations leads to the Brinkman-type problem*

$$\begin{aligned} -\varepsilon^2 \Delta \boldsymbol{u} + \nabla \hat{p} + \boldsymbol{u} &= \boldsymbol{f}, \\ \nabla \cdot \boldsymbol{u} &= 0, \end{aligned}$$

  *with $\varepsilon^2 := \mu \Delta t$, $\Delta t$ being the (uniform) time step, $\hat{p} := \frac{\Delta t}{\rho} p$, and appropriately redefined $\boldsymbol{f}$. For the non-linear case, see also [Sog14] and [Joh16, Ch. 5].*

### On the Validity of the Brinkman Equations and Fields of Application

The Brinkman(–Forchheimer) equations were experimentally backed in, e.g., [GA94] for flow through a wall-bounded open foam with porosity $\phi = 0.972$, for which the authors obtained the relation $\mu_{\text{eff}} = 7.5\mu$. The second order velocity term was identified to be an important extension to Darcy's law in presence of an interface as, e.g., a wall or a pure flow region neighboring a porous region, where this Brinkman term allows to predict the behavior of the flow near this interface (boundary layers), see also [CGW10].

The fields of application associated with the Brinkman model are numerous and versatile. Some examples include:

- Fluid flows through highly porous media and such with distinctive heterogeneities were modeled by the Brinkman model in [ILW11] and [KLM17].

- The Brinkman equations were used to describe flow through natural vuggy reservoirs in [MWY14] and [ILW11], for the simulation of water flow and tracer transport in heterogeneous karstic aquifers in [Joo$^+$09], and the backward erosion of soils in [FMS16]. In [KS12] it was remarked that the effects of taking the viscous stress term into account are most significant in the presence of large cracks or vugs, which is a typical phenomenon in oil reservoirs. Since the effects of solid boundaries are assumed to be most powerful in a neighborhood of the boundary, the Brinkman equation has been sometimes used for modeling the flow in a thin boundary layer of thickness $\left(\mu_{\text{eff}}\mathbb{K}\mu^{-1}\right)^{\frac{1}{2}}$, [LPN07], [NB06, Sect. 1.5]. Similarly, a Brinkman layer between a free surface flow and a porous Darcy flow has been considered in [EFL12], [Ehr$^+$08] concerning fuel cells and for modeling groundwater flow problems in [Das02].
- Moreover, the Brinkman model was used in biomedicine [KV03], in connection with tissue generation in bio-reactors [Pod$^+$14], for modeling blood flow in biological tissue [Nis$^+$18] and [KV03], in particular Casson fluid flow in porous media as blood flow with fatty plaques and clots [DMJ96], for interstitial flow (fluid flow through tissue) and its impact on muscle cells [TT00], and for fluid flow through osteons (part of cortical bones) [MR09].
- The Brinkman equations and generalized forms of it have been also used to model cooling of electronic equipment via convective flows through metal foam-filled pipes [LZT06], and heat and mass transfer in pipes [KG06], [LYT89].
- Flows in highly porous media, e.g., industrial filters (porosity $\geqslant 0.9$) as oil filters [IL04], glass or mineral wool based thermal insulators (porosity 0.99) [ILW09], and open foams (porosity $\geqslant 0.95$) [ILW09] were simulated.
- Also liquid composite moulding [GK10], [Pn$^+$17] and gas diffusion through fuel cell membranes [GLK98] were treated as well.
- Applications, where the porous structure is too complicated for a microscopic (Navier–)Stokes model with no-slip boundary conditions at the solid obstacle boundaries and the flow is non-Darcy in the sense that external viscous shear effects cannot be neglected, include wind farms with closely placed turbines and gas flow in pebble bed nuclear reactors, [Ing11].
- The Brinkman equations were also used for moderate Reynolds number flow around complex geometries and moving or deforming solid boundaries via a penalization approach in order to circumvent body-fitted meshes, [ABF99], [KG01]. Therefore geometric restrictions are indirectly enforced by penalizing, e.g., fluid flow in solid obstacles (volume penalization) or at solid boundaries (surface penalization). Instead of considering for example the Stokes equations in a complicated domain with obstacles, a Brinkman-type problem is posed in a much simpler domain by penalizing velocities appropriately. This approach is strongly related to the *fictitious domain method* (also *immersed boundary approach*). A downside here is the strong variation in coefficients, which might yield an ill-conditioned discrete problem due to significant discontinuities of the coefficients, [KLM17], and

thus necessitates local mesh refinement. The fictitious domain method on the continuous level has been extensively analyzed in [Ang99] for fluid-porous-solid systems by using spatially varying (discontinuous) coefficients (permeability and effective viscosity) which then results in a discrete problem posed in simple (fictitious) domains.

The attempt to identify the domain of validity of the Brinkman model has been a frequent point of discussion:

- In [Tam69, Sect. 6] it is noted, that $\mu\mathbb{K}^{-1}\boldsymbol{u}$ dominates $\mu_{\text{eff}}\Delta\boldsymbol{u}$ (i.e., the Brinkman equation can be replaced by the Darcy equation) whenever $\left(\mu_{\text{eff}}\mu^{-1}\mathbb{K}\right)^{\frac{1}{2}}$ is much smaller than the length scale of interest.
- The authors in [Kro$^+$11] performed numerical simulations of carbonate karst reservoirs in order to identify a Brinkman region in a homogeneous porous medium and concluded it to be defined by $10^{-4} \leqslant \frac{\mathbb{K}}{L^2} \leqslant 10$, where $L$ is a characteristic length.
- According to [Lun72], a porosity $\phi > 0.6$ is required for the validity of the Brinkman model, see also [Nie83, p. 38]. In [DB87] the validity of the Brinkman equations with $\mu_{\text{eff}} = \mu$ for porosities $\geqslant 0.95$ is concluded.
- Nield and Bejan argue in [NB06] that the effective viscosity term should be included if the porosity is close to one or if the flow is near a solid (impermeable) boundary.
- According to [Aur09b], it is admitted that for $\phi \to 1$, the effective viscosity $\mu_{\text{eff}}$ tends to the fluid viscosity $\mu$. In this text, the domain of validity of the Brinkman equations is addressed by considering different separations of scales via the parameter

$$\varepsilon := \frac{\text{pore scale}}{\text{system scale}}$$

for different kinds of porous media. Altogether, Auriault's discussion can be summarized as giving the tendency that for $\varepsilon \ll 1$ (very good separation of scales) the Darcy equations, for $\varepsilon < 1$ the Brinkman equations, and for $\varepsilon \geqslant 1$ (no separation of scales) the (Navier–)Stokes equations should be considered. This can be restated as the pressure gradient balancing the filter velocity for rather large system length scales and the pressure gradient balancing the Laplacian of the velocity for rather small system length scales, see also [DB87].

### The (Navier–)Stokes–Brinkman and Brinkman–Darcy Couplings

A coupling of plain and porous domains can be managed by using the Beavers–Joseph(–Saffman) condition for Brinkman–Darcy couplings similar as for a Stokes–Darcy interface. Note that the slip coefficient $\alpha_{BJ}$ in general has to be determined experimentally or via an auxiliary problem, see [JM00].

Due to the presence of the shear stress term, the Brinkman equation with $\mu_{\text{eff}} \neq 0$ is compatible with boundary and interface layers. Therefore, the

Brinkman equation (with possibly varying effective viscosity, see [Kav02, Sect. 2.11.6]) in contrast to the Darcy equation has the potential to describe the boundary layer in the immediate porous neighborhood of a permeable interface in which the velocity can change very rapidly. This suggests a replacement of the Darcy equations in certain coupling approaches.

The coupling of the (Navier–)Stokes equations with the Brinkman equations (instead of the Darcy equations) has been studied, e.g., in [CGW10], [DZ09], [IL04], [Ang11], [Ehr+08], [OTW95a], [OTW95b], and [HA10]. In this situation, the order of the differential operators does not change across the interface as long as $\mu_{\text{eff}} \neq 0$ - contrary to the Stokes–Darcy coupling. Therefore, the fluid-porous interface conditions are much simpler to address here, since one can pose the condition of continuous velocity and stress across the interface using $\mu_{\text{eff}}$, see, e.g., [GK10, Ch. 3.1], [SAD94, Sect. 2.3.2], [CGW10], or [IL04]. As mentioned before, boundary layers in the porous domain - although the layer thickness is usually of order $\sqrt{\mathbb{K}}$ only [Saf71] - might have a significant impact on external flows, [NN74].

In [NN74], the phenomenological Beavers–Joseph interface condition (slip flow) was related to the Brinkman equation in the following sense. Considering unidirectional, porous-plain, interface-parallel channel flow and comparing a Stokes–Brinkman coupling with a Stokes–Darcy coupling revealed that the velocity field in the plain subdomain (Stokes) coincides if the Stokes–Darcy coupling uses the Beavers-Joseph condition with $\alpha_{BJ} = \sqrt{\mu_{\text{eff}}\mu^{-1}}$.

Also in [SAD94], a comparison, this time between a Navier–Stokes–Darcy coupling and a Navier–Stokes–Brinkman coupling (with $\mu_{\text{eff}} = \mu$) was undertaken in the finite element context. They conclude that the use of the Brinkman equation inside the porous domain instead of the Darcy equation simplifies the implementation of the interface conditions and allows to describe the interface layer inside the porous medium. The latter however necessitates an appropriate resolution of the boundary layer by the mesh in order to prevent numerical instabilities (oscillations). However, this holds in general, namely the mesh in the neighborhood of the interface has to be fine enough to resolve rapidly changing flow characteristics, [Kav02, Sect. 2.11.2].

### The Brinkman Problem as a Single Domain Approach

Besides the identification of appropriate interface conditions matching differential operators of different order - which might pose difficult questions already on the continuous level - there are several further challenges concerned with coupling approaches in the discrete case. For domain decomposition approaches, the detailed location of the interface, and its analytic description have to be known a priori, such that appropriate models can be chosen and assigned to the sub-regions. This might be an awkward task, since the ranges of validity of porous media flow models are not clearly distinguished as we have seen. Further, the interface should be rather simple since this decides about the necessary computational effort, in particular if one wishes to use a fitted mesh.

This complex of requirements raises the interest in alternative ways to treat such a situation. In fact, the Brinkman model could be guessed of as in principle allowing a single-domain approach for modeling multi-domain problems by using one single set of equations with varying coefficients $\mu_{\text{eff}}(\boldsymbol{x})$ and $\left(\mu\mathbb{K}^{-1}\right)(\boldsymbol{x})$. These have to be defined in the whole domain[7], but in contrast to a coupling approach, an analytic description of the interface and its spatial approximation are not required. As we have mentioned in the previous paragraph, the mesh has to be fine enough to resolve the flow characteristics of interest as it is the case in general.

Later on we will analyze a discrete method which is appropriately designed to be applicable for a single-domain approach, but the construction and analysis of this unified discrete setting is nontrivial and will play a key role in Chapters 4 and 5.

#### Extensions of Brinkman's law

Let us finally note that there are also generalizations of the Brinkman equations available, which take into account additional dynamic fluid flow effects. For example, the steady (Darcy–)Brinkman–Forchheimer equation [Vaf00, Ch. 3.5], [Ham94]:

$$-\nabla p = \mu\mathbb{K}^{-1}\boldsymbol{u} + \left(c_F\rho\mathbb{K}^{-\frac{1}{2}}|\boldsymbol{u}|\right)\boldsymbol{u} - \mu_{\text{eff}}\Delta\boldsymbol{u},$$

has been proposed, where $c_F$ is called form drag coefficient and depends on the porous medium. It is an extension of the momentum balance in the Brinkman model, additionally accounting for inertial effects. For more information, in particular concerning the specification of model-appropriate fluid flow regimes, see, e.g., [Vaf00, Ch. 3.5], [JNP82], and references therein.

## 2.4 Summary

Starting with the introduction of basic properties of porous media and fluids, we have introduced the micro- to macro-scale concepts important in the porous media flow context. The Stokes, Darcy, and Brinkman equations have then (partly heuristically as in the original works) been derived and we have noted that various theoretical derivations exist for the porous media flow models.

The Brinkman model, as an extension of the Darcy model, takes into account viscous stresses, is therefore compatible with no-slip boundary conditions, and can represent boundary/internal layers. The corresponding Brinkman parameter, the effective viscosity, is a macroscopic parameter that can be defined in several different ways.

Moreover, we have seen that the ranges of validity of the Stokes, Darcy, and Brinkman equations can be roughly distinguished by considering the relation

[7] More precisely, for a discretization method based on a variational formulation, the application of quadrature formulas requires that the coefficients are known at all quadrature points.

of pore scale and problem scale but are not precisely defined in general. This makes the choice of the appropriate model an issue.

A large field of applications is related to coupling approaches with coexisting porous and plain regions. Interface conditions that are used for coupling the models with each other were studied and disadvantages of such explicit couplings related to the required analytic knowledge of the interface, the presence of unknown coefficients (Beavers–Joseph), and the smearing of the model ranges are pointed out.

Finally, the chapter is concluded with the observation that the Brinkman model has the potential to contribute to a unified approach since it represents a family of models with the Stokes and Darcy models as limits, depending on the physical coefficients.

# 3 The Brinkman Problem and its Weak Formulations

Weak formulations of model problems are the basis for numerical approximations by the finite element method. Their discussion is an important part of the preparation and therefore the central goal of this chapter.

In the beginning of the 20th century, Hadamard shaped the idea and notion of well-posedness, [Had07]. Considering problems based on systems of partial differential equations, he introduced the concept of a *well-posed* problem to satisfy three criteria:

1. Existence of a solution,
2. Uniqueness of the solution,
3. Continuous dependence of the solution on the data.

The analysis of a weak problem concerning these properties requires tools from functional analysis and well-posedness results that will be reviewed. Subsequently, we will figure out appropriate functional settings and examine the well-posedness of the limits of the weak Brinkman problem.

## 3.1 Functional Theory

### 3.1.1 Function Spaces

In the following we introduce function spaces that will be used in the construction and analysis of weak formulations. Thereby, vector-valued versions of spaces, e.g., $\boldsymbol{V} = V^n = \underbrace{V \times \cdots \times V}_{n \text{ times}}$ are considered whenever bold faced symbols are used here and in the rest of this thesis.

#### The Lebesgue Spaces

The spaces of Lebesgue-integrable functions on $\boldsymbol{\Omega}$ are defined as

$$L^p(\boldsymbol{\Omega}) := \left\{ v \colon \boldsymbol{\Omega} \longrightarrow \mathbb{R} \text{ measurable} \;:\; \int_{\boldsymbol{\Omega}} |v|^p < \infty \right\}, \quad p \in [1, \infty),$$
$$L^\infty(\boldsymbol{\Omega}) := \left\{ v \colon \boldsymbol{\Omega} \longrightarrow \mathbb{R} \text{ measurable} \;:\; |v| < \infty \text{ a.e. in } \boldsymbol{\Omega} \right\},$$

and equipped with the norms

$$\|v\|_{L^p(\boldsymbol{\Omega})} := \begin{cases} \left(\int_{\boldsymbol{\Omega}} |v|^p\right)^{\frac{1}{p}}, & p \in [1, \infty), \\ \operatorname{ess\,sup}_{\boldsymbol{\Omega}} |v|, & p = \infty, \end{cases}$$

see, e.g., [Gal11, Ch. II.2]. Of special interest will be the case $p = 2$, i.e., the space of square-integrable functions on $\boldsymbol{\Omega}$ given by

$$L^2(\boldsymbol{\Omega}) := \left\{ v \colon \boldsymbol{\Omega} \longrightarrow \mathbb{R} \text{ measurable} \ : \ \int_{\boldsymbol{\Omega}} |v|^2 < \infty \right\},$$

with corresponding norm

$$\|v\|_{L^2(\boldsymbol{\Omega})} := \left( \int_{\boldsymbol{\Omega}} |v|^2 \right)^{\frac{1}{2}},$$

which is induced by the inner product

$$(u, v)_{L^2(\boldsymbol{\Omega})} := \int_{\boldsymbol{\Omega}} uv.$$

Hence, the space $\left(L^2(\boldsymbol{\Omega}), \|\cdot\|_{L^2(\boldsymbol{\Omega})}\right)$ becomes a Hilbert space. For simplicity of notation, the index of the $L^2$-norm will be omitted and replaced by a zero index, and for the $L^2$ product round brackets solely without any index will be used, i.e.,

$$\|\cdot\|_0 := \|\cdot\|_{L^2(\boldsymbol{\Omega})} \qquad \text{and} \qquad (\cdot,\cdot) := (\cdot,\cdot)_{L^2(\boldsymbol{\Omega})}.$$

For a set $\boldsymbol{A} \subsetneq \boldsymbol{\Omega}$, we will write

$$\|\cdot\|_{0,\boldsymbol{A}} := \|\cdot\|_{L^2(\boldsymbol{A})} \qquad \text{and} \qquad (\cdot,\cdot)_{\boldsymbol{A}} := (\cdot,\cdot)_{L^2(\boldsymbol{A})}.$$

The notation $L^2_0(\boldsymbol{\Omega})$ will be used to denote the subspace of $L^2$-functions with integral mean value zero, i.e.,

$$L^2_0(\boldsymbol{\Omega}) := \left\{ v \in L^2(\boldsymbol{\Omega}) \ : \ \int_{\boldsymbol{\Omega}} v = 0 \right\}.$$

With respect to the boundary $\Gamma := \partial\boldsymbol{\Omega}$, we will use the notation

$$\langle u, v \rangle := \langle u, v \rangle_{L^2(\Gamma)} := \int_{\Gamma} uv$$

and whenever a set $\boldsymbol{B} \subsetneq \Gamma$ is considered, we will write $\langle\cdot,\cdot\rangle_{\boldsymbol{B}} := \langle\cdot,\cdot\rangle_{L^2(\boldsymbol{B})}$ and $\|\cdot\|_{0,\boldsymbol{B}} := \|\cdot\|_{L^2(\boldsymbol{B})}$.

### The Sobolev Spaces

Let $\boldsymbol{\alpha} := (\alpha_1, \ldots, \alpha_n)$ be a multi-index with $\alpha_1, \ldots, \alpha_n \in \mathbb{N}_0$ and magnitude $|\boldsymbol{\alpha}| = \sum_{i=1}^{n} \alpha_i$, and denote by $D^{\boldsymbol{\alpha}}$ the weak derivative with respect to $\boldsymbol{\alpha}$ (see, e.g., [GR86] or [Gal11]). The *Sobolev spaces* $H^k(\boldsymbol{\Omega})$ for $k \in \mathbb{N}_0$ are defined as

$$H^k(\boldsymbol{\Omega}) := \left\{ v \in L^2(\boldsymbol{\Omega}) \ : \ D^{\boldsymbol{\alpha}} v \in L^2(\boldsymbol{\Omega}), \quad \forall |\boldsymbol{\alpha}| \leqslant k \right\},$$

with the standard norms

$$\|v\|_{H^k(\boldsymbol{\Omega})} := \left( \sum_{|\boldsymbol{\alpha}| \leqslant k} \|D^{\boldsymbol{\alpha}} v\|_0^2 \right)^{\frac{1}{2}}.$$

Together with the inner products

$$(u, v)_{H^k(\boldsymbol{\Omega})} := \sum_{|\boldsymbol{\alpha}| \leqslant k} (D^{\boldsymbol{\alpha}} u, D^{\boldsymbol{\alpha}} v),$$

the Sobolev spaces $\left(H^k(\boldsymbol{\Omega}), \|\cdot\|_{H^k(\boldsymbol{\Omega})}\right)$ become Hilbert spaces. Moreover, one can define a semi-norm $|\cdot|_{H^k(\boldsymbol{\Omega})}$ as

$$|v|_{H^k(\boldsymbol{\Omega})} := \left( \sum_{|\boldsymbol{\alpha}| = k} \|D^{\boldsymbol{\alpha}} v\|_0^2 \right)^{\frac{1}{2}}.$$

We will use the abbreviations

$$\|v\|_k = \|v\|_{H^k(\boldsymbol{\Omega})} \qquad \text{and} \qquad |v|_k = |v|_{H^k(\boldsymbol{\Omega})}.$$

For $k = 0$, we have $H^0(\boldsymbol{\Omega}) = L^2(\boldsymbol{\Omega})$, which gives an explanation for the $L^2$-norm notation with index 0. The Sobolev space for $k = 1$ will be particularly important in this thesis (apart from $k = 0$):

$$H^1(\boldsymbol{\Omega}) := \left\{ v \in L^2(\boldsymbol{\Omega}) \ : \ \nabla v \in \boldsymbol{L}^2(\boldsymbol{\Omega}) \right\}$$

with the norm

$$\|v\|_{H^1(\boldsymbol{\Omega})} := \left( \|\nabla v\|_0^2 + \|v\|_0^2 \right)^{\frac{1}{2}},$$

induced by the inner product

$$(u, v)_{H^1(\boldsymbol{\Omega})} := (\nabla u, \nabla v) + (u, v),$$

and the semi-norm

$$|v|_{H^1(\boldsymbol{\Omega})} := \|\nabla v\|_0.$$

The consideration of Sobolev functions $v \in H^1(\boldsymbol{\Omega})$ on the boundary $\Gamma$ is justified by the existence of a continuous surjective (with respect to a specific subspace of $L^2(\boldsymbol{\Omega})$) trace operator such that $v \mapsto v|_\Gamma$. For more information on traces of Sobolev functions, see, e.g., [DD12, Thm. 3.9], [QV08, Thm. 1.3.1], or [BF91, III.1.1]. The subspace $H_0^1(\boldsymbol{\Omega}) \subset H^1(\boldsymbol{\Omega})$ contains only functions from $H^1(\boldsymbol{\Omega})$ which vanish at the boundary in the sense of traces

$$H_0^1(\boldsymbol{\Omega}) := \left\{ v \in H^1(\boldsymbol{\Omega}) \ : \ v|_\Gamma = 0 \right\}.$$

Moreover, let us introduce the space of vector fields whose divergence is in $L^2(\boldsymbol{\Omega})$, i.e.,

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

Equipped with the inner product

$$(\boldsymbol{u}, \boldsymbol{v})_{\boldsymbol{H}^{\mathrm{div}}(\boldsymbol{\Omega})} := (\nabla \cdot \boldsymbol{u}, \nabla \cdot \boldsymbol{v}) + (\boldsymbol{u}, \boldsymbol{v}),$$

$\left(\boldsymbol{H}^{\mathrm{div}}(\boldsymbol{\Omega}), \|\cdot\|_{\boldsymbol{H}^{\mathrm{div}}(\boldsymbol{\Omega})}\right)$ is a Hilbert space with

$$\|\boldsymbol{v}\|_{\boldsymbol{H}^{\mathrm{div}}(\boldsymbol{\Omega})} := \left( \|\nabla \cdot \boldsymbol{v}\|_0^2 + \|\boldsymbol{v}\|_0^2 \right)^{\frac{1}{2}}.$$

Furthermore, the existence of a continuous surjective trace operator for vector fields in $\boldsymbol{H}^{\mathrm{div}}(\boldsymbol{\Omega})$ such that $\boldsymbol{v} \mapsto (\boldsymbol{v} \cdot \boldsymbol{n})|_\Gamma$ is well-defined (for details see [DD12, Prop. 3.58], [QV08, Thm. 1.3.2], or [BF91, III.1.1]) legitimates to consider appropriate boundary values of functions in $\boldsymbol{H}^{\mathrm{div}}(\boldsymbol{\Omega})$. A subspace $\boldsymbol{H}_0^{\mathrm{div}}(\boldsymbol{\Omega}) \subset \boldsymbol{H}^{\mathrm{div}}(\boldsymbol{\Omega})$ will be utilized, where the index 0 here corresponds to the behavior at the boundary (in the sense of traces)

$$\boldsymbol{H}_0^{\mathrm{div}}(\boldsymbol{\Omega}) := \left\{ \boldsymbol{v} \in \boldsymbol{H}^{\mathrm{div}}(\boldsymbol{\Omega}) \ : \ (\boldsymbol{v} \cdot \boldsymbol{n})|_\Gamma = 0 \right\}.$$

#### Duality Pairings

The dual space of a normed vector space $V$ will be denoted by $V'$ and contains all bounded linear functionals on $V$. The action of an element $\phi \in V'$ on an element $v \in V$

$$\phi(v) = [\phi, v]_{V',V}$$

will be referred to as *duality pairing*. The norm on the dual space $V'$ is then defined as

$$\|\phi\|_{V'} := \sup_{v \in V \setminus \{0\}} \frac{[\phi, v]_{V',V}}{\|v\|_V}.$$

For $\phi, v \in L^2(\boldsymbol{\Omega})$, the duality pairing is understood as

$$[\phi, v]_{V',V} := (\phi, v).$$

The dual space of $H_0^1(\boldsymbol{\Omega})$ will be denoted by $H^{-1}(\boldsymbol{\Omega}) := \left(H_0^1(\boldsymbol{\Omega})\right)'$.

### 3.1.2 Some Inequalities

The purpose of this subsection is to introduce several inequalities and results that will be used in the rest of the thesis. In what follows, we will denote the *norm in Euclidean space* $\mathbb{R}^N$ for $N \in \mathbb{N}_{\geqslant 0}$ and $\boldsymbol{a} \in \mathbb{R}^N$ by

$$\|\boldsymbol{a}\| = |\boldsymbol{a}| := \left( \sum_{i=1}^{N} a_i^2 \right)^{\frac{1}{2}}, \tag{3.1.1}$$

which is induced by the inner product $\boldsymbol{a} \cdot \boldsymbol{b} := \sum_{i=1}^{N} a_i b_i$. Both notations $\|\cdot\|$ and $|\cdot|$ are common and will be utilized synonymous.

**Theorem 3.1.1** (Young/Peter–Paul Inequalities)**.**
*For $a, b \in \mathbb{R}$ and any $\varepsilon > 0$ it holds*

$$(i) \quad ab \leqslant \varepsilon a^2 + \frac{1}{4\varepsilon} b^2, \qquad\qquad (ii) \quad ab \leqslant \frac{\varepsilon a^2}{2} + \frac{1}{2\varepsilon} b^2.$$

*Proof:* The proof is straightforward, since $(i)$ is equivalent to $0 \leqslant (2\varepsilon a - b)^2$ and $(ii)$ is implied by $(i)$ by simply replacing $\varepsilon$ with $\frac{\varepsilon}{2}$. □

**Lemma 3.1.2** (Cauchy–Schwarz Inequality for Sums)**.**
*Let $a_i, b_i \in \mathbb{R}$ and $N \in \mathbb{N}_{\geqslant 1}$. Then*

$$\sum_{i=1}^{N} a_i b_i \leqslant \left( \sum_{i=1}^{N} a_i^2 \right)^{\frac{1}{2}} \left( \sum_{i=1}^{N} b_i^2 \right)^{\frac{1}{2}}.$$

*Proof:* Let us introduce two vectors in the Euclidean space by $\boldsymbol{a} = (a_1, \ldots, a_N)^T$ and $\boldsymbol{b} = (b_1, \ldots, b_N)^T$. Then the definition of the dot product yields

$$\boldsymbol{a} \cdot \boldsymbol{b} = |\boldsymbol{a}| \, |\boldsymbol{b}| \cos\left(\angle(\boldsymbol{a}, \boldsymbol{b})\right) \leqslant |\boldsymbol{a}| \, |\boldsymbol{b}| \,,$$

which proves the inequality. □

**Corollary 3.1.3.** *Let $a_i \in \mathbb{R}$ and $N \in \mathbb{N}_{\geqslant 1}$, then*

$$\left( \sum_{i=1}^{N} a_i \right)^2 \leqslant N \sum_{i=1}^{N} a_i^2.$$

*Proof:* This estimate is a direct consequence of the Cauchy–Schwarz inequality for sums (Lem. 3.1.2) with $b_i := 1$, for all $i = 1, \ldots, N$. □

**Corollary 3.1.4.** *For $\boldsymbol{v} \in \boldsymbol{H}^1(\boldsymbol{\Omega}) = \left[H^1(\boldsymbol{\Omega})\right]^N$, $N \in \mathbb{N}_{>1}$, it holds* [1]

$$\|\nabla \cdot \boldsymbol{v}\|_0 \leqslant N^{\frac{1}{2}} \|\nabla \boldsymbol{v}\|_0.$$

*Proof:* The inequality is a consequence of Corollary 3.1.3, since

$$\|\nabla \cdot \boldsymbol{v}\|_0^2 = \int_{\boldsymbol{\Omega}} \left| \sum_{i=1}^{N} \frac{\partial v_i}{\partial x_i} \right|^2 \leqslant N \int_{\boldsymbol{\Omega}} \sum_{i=1}^{N} \left( \frac{\partial v_i}{\partial x_i} \right)^2 = N \|\nabla \boldsymbol{v}\|_0^2.$$

□

**Theorem 3.1.5** (Cauchy–Schwarz Inequality)**.**
*The Cauchy–Schwarz inequality is a special case of the Hölder inequality [Gal11, (II.2.3)] for functions in $L^2(\boldsymbol{\Omega})$, stating*

$$|(u, v)| \leqslant \|u\|_0 \|v\|_0, \qquad \text{for } u, v \in L^2(\boldsymbol{\Omega}).$$

[1] For $\boldsymbol{v} \in \boldsymbol{H}_0^1(\boldsymbol{\Omega})$ the inequality is valid without the factor $N^{\frac{1}{2}}$, see [Joh16, Lem. 3.179].

*Proof:* If $v = 0$ the inequality trivially becomes an equality. If $v \neq 0$ we have for any $t \in \mathbb{R}$

$$\begin{aligned} 0 \leqslant \|u - tv\|_0^2 &= (u - tv, u - tv) = (u, u) - 2t\,(u, v) + t^2\,(v, v) \\ &= \|u\|_0^2 - 2t\,(u, v) + t^2 \|v\|_0^2. \end{aligned}$$

The choice $t := \frac{(u,v)}{\|v\|_0^2}$ yields

$$0 \leqslant \|u\|_0^2 - \frac{(u, v)^2}{\|v\|_0^2}$$

and thus

$$(u, v) \leqslant \|u\|_0 \|v\|_0,$$

which is valid for $u$ and $-u$, hence implies the desired result. □

**Theorem 3.1.6** (Minkowski Inequality)**.**
*The Minkowski inequality is the equivalent of the triangle inequality in terms of Lebesgue norms. Let $u, v \in L^2(\mathbf{\Omega})$, then*

$$\|u + v\|_0 \leqslant \|u\|_0 + \|v\|_0.$$

*Proof:* Using the triangle inequality and the Cauchy–Schwarz inequality (Thm. 3.1.5), we can estimate

$$\begin{aligned} \|u + v\|_0^2 &\leqslant \int_{\mathbf{\Omega}} (|u| + |v|)\,(|u + v|) = \int_{\mathbf{\Omega}} |u|\,(|u + v|) + \int_{\mathbf{\Omega}} |v|\,(|u + v|) \\ &\leqslant \|u\|_0 \|u + v\|_0 + \|v\|_0 \|u + v\|_0. \end{aligned}$$

Dividing by $\|u + v\|_0$ gives the desired result. □

The Minkowski inequality will be used in this text without reference.

**Theorem 3.1.7** (Inequality for Duality Pairings)**.**
*Let $v \in V$, $\phi \in V'$, and $V$ a Hilbert space with dual space $V'$ (which is a Hilbert space as well). Then*

$$\left|[\phi, v]_{V',V}\right| \leqslant \|\phi\|_{V'} \|v\|_V.$$

*Proof:* By the definition of the norm for the dual space we have

$$\|\phi\|_{V'} \|v\|_V = \left( \sup_{v^\star \in V \setminus \{0\}} \frac{[\phi, v^\star]_{V',V}}{\|v^\star\|_V} \right) \|v\|_V \geqslant \frac{\left|[\phi, v]_{V',V}\right|}{\|v\|_V} \|v\|_V = \left|[\phi, v]_{V',V}\right|.$$

□

**Theorem 3.1.8** (The Classical Poincaré Inequality)**.**
*Let $\mathbf{\Omega} \subset \mathbb{R}^n$ be a bounded domain with Lipschitz boundary $\Gamma$. Then*

$$\forall v \in H_0^1(\mathbf{\Omega}): \qquad \|v\|_0 \leqslant C_P \|\nabla v\|_0,$$

*with a constant $C_P = C_P(\operatorname{diam}(\mathbf{\Omega})) > 0$.*

*Proof:* For details see the references cited in [Joh16, Thm. A.36], [GR86, p. 3], and [EG04, Lem. B.61]. □

**Remark 3.1.9** (On the Poincaré Inequality)**.**
*The classical Poincaré inequality (sometimes called Poincaré–Friedrichs inequality) implies that, in $H_0^1(\boldsymbol{\Omega})$, the standard $H^1$-norm and the $H^1$-semi-norm are equivalent:*

$$|v|_{H^1(\boldsymbol{\Omega})}^2 = \|\nabla v\|_0^2 \leqslant \|v\|_{H^1(\boldsymbol{\Omega})}^2 \leqslant \left(1 + C_P^2\right) \|\nabla v\|_0^2 = \left(1 + C_P^2\right) |v|_{H^1(\boldsymbol{\Omega})}^2 .$$

*Hence, in particular, the semi-norm $|\cdot|_{H^1(\boldsymbol{\Omega})}$ becomes a norm on $H_0^1(\boldsymbol{\Omega})$.*

**Theorem 3.1.10** (Isomorphism of the Divergence Operator)**.**
*Define the subspaces $\widetilde{\boldsymbol{V}}_0, \widetilde{\boldsymbol{V}}_0^{\perp} \subset \boldsymbol{H}_0^1(\boldsymbol{\Omega})$ as*

$$\widetilde{\boldsymbol{V}}_0 := \left\{ \boldsymbol{v} \in \boldsymbol{H}_0^1(\boldsymbol{\Omega}) \ : \ (\nabla \cdot \boldsymbol{v}, q) = 0, \quad \forall q \in L_0^2(\boldsymbol{\Omega}) \right\},$$
$$\widetilde{\boldsymbol{V}}_0^{\perp} := \left\{ \boldsymbol{v} \in \widetilde{\boldsymbol{V}}_0 \ : \ (\nabla \boldsymbol{v}, \nabla \boldsymbol{w}) = 0, \quad \forall \boldsymbol{w} \in \widetilde{\boldsymbol{V}}_0 \right\}.$$

*Then the divergence operator*

$$\nabla \cdot (\cdot) \ : \ \widetilde{\boldsymbol{V}}_0^{\perp} \longrightarrow L_0^2(\boldsymbol{\Omega})$$

*is an isomorphism. As a consequence, for any $q \in L_0^2(\boldsymbol{\Omega})$ $\exists!\, \widetilde{\boldsymbol{v}} \in \widetilde{\boldsymbol{V}}_0^{\perp}$ such that*

$$\nabla \cdot \widetilde{\boldsymbol{v}} = q \quad \text{and} \quad \|q\|_0 \leqslant |\widetilde{\boldsymbol{v}}|_{\boldsymbol{H}^1(\boldsymbol{\Omega})} \leqslant C \|q\|_0,$$

*where the constant $C > 0$ does not depend on $q$ and $\widetilde{\boldsymbol{v}}$.*[2][3]

*Proof:* See [Joh16, Lem. 3.43, Cor. 3.44] or [GR86, Cor. 2.4]. □

### 3.1.3 Existence and Uniqueness Theory

This subsection is used to present existence and uniqueness results for certain abstract problems, in particular so-called saddle point problems, posed in Hilbert spaces. Therefore we state necessary and sufficient conditions that guarantee their well-posedness in the sense of Hadamard. This theory will then be used to analyze the weak Brinkman problems in the subsequent sections.

Let us denote by $(V, \|\cdot\|_V)$ and $(Q, \|\cdot\|_Q)$ two real Hilbert spaces with their norms induced by the inner products $(\cdot,\cdot)_V$ and $(\cdot,\cdot)_Q$, respectively. Their dual spaces are $V'$ and $Q'$, respectively, and equipped with the norms $\|\cdot\|_{V'}$ and $\|\cdot\|_{Q'}$.

Most of the following theorems consider bounded/continuous bilinear forms. A bilinear form $b(\cdot,\cdot) : V \times Q \to \mathbb{R}$ on Hilbert spaces $(V, \|\cdot\|_V)$ and $(Q, \|\cdot\|_Q)$ is called **continuous/bounded** if

$$\exists\, M > 0 : \quad |b(v,q)| \leqslant M \|v\|_V \|q\|_Q, \quad \forall\, (v,q) \in V \times Q.$$

[2] In a few occasions, the letters $c$ and $C$ are used in this dissertation to represent generic constants.

[3] Using the equivalence of the $\boldsymbol{H}^1(\boldsymbol{\Omega})$-norm and the $\boldsymbol{H}^1(\boldsymbol{\Omega})$-semi norm on $\boldsymbol{H}_0^1(\boldsymbol{\Omega})$ the inequality can be formulated with the norm instead of the semi norm, in general with different constants.

### General Abstract Problem

Let us define the **general abstract problem**:
Given a bounded bilinear form $a(\cdot,\cdot) : V \times V \to \mathbb{R}$ and a functional $f \in V'$, find an element $u \in V$ such that

$$a(u,v) = [f,v]_{V',V}, \quad \forall v \in V. \tag{3.1.2}$$

In order to discuss the well-posedness of problems of the form (3.1.2) let us clarify what this means in detail.

**Definition 3.1.11** (Well-Posedness)**.**
*The problem* (3.1.2) *is* ***well-posed*** *if it has exactly one solution and the following a priori estimate holds*

$$\exists\, C > 0 \text{ such that } \forall f \in V' : \qquad \|u\|_V \leqslant C \|f\|_{V'}.$$

The purpose of the rest of this paragraph is to introduce necessary and sufficient conditions for the well-posedness of the problem (3.1.2), which will reveal crucial properties of $a(\cdot,\cdot)$.

Let us begin with the following fundamental result:

**Theorem 3.1.12** (Lax–Milgram)**.**
*In the case that*

$$a(\cdot,\cdot) \text{ is } \textbf{coercive}, \text{ i.e., } \exists\, m > 0 : \quad a(u,u) \geqslant m\|u\|_V^2, \quad \forall u \in V,$$

*the problem* (3.1.2) *is well-posed. Additionally, for all $f \in V'$, the a priori estimate*

$$\|u\|_V \leqslant \frac{1}{m}\|f\|_{V'}$$

*is valid, where $m > 0$ is the coercivity constant.*

*Proof:* See [BF91, Prop. 1.1], [BS08, (2.7.7)], [Cia02, Thm. 1.1.3], [EG04, Sect. 2.1], or [Joh16, Thm. B.4]. □

The theorem of Lax–Milgram describes only a sufficient condition for the well-posedness. The following theorem states necessary conditions for the well-posedness of an abstract problem of the form (3.1.2).

**Theorem 3.1.13** (Generalized Lax–Milgram/BNB)**.**
*The problem* (3.1.2) *is well-posed if and only if the following properties hold:*
*1.) The bilinear form $a(\cdot,\cdot)$ fulfills the so-called* ***inf-sup condition****, i.e.,*

$$\exists\, \beta > 0 : \qquad \inf_{u \in V\setminus\{0\}} \left( \sup_{v \in V\setminus\{0\}} \frac{a(u,v)}{\|u\|_V \|v\|_V} \right) \geqslant \beta.$$

*2.) Let $v \in V$. If $a(u,v) = 0,\ \forall u \in V,$ then $v = 0$.*

*Furthermore, for all $f \in V'$, it holds*

$$\|u\|_V \leqslant \frac{1}{\beta}\|f\|_{V'}.$$

*Proof:* See [EG04, Thm 2.6], [Bra07, Thm. 3.6], and references in [BO80]. □

**Remark 3.1.14** (On the BNB Theorem)**.**
*Theorem 3.1.13 is also called Banach–Nečas–Babuška theorem, which justifies its abbreviation by BNB. The theorem of Lax-Milgram is in fact a special case of the BNB theorem, since coercivity of $a(\cdot,\cdot)$ implies the conditions in Theorem 3.1.13*

- *coercivity $\Longrightarrow$ 1.):*

$$\sup_{v\in V\setminus\{0\}} \frac{a(u,v)}{\|u\|_V\|v\|_V} \geqslant \frac{a(u,u)}{\|u\|_V\|u\|_V} \geqslant m\frac{\|u\|_V^2}{\|u\|_V^2} = m, \quad \forall u \in V,$$

- *coercivity $\Longrightarrow$ 2.):*

$$a(u,v) = 0, \quad \forall u \in V \quad \overset{u:=v}{\Longrightarrow} \quad 0 = a(v,v) \geqslant \underbrace{m}_{>0}\underbrace{\|v\|_V^2}_{\geqslant 0} \quad \Longrightarrow \quad v = 0.$$

*The opposite direction is in general not satisfied.*

### Linear Saddle Point Problem

The saddle point problem theory was developed for abstract problems with a specific structure. Saddle point problems arise, e.g., when weakly incorporating side conditions via a Lagrange multiplier approach, artificially or naturally as it is for example the case for the mass conservation in fluid dynamics.

Let us introduce the **general abstract linear saddle point problem**: Given two bounded bilinear forms $a(\cdot,\cdot) : V \times V \to \mathbb{R}$ and $b(\cdot,\cdot) : V \times Q \to \mathbb{R}$, $f(\cdot) : V \to \mathbb{R}$ with $f \in V'$, and $g(\cdot) : Q \to \mathbb{R}$ satisfying $g \in Q'$, find a pair $(u,p) \in V \times Q$ such that

$$a(u,v) + b(v,p) = [f,v]_{V',V}, \quad \forall v \in V, \tag{3.1.3a}$$
$$b(u,q) = [g,q]_{Q',Q}, \quad \forall q \in Q. \tag{3.1.3b}$$

**Remark 3.1.15** (Connection to Saddle Point Problems)**.**
*In the special case of a symmetric (i.e., $a(u,v) = a(v,u), \forall u,v \in V$) and positive (i.e., $a(u,u) \geqslant 0, \quad \forall u \in V$) bilinear form $a(\cdot,\cdot)$, one can show that the problem* (3.1.3) *is equivalent to the saddle point problem*

$$\inf_{v\in V}\sup_{q\in Q}\left(\frac{1}{2}a(v,v) + b(v,q) - [f,v]_{V',V} - [g,q]_{Q',Q}\right).$$

*For a proof of this statement, [EG04, Prop. 2.39] or [GR86, Thm. 4.2] can be consulted. This is the origin of the naming convention which has established itself even for more general problems of the form* (3.1.3)*.*

**Remark 3.1.16** (Saddle Point Problems as General Abstract Problems)**.** *The well-posedness theory of the general abstract problem* (3.1.2) *can be applied to* (3.1.3) *as follows. Let* $\left(V \times Q, \|(v,q)\|^2_{V\times Q} := \|v\|^2_V + \|q\|^2_Q\right)$ *be a Hilbert space and* $(f,g) \in V' \times Q'$*. Define*

$$\begin{aligned} \mathcal{A}[(\cdot,\cdot);(\cdot,\cdot)]\colon\ & (V\times Q)\times(V\times Q)\to\mathbb{R},\\ \mathcal{A}[(u,p);(v,q)] :=\ & a(u,v) + b(v,p) \pm b(u,q),\\ \mathcal{F}(\cdot,\cdot)\colon\ & V\times Q\to\mathbb{R},\\ \mathcal{F}(v,q) :=\ & f(v) \pm g(q), \end{aligned}$$

*where* $a(\cdot,\cdot) : V\times V\to\mathbb{R}$ *and* $b(\cdot,\cdot) : V\times Q\to\mathbb{R}$ *are bilinear forms thus* $\mathcal{A}$ *is a bilinear form. Since* $(f,g)\in V'\times Q'$ *it is* $\mathcal{F}\in(V\times Q)'$*. Then, it can be proven that the problem:*
*Find a pair* $(u,p)\in V\times Q$ *such that*

$$\mathcal{A}[(u,p);(v,q)] = \mathcal{F}(v,q), \quad \forall\,(v,q)\in V\times Q,$$

*is well-posed using Theorem 3.1.12 or 3.1.13. In fact, the sum (respectively the difference) of* (3.1.3a) *and* (3.1.3b) *yields the variational equation above, while the reverse direction can be realized by testing with* $(v,q) = (v,0)$ *and* $(v,q) = (0,q)$ *(respectively* $(v,q) = (0,-q)$ *for* $-q\in Q$*), as long as* $0\in V,Q$*.*

In the analysis of the problem (3.1.3), one can exploit the special structure to deduce well-posedness in an elegant manner based on conditions on the bilinear forms $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$.

Let us begin with motivating, preliminary observations. To that end, we define a subset of the space $V$ as

$$V_g := \left\{v\in V\ :\ b(v,q) = [g,q]_{Q',Q},\quad \forall q\in Q\right\}. \tag{3.1.4}$$

In general, $V_g$ is not a linear subspace of $V$. To see that let $v_1, v_2\in V_g$ and consider their sum $v_1+v_2$. Bilinearity of $b(\cdot,\cdot)$ and linearity of $g(\cdot)$ yield

$$b(v_1+v_2,q) = b(v_1,q) + b(v_2,q) = g(q) + g(q) = g(2q).$$

Therefore, $v_1+v_2\notin V_g$, for general $g\in Q'$. However, $V_0$ is a linear subspace. Moreover, $(V_0,(\cdot,\cdot)_V)$ is a Hilbert space, since it is a closed subspace, [Bre74].

($\star_1$) The equation (3.1.3b) poses a constraint on $u$. Let us firstly assume that there exists a unique solution to that equation[4].

($\star_2$) Then, in a second step, this condition can be enforced in (3.1.3a) by changing the solution space for $u$, such that it includes only functions that satisfy the constraint (3.1.3b):
Find $u\in V_g$ such that

$$a(u,v) = f(v), \quad \forall v\in V_0. \tag{3.1.5}$$

[4] More precisely, orthogonally decomposing $V = V_0 + V_0^\perp$ with respect to the inner product in $V$, we assume that (3.1.3b) has an exact solution in $V_0^\perp$.

Apparently, the result is an equation that solely explicitly contains the partial solution $u$ and not the partial solution $p$ anymore. It is straightforward to see that the partial solution $u$ of (3.1.3) then also solves the problem (3.1.5). Consequently, for the other way around, it suffices to show that (3.1.5) is uniquely solvable. Decomposing $u = u_0 + u_g$[5] with $u_0 \in V_0$, and using the theorem of Lax–Milgram (Thm. 3.1.12, which is a sufficient criterion), we can prove the existence of a unique solution $u_0 \in V_0$ of

$$a(u_0, v) = f(v) - a(u_g, v), \quad \forall v \in V_0,$$

if $a(\cdot,\cdot)$ is coercive and $f(v) - a(u_g, v)$ is bounded. This is a particularly simple ansatz in case of $g = 0$.

($\star_3$) In a third step we would like to assure the existence of a unique solution $p \in Q$ (corresponding to the previously derived partial solution $u$) solving (3.1.3a), i.e.,

$$b(v, p) = f(v) - a(u, v), \quad \forall v \in V.$$

The validity of ($\star_1$) and ($\star_3$) can be assured by the inf-sup condition for $b(\cdot,\cdot)$.

In 1974, Brezzi published an abstract theory tackling the existence and uniqueness of solutions to saddle point problems, [Bre74]. The final result in the continuous case can be formulated as follows:

**Theorem 3.1.17** (Brezzi's Splitting)**.**
*In the case that*

*(i) the bilinear form $a(\cdot,\cdot)$ is coercive on $V_0$ (defined in* (3.1.4)*), i.e.,*

$$\exists\, m > 0 : \quad a(v, v) \geqslant m\|v\|_V^2, \quad \forall v \in V_0,$$

*(ii) the bilinear form $b(\cdot,\cdot)$ fulfills the inf-sup condition*

$$\exists\, \beta > 0 : \qquad \inf_{q \in Q\setminus\{0\}} \left( \sup_{v \in V\setminus\{0\}} \frac{b(v, q)}{\|v\|_V \|q\|_Q} \right) \geqslant \beta,$$

*the problem* (3.1.3) *is well-posed.*

*Proof:* The proof is given, e.g., in [Bre74, Cor. 1.1], [BF91, Prop. 1.1], [Joh16, Lem. 3.19], and [GR86, Cor. 4.1]. Note that the a priori stability estimate depends inversely on the coercivity and inf-sup constants, see e.g., [BF91, Prop. 1.3]. □

**Remark 3.1.18** (On Brezzi's Splitting Theorem and the Inf-Sup Condition)**.**
*The inf-sup condition (inequality (ii) in Theorem 3.1.17) is also known as Babuška–Brezzi condition or Ladyzhenskaya–Babuška–Brezzi (LBB) condition. Brezzi's splitting theorem describes conditions that are sufficient for the well-posedness of the saddle point problem* (3.1.3)*. Necessary conditions are given, e.g., in [BF91, Thm. 1.1].*

[5] One can consider the orthogonal decomposition $V = V_0 + V_0^\perp$ (with respect to the inner product in $V$), $u_0 \in V_0$, $u_g = u - u_0 \in V_0^\perp$ according to ($\star_1$), and obtains two problems by testing (3.1.3) with $V_0$ and $V_0^\perp$ separately.

## 3.2 The Brinkman Problem

Remind that the Brinkman equations are given by

$$\begin{aligned} -2\nabla \cdot (\mu_{\text{eff}} \mathbb{D}(\boldsymbol{u})) + \nabla p + \sigma \boldsymbol{u} &= \boldsymbol{f} \quad \text{in } \boldsymbol{\Omega}, \\ \nabla \cdot \boldsymbol{u} &= g \quad \text{in } \boldsymbol{\Omega}, \end{aligned}$$

with $\mu_{\text{eff}} \geqslant 0$ denoting the effective viscosity, $\sigma \geqslant 0$ standing for the ratio of the dynamic viscosity and the permeability, i.e.,

$$\sigma := \mu \, \mathbb{K}^{-1},$$

satisfying $\mu_{\text{eff}} + \sigma > 0$, and

- a vectorial velocity field $\boldsymbol{u} \colon \boldsymbol{\Omega} \to \mathbb{R}^N$,
- a scalar pressure field $p \colon \boldsymbol{\Omega} \to \mathbb{R}$,
- vectorial body forces $\boldsymbol{f} \colon \boldsymbol{\Omega} \to \mathbb{R}^N$,
- a scalar source/sink function $g \colon \boldsymbol{\Omega} \to \mathbb{R}$.

This is a system of two steady-state partial differential equations, associated with momentum balance and mass conservation, respectively, consisting of $N+1$ scalar equations.

**Remark 3.2.1** (Variants of the Momentum Balance)**.**
*Different formulations might appear due to the reformulation of the first term in the momentum balance equation for special cases. For a (continuously) differentiable scalar effective viscosity we can rewrite*

$$-2\nabla \cdot (\mu_{\text{eff}} \mathbb{D}(\boldsymbol{u})) = -2 (\nabla \mu_{\text{eff}}) \mathbb{D}(\boldsymbol{u}) - 2\mu_{\text{eff}} (\nabla \cdot \mathbb{D}(\boldsymbol{u})),$$

*using the product rule and the linearity. If $\mu_{\text{eff}}(\boldsymbol{x})$ is tensorial, we can use a similar transformation, where $\nabla \mu_{\text{eff}}(\boldsymbol{x})$ has to be replaced by $\nabla \cdot \mu_{\text{eff}}(\boldsymbol{x})$.*
*Under the assumption of $\mu_{\text{eff}}(\boldsymbol{x}) = const.$, the previously derived expression leads to:*

$$\begin{aligned} -2\nabla \cdot (\mu_{\text{eff}} \mathbb{D}(\boldsymbol{u})) &= -2\mu_{\text{eff}} (\nabla \cdot \mathbb{D}(\boldsymbol{u})) = -\mu_{\text{eff}} (\nabla \cdot \nabla \boldsymbol{u} + \nabla (\nabla \cdot \boldsymbol{u})) \\ &= -\mu_{\text{eff}} (\nabla \cdot \nabla \boldsymbol{u} + \nabla g) = -\mu_{\text{eff}} (\Delta \boldsymbol{u} + \nabla g), \end{aligned}$$

*where also the equation of mass conservation and Lemma 2.1.2 were used. This results in the momentum balance*

$$-\mu_{\text{eff}} \Delta \boldsymbol{u} + \nabla p + \sigma \boldsymbol{u} = \boldsymbol{f} + \mu_{\text{eff}} \nabla g.$$

### The Dimensionless Brinkman Equations

The *non-dimensionalization* has the scope to remove the physical units from the considered equations via an appropriate scaling with physical characteristic quantities. Furthermore, this process allows to identify dimensionless coefficients that might help to describe the physical regimes (e.g., the Reynolds number for the Navier–Stokes equations).

The derivation of a dimensionless formulation starts with defining so-called characteristic quantities (whose precise values are not uniquely defined) for the dimensional variables. Depending on the choice of the characteristic quantities (and the definition of new variables), various non-dimensional formulations of one and the same dimensional system of equations might be derived. Here we consider

- U - characteristic velocity [$^{\mathrm{m}}/_{\mathrm{s}}$],
- P - characteristic pressure [$^{\mathrm{kg}}/_{\mathrm{m\,s^2}}$],
- L - characteristic length [m].

Let us equip dimensionless quantities with a hat in this paragraph. The problem is then formulated in terms of the dimensionless velocity $\hat{\boldsymbol{u}}$, the dimensionless pressure $\hat{p}$, and the dimensionless spatial variable $\hat{\boldsymbol{x}}$, given by

$$\hat{\boldsymbol{u}} := \frac{\boldsymbol{u}}{U}, \qquad \hat{p} := \frac{p}{P}, \qquad \text{and} \qquad \hat{\boldsymbol{x}} := \frac{\boldsymbol{x}}{L}.$$

Inserting the resulting expressions for the dimensional quantities, i.e., $\boldsymbol{u} = U\hat{\boldsymbol{u}}$, $p = P\hat{p}$, and $\boldsymbol{x} = L\hat{\boldsymbol{x}}$, and transforming the differential operators appropriately ($\nabla = \frac{1}{L}\hat{\nabla}$) yields

$$\begin{aligned} -2\hat{\nabla}\cdot\left(\mu_{\text{eff}}\frac{U}{L^2}\hat{\mathbb{D}}\left(\hat{\boldsymbol{u}}\right)\right) + \frac{P}{L}\hat{\nabla}\hat{p} + \sigma U\hat{\boldsymbol{u}} &= \boldsymbol{f} \quad \text{in } \hat{\boldsymbol{\Omega}}, \\ \frac{U}{L}\hat{\nabla}\cdot\hat{\boldsymbol{u}} &= g \quad \text{in } \hat{\boldsymbol{\Omega}}. \end{aligned}$$

Dividing the momentum balance equation by any coefficient on the left-hand side results in a non-dimensional summand and thus in a dimensionless equation. In order to prevent non-zero constraints on the physical coefficients $\mu_{\text{eff}}$ and $\sigma$, we multiply the first equation by the inverse coefficient of the pressure gradient and obtain

$$\begin{aligned} -2\hat{\nabla}\cdot\left(\mu_{\text{eff}}\frac{U}{LP}\hat{\mathbb{D}}\left(\hat{\boldsymbol{u}}\right)\right) + \hat{\nabla}\hat{p} + \sigma\frac{LU}{P}\hat{\boldsymbol{u}} &= \frac{L}{P}\boldsymbol{f} \quad \text{in } \hat{\boldsymbol{\Omega}}, \\ \hat{\nabla}\cdot\hat{\boldsymbol{u}} &= \frac{L}{U}g \quad \text{in } \hat{\boldsymbol{\Omega}}, \end{aligned}$$

where the second equation was multiplied by $\frac{L}{U}$. As we can see, the velocity terms that contribute to the transport of momentum scale inversely with respect to the characteristic length $L$. Whenever $L$ is large, the bulk resistance dominates the viscous shear stress, and vice versa in case of a small $L$. Now one typically introduces dimensionless constituents

$$\hat{\mu}_{\text{eff}} := \mu_{\text{eff}}\frac{U}{LP}, \qquad \hat{\sigma} := \sigma\frac{LU}{P},$$
$$\hat{\boldsymbol{f}} := \frac{L}{P}\boldsymbol{f}, \qquad \hat{g} := \frac{L}{U}g,$$

yielding the *dimensionless Brinkman equations* as

$$\begin{aligned} -2\hat{\nabla}\cdot\left(\hat{\mu}_{\text{eff}}\hat{\mathbb{D}}\left(\hat{\boldsymbol{u}}\right)\right) + \hat{\nabla}\hat{p} + \hat{\sigma}\hat{\boldsymbol{u}} &= \hat{\boldsymbol{f}} \quad \text{in } \hat{\boldsymbol{\Omega}}, \\ \hat{\nabla}\cdot\hat{\boldsymbol{u}} &= \hat{g} \quad \text{in } \hat{\boldsymbol{\Omega}}. \end{aligned}$$

**Remark 3.2.2** (Another Popular Form of the Brinkman Equations)**.**
*Later on, we will come across a form of the momentum balance in the Brinkman equations, where the physical problem coefficients ($\mu_{\text{eff}}$ and $\sigma$) are combined into one coefficient in front of the second order derivative term. This can be formally obtained by using the characteristic velocity and the characteristic length only, and scaling by the lowest order velocity coefficient:*

$$-2\frac{1}{\sigma}\hat{\nabla}\cdot\left(\frac{\mu_{\text{eff}}}{L^2}\hat{\mathbb{D}}\left(\hat{\boldsymbol{u}}\right)\right)+\frac{1}{\sigma UL}\hat{\nabla}p+\hat{\boldsymbol{u}}=\frac{1}{\sigma U}\boldsymbol{f}\qquad \text{in } \hat{\boldsymbol{\Omega}}.$$

*This necessitates the exclusion of the case $\sigma = 0$ (strict Stokes limit). Under the further assumption of $\sigma$ being constant, we can redefine*

$$\hat{p}:=\frac{1}{\sigma UL}p\qquad \text{and}\qquad \hat{\boldsymbol{f}}:=\frac{1}{\sigma U}\boldsymbol{f}$$

*in*

$$-2\hat{\nabla}\cdot\left(t^2\hat{\mathbb{D}}\left(\hat{\boldsymbol{u}}\right)\right)+\hat{\nabla}\hat{p}+\hat{\boldsymbol{u}}=\hat{\boldsymbol{f}}\qquad \text{in } \hat{\boldsymbol{\Omega}},$$

*where $t^2 := \frac{\mu_{\text{eff}}}{\sigma L^2}$ combines now all physical coefficients of the problem. On a first glance, only $t^2$ seems to contain the model parameters ($\mu_{\text{eff}}$ and $\sigma$) present in this equation, but the pressure and the right-hand side actually also contain $\sigma^{-1}$.*

Note that the process of non-dimensionalization might restrict the set of admissible coefficient values (see Rem. 3.2.2) and reduce the visibility of dependencies.

### The Brinkman Boundary Value Problem

To obtain a well-defined problem, the Brinkman equations have to be equipped with appropriate boundary conditions such that the resulting problem can be embedded into the theory of boundary value problems. A commonly used classification is that into Dirichlet and Neumann boundary conditions (or a combination, referred to as mixed or Robin boundary conditions). However, we will make use of the universal classification into *essential* and *natural boundary conditions*. This concept is directly related to the weak solution theory and will be presented in the corresponding Section 3.3 in detail. For the analysis presented in this thesis, we will focus on homogeneous essential boundary conditions in order to circumvent unnecessary technicalities.

**Assumption V:** Unless otherwise stated, we will concentrate on coefficients $\mu_{\text{eff}} \geqslant 0$ and $\sigma \geqslant 0$, that are constant throughout $\boldsymbol{\Omega}$ and scalar (see also Assumption III), noting that a generalization to tensors $\mu_{\text{eff}}\mathbb{I}$ and $\sigma\mathbb{I}$ is straightforward.

For the remainder of this thesis, we will focus on the **Brinkman problem with constant coefficients and homogeneous boundary conditions**:
For sufficiently smooth given data $\boldsymbol{f} : \boldsymbol{\Omega} \to \mathbb{R}^N$, $g : \boldsymbol{\Omega} \to \mathbb{R}$ find $\boldsymbol{u} : \boldsymbol{\Omega} \to \mathbb{R}^N$ and $p : \boldsymbol{\Omega} \to \mathbb{R}$ such that

$$-\mu_{\text{eff}}\Delta\boldsymbol{u}+\nabla p+\sigma\boldsymbol{u}=\boldsymbol{f}\qquad \text{in } \boldsymbol{\Omega}, \tag{3.2.8a}$$

$$\nabla\cdot\boldsymbol{u}=g\qquad \text{in } \boldsymbol{\Omega}, \tag{3.2.8b}$$

subject to the boundary conditions

$$\begin{aligned} \mu_{\text{eff}} > 0: &\quad \boldsymbol{u} = \boldsymbol{0} && \text{on } \Gamma, \\ \mu_{\text{eff}} = 0: &\quad \boldsymbol{u} \cdot \boldsymbol{n} = 0 && \text{on } \Gamma. \end{aligned} \tag{3.2.8c}$$

Let us point out the fact that the Brinkman equations with $\mu_{\text{eff}} > 0$ allow the imposition of boundary conditions for the velocity (normal and tangential component) on the boundary, e.g., the prominent no-slip boundary condition given above. Such boundary conditions are not compatible with the Darcy problem ($\mu_{\text{eff}} = 0$), where only the normal component of the velocity can be imposed.

In the following we will refer to the Brinkman problem as

- *Stokes limit* problem for $\sigma = 0$,
- *Stokes-type* problem or *Stokes case*[6] for $\mu_{\text{eff}} > 0$,
- *Stokes regime* problem for $\mu_{\text{eff}} \gg \sigma$,
- *Darcy regime* problem for $\mu_{\text{eff}} \ll \sigma$,
- *Darcy limit* problem for $\mu_{\text{eff}} = 0$.

## 3.3 Weak Formulations of the Brinkman Problem

The approach to solve weak formulations of systems of partial differential equations instead of their classical forms allows to also consider solutions, which fulfill the partial differential equations in a weak sense only, i.e., with reduced regularity assumptions compared to the classical solutions ([Bra07], [RR04]). Therefore, some problems which do not have a classical solution due to, e.g., discontinuous data (see, e.g., [ESW14, p. 13]), can be solved using the concept of weak solutions. Moreover, the weak formulation of a problem is the starting point for deriving a finite element formulation.

The weak form of a partial differential equation can be derived by multiplying it by a *test function*, integrating the result over the domain, and using integration by parts to transfer derivatives from the solution to the test functions. Hence, depending on the way integration by parts is applied, weak formulations that differ in the necessary regularity assumptions on the constituents may arise.

As we will see, the appropriate functional setting for the Brinkman problem changes when the second order derivative disappears. Therefore, the analysis will be performed for the limit cases separately.

### Classification of Boundary Conditions

As already mentioned, we will distinguish between *essential boundary conditions* and *natural boundary conditions*. This classification corresponds to the standard treatment of boundary conditions in the weak solution theory of boundary value

[6] The Stokes-type problem or Stokes case is associated with the functional setting which is for $\mu_{\text{eff}} > 0$ similar to that of a Stokes problem.

problems. Considering variational formulations, obtained using integration by parts, boundary integrals might be present. Natural boundary conditions are those which determine the value of a boundary term in the weak formulation without imposing conditions on the test and solution spaces. In contrast to that, essential boundary conditions enter the definition of the solution and test spaces and thus pose explicit conditions on the set of admissible functions.

Example: Let us briefly discuss the difference between natural and essential boundary conditions for the *Poisson problem* with homogeneous essential and inhomogeneous natural boundary conditions given by

$$-\Delta u = f \quad \text{in } \boldsymbol{\Omega}, \tag{3.3.1a}$$
$$u = 0 \quad \text{on } \Gamma_E, \tag{3.3.1b}$$
$$\nabla u \cdot \boldsymbol{n} = u_N \quad \text{on } \Gamma_N, \tag{3.3.1c}$$

where $\Gamma = \overline{\Gamma_E \cup \Gamma_N}$, $\Gamma_E \cap \Gamma_N = \emptyset$, and $\operatorname{meas}^{n-1}(\Gamma_E) > 0$, with the indices indicating the boundary types, $E$ for essential and $N$ for natural. Note that $\Gamma_E \cap \Gamma_N \neq \emptyset$ is in general not excluded, but not considered here.

Multiplying the partial differential equation (3.3.1a) by a test function $v$, then integrating over the domain $\boldsymbol{\Omega}$ and using integration by parts, we obtain

$$\begin{aligned}\int_{\boldsymbol{\Omega}} f v &= \int_{\boldsymbol{\Omega}} \nabla u \cdot \nabla v - \int_{\Gamma} (\nabla u \cdot \boldsymbol{n})\, v \\ &= \int_{\boldsymbol{\Omega}} \nabla u \cdot \nabla v - \int_{\Gamma_E} (\nabla u \cdot \boldsymbol{n})\, v - \int_{\Gamma_N} (\nabla u \cdot \boldsymbol{n})\, v. \end{aligned} \tag{3.3.2}$$

Defining the solution and test space as $V = H^1(\boldsymbol{\Omega})$ and assuming $f \in V'$ guarantees sufficient regularity such that the bulk integrals are well-defined. The boundary condition (3.3.1c) can be inserted into the weak formulation and is therefore called *natural boundary condition*. As a result, the integral on $\Gamma_N$ can be formally incorporated in the force term on the right-hand side[7]. The boundary condition (3.3.1b) cannot be inserted into the weak formulation (3.3.2) directly, since the condition $u|_{\Gamma_E} = 0$ does not determine the behavior of $(\nabla u \cdot \boldsymbol{n})|_{\Gamma_E}$. This condition is therefore called essential boundary condition. Explicit incorporation into the solution and test spaces by defining $V := \left\{ v \in H^1(\boldsymbol{\Omega}) : v|_{\Gamma_E} = 0 \right\}$ forces the test functions $v \in V$ (and the solution $u \in V$) to vanish on $\Gamma_E$ and hence the integral on $\Gamma_E$ in (3.3.2) also vanishes. Thus, (3.3.2) becomes

$$\int_{\boldsymbol{\Omega}} \nabla u \cdot \nabla v = \int_{\boldsymbol{\Omega}} f v + \int_{\Gamma_N} u_N\, v, \quad \forall v \in V.$$

### 3.3.1 A Weak Formulation for the Stokes Case

As long as $\mu_{\text{eff}} > 0$, the Brinkman problem (3.2.8) can be treated analogous to a Stokes problem with an additional zeroth order velocity term.

[7] In this dissertation, we refer to the right-hand side of a classical, weak, or discrete formulation as the side containing the given data, which usually coincides with the physical side. Note that in (3.3.2), the right-hand side (containing $f$) appears on the physical left side.

The weak formulation is derived using integration by parts for the Laplacian of the velocity and for the gradient of the pressure. As a result, the necessary regularity of the velocity solution $\boldsymbol{u}$ is reduced whereas the necessary regularity of the velocity test function is increased

$$-\int_{\boldsymbol{\Omega}} (\mu_{\text{eff}} \Delta \boldsymbol{u}) \cdot \boldsymbol{v} = \int_{\boldsymbol{\Omega}} \mu_{\text{eff}} \nabla \boldsymbol{u} : \nabla \boldsymbol{v} - \int_{\Gamma} \mu_{\text{eff}} (\nabla \boldsymbol{u} \cdot \boldsymbol{n}) \cdot \boldsymbol{v}. \tag{3.3.3}$$

Moreover, the transformation

$$\int_{\boldsymbol{\Omega}} \nabla p \cdot \boldsymbol{v} = -\int_{\boldsymbol{\Omega}} p (\nabla \cdot \boldsymbol{v}) + \int_{\Gamma} (p \boldsymbol{n}) \cdot \boldsymbol{v} = -\int_{\boldsymbol{\Omega}} (\nabla \cdot \boldsymbol{v}) p + \int_{\Gamma} p (\boldsymbol{v} \cdot \boldsymbol{n}) \tag{3.3.4}$$

can be used, such that the necessary regularity of the pressure $p$ reduces. Assuming $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{H}^1(\boldsymbol{\Omega})$ and $p, q \in L^2(\boldsymbol{\Omega})$, all bulk integrals are well-defined.

Due to the essential character of the boundary condition for the Brinkman problem (3.2.8) with $\mu_{\text{eff}} > 0$, namely

$$\boldsymbol{u} = \boldsymbol{0} \quad \text{on } \Gamma, \tag{3.3.5}$$

it is incorporated into the solution and test spaces for the velocity by using the space $\boldsymbol{H}_0^1(\boldsymbol{\Omega})$. For $\boldsymbol{v} \in \boldsymbol{H}_0^1(\boldsymbol{\Omega})$, the boundary integrals appearing in (3.3.3) and (3.3.4) vanish.

Since the Brinkman problem (3.2.8) depends only on the gradient of the pressure, $p$ is not uniquely determined (since the boundary conditions do not fix it). If $(\boldsymbol{u}, p)$ is a solution, then $(\boldsymbol{u}, p + c)$ is a solution for any constant $c \in \mathbb{R}$, too. Hence a condition that fixes the additive constant is necessary in order to enable the existence of a unique pressure solution. A common way to do that is by enforcing

$$\int_{\boldsymbol{\Omega}} p = 0. \tag{3.3.6}$$

Apparently, the boundary condition (3.3.5) does not concern the pressure, thus the condition (3.3.6) has to be incorporated into the solution (and test) spaces for the pressure, yielding $p, q \in L_0^2(\boldsymbol{\Omega})$.

By the Gaussian theorem, the boundary condition on the velocity and the right-hand side of the divergence constraint have to be compatible in the sense that

$$\int_{\boldsymbol{\Omega}} g = \int_{\boldsymbol{\Omega}} \nabla \cdot \boldsymbol{u} = \int_{\Gamma} \boldsymbol{u} \cdot \boldsymbol{n}. \tag{3.3.7}$$

Together with the boundary condition (3.3.5), this results in the compatibility condition

$$\int_{\boldsymbol{\Omega}} g = 0,$$

since $\boldsymbol{u}|_{\Gamma} = \boldsymbol{0} \Longrightarrow (\boldsymbol{u} \cdot \boldsymbol{n})|_{\Gamma} = 0$.

The resulting **weak formulation of the Brinkman problem for constant coefficients, $\mu_{\text{eff}} > 0$, and $\sigma \geqslant 0$** can be stated as follows.
Given $(\boldsymbol{f}, g) \in \boldsymbol{H}^{-1}(\boldsymbol{\Omega}) \times L_0^2(\boldsymbol{\Omega})$, find $(\boldsymbol{u}, p) \in \boldsymbol{H}_0^1(\boldsymbol{\Omega}) \times L_0^2(\boldsymbol{\Omega})$ such that

$$\mu_{\text{eff}}(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}) - (\nabla \cdot \boldsymbol{v}, p) + \sigma(\boldsymbol{u}, \boldsymbol{v}) = [\boldsymbol{f}, \boldsymbol{v}]_{\boldsymbol{H}^{-1}, \boldsymbol{H}_0^1}, \quad \forall \boldsymbol{v} \in \boldsymbol{H}_0^1(\boldsymbol{\Omega}), \tag{3.3.8a}$$

$$(\nabla \cdot \boldsymbol{u}, q) = (g, q), \qquad \forall q \in L_0^2(\boldsymbol{\Omega}). \tag{3.3.8b}$$

### Well-Posedness for $\mu_{\text{eff}} > 0$

In order to apply the existence and uniqueness theory presented in Subsection 3.1.3, the weak Brinkman problem (3.3.8) is embedded into the abstract framework as follows: Define

$$\boxed{\boldsymbol{V} \times Q := \boldsymbol{H}_0^1(\boldsymbol{\Omega}) \times L_0^2(\boldsymbol{\Omega})} \tag{3.3.9a}$$

and linear respectively bilinear forms

- $a(\cdot,\cdot) : \boldsymbol{V} \times \boldsymbol{V} \to \mathbb{R}$, $a(\boldsymbol{u}, \boldsymbol{v}) := \mu_{\text{eff}}(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}) + \sigma(\boldsymbol{u}, \boldsymbol{v})$,
- $f(\cdot) : \boldsymbol{V} \to \mathbb{R}$, $f(\boldsymbol{v}) := [\boldsymbol{f}, \boldsymbol{v}]_{V',V}$,
- $b(\cdot,\cdot) : \boldsymbol{V} \times Q \to \mathbb{R}$, $b(\boldsymbol{v}, p) := -(\nabla \cdot \boldsymbol{v}, p)$,
- $g(\cdot) : Q \to \mathbb{R}$, $g(q) := [g, q]_{Q',Q}$.

(3.3.9b)

With the definitions above, the problem (3.3.8) is equivalent to:
Find $(\boldsymbol{u}, p) \in \boldsymbol{V} \times Q$ such that

$$a(\boldsymbol{u}, \boldsymbol{v}) + b(\boldsymbol{v}, p) = f(\boldsymbol{v}), \quad \forall \boldsymbol{v} \in \boldsymbol{V}, \tag{3.3.9c}$$

$$b(\boldsymbol{u}, q) = g(q), \quad \forall q \in Q. \tag{3.3.9d}$$

**Theorem 3.3.1** (Well-Posedness in the Stokes Case)**.**
*The weak Brinkman problem* (3.3.9) *is well-posed for all* $\mu_{\text{eff}} > 0$.

*Proof:* Before we start, let us note that the bilinear forms $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ are still well-defined for the velocity space $\boldsymbol{H}^1(\boldsymbol{\Omega})$ and the pressure space $L^2(\boldsymbol{\Omega})$ such that one can also consider these supersets and possibly infer the validity of the arguments for $\boldsymbol{H}_0^1(\boldsymbol{\Omega}) \subsetneq \boldsymbol{H}^1(\boldsymbol{\Omega})$ and $L_0^2(\boldsymbol{\Omega}) \subsetneq L^2(\boldsymbol{\Omega})$.

The proof is based on Brezzi's splitting theorem (Thm. 3.1.17). Therefore we have to show that $a(\cdot,\cdot)$, $b(\cdot,\cdot)$, $f(\cdot)$, and $g(\cdot)$ are continuous/bounded (with the generic continuity constant denoted by $M$), that $a(\cdot,\cdot)$ is coercive on a special subset of $\boldsymbol{V}$ (with coercivity constant $m$), and that $b(\cdot,\cdot)$ satisfies the inf-sup condition (with inf-sup constant $\beta$).
(i) The continuity of $f(\cdot)$ and $g(\cdot)$ can be proven by applying the inequality for duality pairings (Thm. 3.1.7) resulting in

$$|f(\boldsymbol{v})| \leqslant \|\boldsymbol{f}\|_{\boldsymbol{V}'} \|\boldsymbol{v}\|_{\boldsymbol{V}} \qquad \text{and} \qquad |g(q)| \leqslant \|g\|_{Q'} \|q\|_{Q}.$$

Since $\boldsymbol{f}$ and $g$ are given data, choosing the continuity constants as $\|\boldsymbol{f}\|_{\boldsymbol{V}'}$ respectively $\|g\|_{Q'}$ is appropriate.

(ii) Using the triangle inequality and the Cauchy–Schwarz inequality (Thm. 3.1.5), followed by the binomial theorem, one obtains

$$\begin{aligned} |a(\boldsymbol{u},\boldsymbol{v})| &\leqslant \max\{\mu_{\text{eff}},\sigma\}\left(\|\nabla\boldsymbol{u}\|_0\|\nabla\boldsymbol{v}\|_0+\|\boldsymbol{u}\|_0\|\boldsymbol{v}\|_0\right) \\ &\leqslant \max\{\mu_{\text{eff}},\sigma\}\|\boldsymbol{u}\|_{\boldsymbol{H}^1(\boldsymbol{\Omega})}\|\boldsymbol{v}\|_{\boldsymbol{H}^1(\boldsymbol{\Omega})}, \qquad \forall \boldsymbol{u},\boldsymbol{v}\in\boldsymbol{H}^1(\boldsymbol{\Omega}). \end{aligned}$$

Thus, $M := \max\{\mu_{\text{eff}},\sigma\}$ yields continuity.[8] Continuity of $b(\cdot,\cdot)$ is proven by combining the Cauchy–Schwarz inequality (Thm. 3.1.5) with Lemma 3.1.4 as

$$\begin{aligned} |b(\boldsymbol{v},p)| &= |-(\nabla\cdot\boldsymbol{v},p)| \leqslant N^{\frac{1}{2}}\|\nabla\boldsymbol{v}\|_0\|p\|_0 \\ &\leqslant N^{\frac{1}{2}}\|\boldsymbol{v}\|_{\boldsymbol{H}^1(\boldsymbol{\Omega})}\|p\|_0, \qquad \forall(\boldsymbol{v},p)\in\boldsymbol{H}^1(\boldsymbol{\Omega})\times L^2(\boldsymbol{\Omega}). \end{aligned}$$

Hence $M := N^{\frac{1}{2}}$, where $N$ is the length of the vector $\boldsymbol{v}$, shows continuity.
(iii) Coercivity can be shown with $m := \min\{\mu_{\text{eff}},\sigma\} > 0$, as long as $\mu_{\text{eff}} > 0$ and $\sigma > 0$, since

$$a(\boldsymbol{u},\boldsymbol{u}) = \mu_{\text{eff}}\|\nabla\boldsymbol{u}\|_0^2+\sigma\|\boldsymbol{u}\|_0^2 \geqslant \min\{\mu_{\text{eff}},\sigma\}\|\boldsymbol{u}\|^2_{\boldsymbol{H}^1(\boldsymbol{\Omega})}, \quad \forall\boldsymbol{u}\in\boldsymbol{H}^1(\boldsymbol{\Omega}).$$

Using the fact that the semi-norm $|\cdot|_1$ is a norm on $\boldsymbol{H}^1_0$ (see Rem. 3.1.9), coercivity can be proven also for $\sigma = 0$, in which case the coercivity constant is $m := \mu_{\text{eff}}$.[9] Note that $\boldsymbol{v}|_\Gamma = \boldsymbol{0}$, i.e., $\boldsymbol{v}\in\boldsymbol{H}^1_0(\boldsymbol{\Omega})$ is a crucial ingredient for the given proof of $\sigma$-robust coercivity of $a(\cdot,\cdot)$.
(iv) Finally it remains to prove the inf-sup stability of $b(\cdot,\cdot)$ with respect to $\boldsymbol{V}$ and $Q$. Therefore, we apply Theorem 3.1.10: For any fixed $q\in L^2_0(\boldsymbol{\Omega})$ there exists a function $\widetilde{\boldsymbol{v}}\in\widetilde{\boldsymbol{V}_0}^{\perp}\subset\boldsymbol{H}^1_0(\boldsymbol{\Omega})$ such that

$$\nabla\cdot\widetilde{\boldsymbol{v}} = q \quad \text{and} \quad \|\widetilde{\boldsymbol{v}}\|_{\boldsymbol{H}^1(\boldsymbol{\Omega})} \leqslant C\|q\|_0,$$

for a constant $C > 0$. Thus, we can infer for $\boldsymbol{V}\times Q := \boldsymbol{H}^1_0(\boldsymbol{\Omega})\times L^2_0(\boldsymbol{\Omega})$ and any $q\in L^2_0(\boldsymbol{\Omega})$ that

$$\sup_{\boldsymbol{v}\in\boldsymbol{V}\setminus\{\boldsymbol{0}\}}\frac{(\nabla\cdot\boldsymbol{v},q)}{\|\boldsymbol{v}\|_{\boldsymbol{V}}} \geqslant \frac{(\nabla\cdot\widetilde{\boldsymbol{v}},q)}{\|\widetilde{\boldsymbol{v}}\|_{\boldsymbol{V}}} = \frac{(q,q)}{\|\widetilde{\boldsymbol{v}}\|_{\boldsymbol{V}}} = \frac{\|q\|_0^2}{\|\widetilde{\boldsymbol{v}}\|_{\boldsymbol{V}}} \geqslant \frac{1}{C}\frac{\|q\|_0^2}{\|q\|_0} = \frac{1}{C}\|q\|_0.$$

Since this is valid for any $q\in L^2_0(\boldsymbol{\Omega})$, the above estimate is also satisfied taking the infimum over all elements in $Q$, yielding

$$\inf_{q\in Q\setminus\{0\}}\left(\sup_{\boldsymbol{v}\in\boldsymbol{V}\setminus\{\boldsymbol{0}\}}\frac{b(\boldsymbol{v},q)}{\|\boldsymbol{v}\|_{\boldsymbol{V}}\|q\|_Q}\right) \geqslant \frac{1}{C} =: \beta.$$

□

**Remark 3.3.2** (Well-Posedness for a Deformation Rate Tensor Formulation)**.** *The proof of the existence and uniqueness of a solution to the more general problem in deformation tensor form, where* $a(\boldsymbol{u},\boldsymbol{v}) := \mu_{\text{eff}}(\mathbb{D}(\boldsymbol{u}),\mathbb{D}(\boldsymbol{v})) + \sigma(\boldsymbol{u},\boldsymbol{v})$,

[8] Note that $M > 0$, for all values of $\mu_{\text{eff}}$ and $\sigma$, especially for $\sigma = 0$, but the larger $\max\{\mu_{\text{eff}},\sigma\}$ becomes the weaker the estimate gets.
[9] This bound would degenerate if $\mu_{\text{eff}}$ would tend to zero.

*is very similar to that for the Laplacian formulation. The main difference is, that a Korn's inequality (see [Joh16, Lem. 3.37]) is used in order to bound the deformation rate tensor in the velocity by a multiple of the velocity gradient from below. Together with the arguments used in the proof of Theorem 3.3.1, this yields coercivity of* $a(\cdot,\cdot)$.

**Remark 3.3.3** (The Case of Non-Constant Coefficients)**.**
*The discussion could be generalized to non-constant coefficients assuming*

$$\begin{aligned} \mu_{\text{eff}} \in L^\infty(\boldsymbol{\Omega}) \quad & \text{and} \quad \exists\, \mu_0 > 0\colon \quad \mu_{\text{eff}}(\boldsymbol{x}) > \mu_0 > 0, \ \text{a.e. in } \boldsymbol{\Omega}, \\ \sigma \in L^\infty(\boldsymbol{\Omega}) \quad & \text{and} \quad \exists\, \sigma_0 > 0\colon \quad \sigma(\boldsymbol{x}) > \sigma_0 \geqslant 0, \ \text{a.e. in } \boldsymbol{\Omega}, \\ & \qquad\qquad\quad \text{(or a symm., pos. def. tensor)}, \end{aligned}$$

*and using the respective bounds in the analysis.*

### 3.3.2 A Weak Formulation for the Darcy Case

For $\mu_{\text{eff}} = 0$, the Brinkman equations reduce to the Darcy equations. The boundary condition considered in this situation is, as stated in (3.2.8), given by

$$\boldsymbol{u} \cdot \boldsymbol{n} = 0 \quad \text{on } \Gamma. \tag{3.3.10}$$

For the derivation of the weak formulation, the integration by parts formula can be used in the momentum balance equation for the transformation

$$\int_{\boldsymbol{\Omega}} \nabla p \cdot \boldsymbol{v} = -\int_{\boldsymbol{\Omega}} p\,(\nabla \cdot \boldsymbol{v}) + \int_{\Gamma} p\,(\boldsymbol{v} \cdot \boldsymbol{n}). \tag{3.3.11}$$

The velocity space $\boldsymbol{H}^{\text{div}}(\boldsymbol{\Omega})$ and the pressure space $L^2_0(\boldsymbol{\Omega})$ yield well-defined bulk integrals.

The boundary condition (3.3.10) is of essential type and is therefore incorporated into the solution and test space $\boldsymbol{H}^{\text{div}}_0(\boldsymbol{\Omega})$, such that the boundary integral in (3.3.11) vanishes.

Since the pressure is not explicitly restricted by the boundary condition, uniqueness is enforced here as well by the condition

$$\int_{\boldsymbol{\Omega}} p = 0.$$

Also in this case, the compatibility condition (3.3.7) has to be satisfied, enforcing

$$\int_{\boldsymbol{\Omega}} g = 0,$$

i.e., $g \in L^2_0(\boldsymbol{\Omega})$.

Altogether, we obtain the following **mixed weak formulation of the Brinkman problem** (3.2.8) **for constant coefficients,** $\mu_{\text{eff}} = 0$**, and** $\sigma > 0$:

Given $(\boldsymbol{f}, g) \in \left(\boldsymbol{H}_0^{\mathrm{div}}(\boldsymbol{\Omega})\right)' \times \left(L_0^2(\boldsymbol{\Omega})\right)'$, find $(\boldsymbol{u}, p) \in \boldsymbol{H}_0^{\mathrm{div}}(\boldsymbol{\Omega}) \times L_0^2(\boldsymbol{\Omega})$ such that

$$\sigma(\boldsymbol{u}, \boldsymbol{v}) - (\nabla \cdot \boldsymbol{v}, p) = [\boldsymbol{f}, \boldsymbol{v}]_{\left(\boldsymbol{H}_0^{\mathrm{div}}(\boldsymbol{\Omega})\right)', \boldsymbol{H}_0^{\mathrm{div}}(\boldsymbol{\Omega})}, \quad \forall \boldsymbol{v} \in \boldsymbol{H}_0^{\mathrm{div}}(\boldsymbol{\Omega}), \tag{3.3.12a}$$

$$(\nabla \cdot \boldsymbol{u}, q) = (g, q), \qquad \forall q \in L_0^2(\boldsymbol{\Omega}). \tag{3.3.12b}$$

This problem is also called *dual Darcy problem*, e.g., in [Cod15]. The definitions of the solution, test, and data spaces guarantee that all the terms in the weak formulation (3.3.12) are well-defined.

**Remark 3.3.4** (Other Weak Formulations of the Darcy Problem)**.**
*The formulation* (3.3.12) *is not the only weak problem that can be derived.*

*The integration by parts formula can be used in the equation modeling the conservation of mass yielding*

$$(\nabla \cdot \boldsymbol{u}, q) = -(\boldsymbol{u}, \nabla q) + \langle \boldsymbol{u} \cdot \boldsymbol{n}, q \rangle \overset{\boldsymbol{u} \cdot \boldsymbol{n}|_\Gamma \equiv 0}{=} -(\boldsymbol{u}, \nabla q).$$

*In this case, the boundary condition changes its role and becomes a natural boundary condition. The corresponding problem is referred to as primal (mixed) Darcy problem (see, e.g., [BC12]) and reads as:*
*Given* $(\boldsymbol{f}, g) \in \boldsymbol{V}' \times Q'$*, find* $(\boldsymbol{u}, p) \in \boldsymbol{V} \times Q$ *such that*

$$\sigma(\boldsymbol{u}, \boldsymbol{v}) + (\nabla p, \boldsymbol{v}) = [\boldsymbol{f}, \boldsymbol{v}]_{V', V}, \quad \forall \boldsymbol{v} \in \boldsymbol{V}, \tag{3.3.13a}$$

$$-(\boldsymbol{u}, \nabla q) = [g, q]_{Q', Q}, \quad \forall q \in Q, \tag{3.3.13b}$$

*with* $\boldsymbol{V} \times Q := \boldsymbol{L}^2(\boldsymbol{\Omega}) \times \left(H^1(\boldsymbol{\Omega}) \cap L_0^2(\boldsymbol{\Omega})\right)$.

*Let us instead consider a boundary condition on the pressure, e.g.,* $p|_\Gamma = 0$*, and assume that the data satisfies* $\boldsymbol{f} \in \boldsymbol{V}'$ *and* $g \in Q'$*. Then the integration by parts formula can be used in the equation modeling the conservation of mass similarly as above, yielding*

$$(\nabla \cdot \boldsymbol{u}, q) = -(\boldsymbol{u}, \nabla q) + \langle \boldsymbol{u} \cdot \boldsymbol{n}, q \rangle \overset{q|_\Gamma \equiv 0}{=} -(\boldsymbol{u}, \nabla q).$$

*This allows to derive the mixed Darcy problem* (3.3.13) *with* $\boldsymbol{V} \times Q := \boldsymbol{L}^2(\boldsymbol{\Omega}) \times H_0^1(\boldsymbol{\Omega})$*. Such formulations are considered, e.g., in [BGR08], [BC09], [BC12], and [JS10]. By the definition of the pressure space* $Q$*, the gradient of any pressure test function can be used as velocity test function, i.e., replacing* $\boldsymbol{v} := \nabla q$ *with the premise that* $q \in Q$*, the equations still have to be satisfied. Hence we get*

$$(\nabla p, \nabla q) + \sigma(\boldsymbol{u}, \nabla q) = (\boldsymbol{f}, \nabla q), \quad \forall q \in Q,$$
$$-(\boldsymbol{u}, \nabla q) = [g, q]_{Q', Q}, \quad \forall q \in Q.$$

*Inserting the second equation into the first one, we get the problem:*
*Find* $p \in H_0^1(\boldsymbol{\Omega})$ *such that*

$$(\nabla p, \nabla q) = (\boldsymbol{f}, \nabla q) + \sigma [g, q]_{Q', Q}, \quad \forall q \in H_0^1(\boldsymbol{\Omega}). \tag{3.3.15}$$

*Hence, the mixed formulation reduces to a Poisson problem for the pressure (for the given data regularities). Equation* (3.3.15) *can formally be obtained by*

*applying the (negative) divergence operator to the momentum balance of the Darcy equations, i.e.,*

$$-\nabla \cdot (\nabla p + \sigma \boldsymbol{u}) = -\nabla \cdot \boldsymbol{f},$$
$$\nabla \cdot \boldsymbol{u} = g.$$

*Note that this operation increases the regularity assumptions posed on the functions $p$ and $\boldsymbol{f}$ in the classical Darcy problem. Using the identity $(i)$ in Lemma 2.1.2 and inserting the divergence constraint yields*

$$-\Delta p = -\nabla \cdot \boldsymbol{f} + \sigma g. \tag{3.3.16}$$

*Note that* (3.3.15) *is a weak formulation of* (3.3.16) *with the boundary condition $p|_\Gamma = 0$. Once one has found a solution $p$ to* (3.3.16) *the corresponding velocity $\boldsymbol{u}$ can be computed by*

$$\sigma \boldsymbol{u} = \boldsymbol{f} - \nabla p. \tag{3.3.17}$$

*Applying the divergence operator and using* (3.3.16)*, it is straightforward that the velocity solution of* (3.3.17) *satisfies the divergence constraint, since*

$$\sigma \nabla \cdot \boldsymbol{u} = \nabla \cdot (\boldsymbol{f} - \nabla p) = \sigma g.$$

*Note that the number of unknowns in the variational problem has reduced here, such that only the pressure is determined in a first step. This approach is usually called pressure Poisson problem. As a post-processing step, the velocity is computed based on the pressure solution.*

*The appropriate weak form of the Darcy equations is a modeling choice, depending on the regularity of the given data.*

**Remark 3.3.5** (On the Primal and the Dual Darcy Problems)**.**
*The fulfillment of the inf-sup condition can be circumvented by invoking the conservation of mass equation in the momentum balance, such that a Poisson problem for the pressure has to be solved. Note that its approximate solution is a much simpler task because of the decoupling of the velocity and the pressure. However, in the literature it is argued that the pressure Poisson formulation is rather disadvantageous in practice, since the velocity is in general of primary interest, but might be discontinuous and is computed by essentially taking the gradient of the pressure multiplied by a possibly rough coefficient. This might result in a reduced accuracy of discrete solutions, see [MH02], [ML98], [BC12, Sect. 2.3.3], and [CJW14]. Also the conservation of mass is poor in practice, as mentioned in [BC10].*

*In this thesis we consider the Darcy problem as a singular limit of the Brinkman problem for $\mu_{\text{eff}} \to 0$, such that the naturally resulting system is of the form* (3.3.12)*, which motivates the focus on this formulation in the context of the Brinkman equations. This is also done in, e.g., [MH02] and [BC09].*

**Well-Posedness for $\mu_{\text{eff}} = 0$**

It remains to investigate well-posedness for $\mu_{\text{eff}} = 0$. Therefore, we again pose our problem in an adequate form to apply the abstract well-posedness theory of Subsection 3.1.3.

We state the weak Brinkman problem (3.3.12) in the form of a saddle point problem, i.e., we consider velocity and pressure spaces, $\boldsymbol{V}$ and $Q$ given by

$$\boxed{\boldsymbol{V} \times Q := \boldsymbol{H}_0^{\mathrm{div}}(\boldsymbol{\Omega}) \times L_0^2(\boldsymbol{\Omega})}, \tag{3.3.18a}$$

linear and bilinear forms defined by

$$\begin{aligned}
&\bullet\ a(\cdot,\cdot) : \boldsymbol{V} \times \boldsymbol{V} \to \mathbb{R}, && \bullet\ f(\cdot) : \boldsymbol{V} \to \mathbb{R}, \\
&\quad a(\boldsymbol{u},\boldsymbol{v}) := \sigma(\boldsymbol{u},\boldsymbol{v}), && \quad f(\boldsymbol{v}) := [\boldsymbol{f},\boldsymbol{v}]_{V',V}, \\
&\bullet\ b(\cdot,\cdot) : \boldsymbol{V} \times Q \to \mathbb{R}, && \bullet\ g(\cdot) : Q \to \mathbb{R}, \\
&\quad b(\boldsymbol{v},p) := -(\nabla\cdot\boldsymbol{v},p), && \quad g(q) := [g,q]_{Q',Q},
\end{aligned} \tag{3.3.18b}$$

and the following problem:
Find $(\boldsymbol{u},p) \in \boldsymbol{V} \times Q$ such that for given $(\boldsymbol{f},g) \in \boldsymbol{V}' \times Q'$ it holds

$$a(\boldsymbol{u},\boldsymbol{v}) + b(\boldsymbol{v},p) = f(\boldsymbol{v}), \quad \forall \boldsymbol{v} \in \boldsymbol{V}, \tag{3.3.18c}$$

$$b(\boldsymbol{u},q) = g(q), \quad \forall q \in Q. \tag{3.3.18d}$$

**Theorem 3.3.6** (Well-Posedness in the Darcy Case)**.**
*The weak Brinkman problem* (3.3.18) *is well-posed.*

*Proof:* Again we will make use of Brezzi's splitting theorem (Thm. 3.1.17). Hence, we have to show that $a(\cdot,\cdot)$, $b(\cdot,\cdot)$, $f(\cdot)$, and $g(\cdot)$ are continuous/bounded (with the generic continuity constant denoted by $M$), that $a(\cdot,\cdot)$ is coercive on a special subset of $\boldsymbol{V}$ (with coercivity constant $m$), and that $b(\cdot,\cdot)$ satisfies the inf-sup condition (with inf-sup constant $\beta$).
(i) The boundedness of $f(\cdot)$ and $g(\cdot)$ is immediate by the inequality for dual pairings (Thm. 3.1.7).
(ii) Continuity of $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ is obtained by using the Cauchy–Schwarz inequality (Thm. 3.1.5):

$$\begin{aligned}
|a(\boldsymbol{u},\boldsymbol{v})| &\leqslant \sigma\|\boldsymbol{u}\|_0\|\boldsymbol{v}\|_0 \leqslant \sigma\|\boldsymbol{u}\|_{\boldsymbol{H}^{\mathrm{div}}(\boldsymbol{\Omega})}\|\boldsymbol{v}\|_{\boldsymbol{H}^{\mathrm{div}}(\boldsymbol{\Omega})}, \quad \forall \boldsymbol{u},\boldsymbol{v} \in \boldsymbol{H}^{\mathrm{div}}(\boldsymbol{\Omega}), \\
|b(\boldsymbol{v},p)| &\leqslant \|\nabla\cdot\boldsymbol{v}\|_0\|p\|_0 \leqslant \|\boldsymbol{v}\|_{\boldsymbol{H}^{\mathrm{div}}(\boldsymbol{\Omega})}\|p\|_0, \quad \forall (\boldsymbol{v},p) \in \boldsymbol{H}^{\mathrm{div}}(\boldsymbol{\Omega}) \times L^2(\boldsymbol{\Omega}),
\end{aligned}$$

hence the continuity constant $M := \sigma$ is sufficient for $a(\cdot,\cdot)$ and $M := 1$ is appropriate for $b(\cdot,\cdot)$.
(iii) Coercivity of $a(\cdot,\cdot)$ on $\boldsymbol{V}_0 := \left\{\boldsymbol{v} \in \boldsymbol{H}_0^{\mathrm{div}}(\boldsymbol{\Omega}) \ :\ b(\boldsymbol{v},q) = 0, \forall q \in L_0^2(\boldsymbol{\Omega})\right\}$:

$$a(\boldsymbol{v},\boldsymbol{v}) = \sigma\|\boldsymbol{v}\|_0^2 = \sigma\|\boldsymbol{v}\|_{\boldsymbol{H}^{\mathrm{div}}(\boldsymbol{\Omega})}^2, \quad \forall \boldsymbol{v} \in \boldsymbol{V}_0.$$

Here, we used that $\boldsymbol{v} \in \boldsymbol{H}^{\mathrm{div}}(\boldsymbol{\Omega}) \Longrightarrow \nabla \cdot \boldsymbol{v} \in \boldsymbol{L}^2(\boldsymbol{\Omega})$ and the Gaussian theorem to deduce $\boldsymbol{v} \in \boldsymbol{H}_0^{\mathrm{div}}(\boldsymbol{\Omega}) \Longrightarrow (\nabla \cdot \boldsymbol{v}, 1) = 0$. Hence $\nabla \cdot \boldsymbol{v} \in L_0^2(\boldsymbol{\Omega})$ and we can choose $q := \nabla \cdot \boldsymbol{v}$ to obtain $0 = b(\boldsymbol{v}, q) = (\nabla \cdot \boldsymbol{v}, \nabla \cdot \boldsymbol{v}) = \|\nabla \cdot \boldsymbol{v}\|_0^2$. Thus, the coercivity constant $m := \sigma$ can be set.
(iv) For proving inf-sup stability of $b(\cdot,\cdot)$, Theorem 3.1.10 is used as follows: Since $\widetilde{\boldsymbol{V}_0}^{\perp} \subset \boldsymbol{H}_0^1(\boldsymbol{\Omega}) \subset \boldsymbol{H}_0^{\mathrm{div}}(\boldsymbol{\Omega})$ we can choose for any fixed $q \in L_0^2(\boldsymbol{\Omega})$ a corresponding $\widetilde{\boldsymbol{v}} \in \widetilde{\boldsymbol{V}_0}^{\perp} \subset \boldsymbol{H}_0^{\mathrm{div}}(\boldsymbol{\Omega})$ such that

$$\nabla \cdot \widetilde{\boldsymbol{v}} = q \qquad \text{and} \qquad \begin{aligned} \|\widetilde{\boldsymbol{v}}\|_{\boldsymbol{H}^{\mathrm{div}}(\boldsymbol{\Omega})} &\leqslant N^{\frac{1}{2}} \|\widetilde{\boldsymbol{v}}\|_{\boldsymbol{H}^1(\boldsymbol{\Omega})} \\ &\leqslant \left(1 + C_P^2\right)^{\frac{1}{2}} N^{\frac{1}{2}} \|\nabla \widetilde{\boldsymbol{v}}\|_0 \leqslant \widetilde{C} \|q\|_0, \end{aligned}$$

where Corollary 3.1.4, $1 \leqslant N$, Theorem 3.1.8, and $\widetilde{C} := C\left(1 + C_P^2\right)^{\frac{1}{2}} N^{\frac{1}{2}}$ were used. Consequently,

$$\begin{aligned} \sup_{\boldsymbol{v} \in \boldsymbol{V} \setminus \{\boldsymbol{0}\}} \frac{b(\boldsymbol{v}, q)}{\|\boldsymbol{v}\|_{\boldsymbol{V}}} &= \sup_{\boldsymbol{v} \in \boldsymbol{V} \setminus \{\boldsymbol{0}\}} \frac{(\nabla \cdot \boldsymbol{v}, q)}{\|\boldsymbol{v}\|_{\boldsymbol{V}}} \\ &\geqslant \frac{(\nabla \cdot \widetilde{\boldsymbol{v}}, q)}{\|\widetilde{\boldsymbol{v}}\|_{\boldsymbol{V}}} = \frac{(q, q)}{\|\widetilde{\boldsymbol{v}}\|_{\boldsymbol{V}}} \geqslant \frac{\|q\|_0^2}{\widetilde{C}\|q\|_0} = \frac{1}{\widetilde{C}} \|q\|_0, \end{aligned}$$

and $\beta = \frac{1}{\widetilde{C}}$ finishes the proof. □

**Remark 3.3.7** (On the Well-Posedness of the Primal (Mixed) Darcy Problem)**.**

*For the sake of completeness, let us secondly discuss the well-posedness of the problem* (3.3.13)*, i.e.,* $\boldsymbol{V} \times Q := \boldsymbol{L}^2(\boldsymbol{\Omega}) \times \left(H^1(\boldsymbol{\Omega}) \cap L_0^2(\boldsymbol{\Omega})\right)$*. Therefore we only have to redefine the bilinear form*

$$b(\boldsymbol{v}, p) := (\boldsymbol{v}, \nabla p).$$

*(i) The continuity of* $f(\cdot)$ *and* $g(\cdot)$ *is proven by the inequality for dual pairings (Thm. 3.1.7).*
*(ii) Continuity of* $a(\cdot,\cdot)$ *and* $b(\cdot,\cdot)$ *are immediate since the Cauchy–Schwarz inequality (Thm 3.1.5) gives*

$$\begin{aligned} &|a(\boldsymbol{u}, \boldsymbol{v})| \leqslant \sigma \|\boldsymbol{u}\|_0 \|\boldsymbol{v}\|_0, \quad \forall \boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{L}^2(\boldsymbol{\Omega}), \\ &|b(\boldsymbol{v}, p)| = |(\boldsymbol{v}, \nabla p)| \leqslant \|\boldsymbol{v}\|_0 \|\nabla p\|_0 \leqslant \|\boldsymbol{v}\|_0 \|p\|_{H^1(\boldsymbol{\Omega})}, \quad \forall (\boldsymbol{v}, p) \in \boldsymbol{L}^2(\boldsymbol{\Omega}) \times H^1(\boldsymbol{\Omega}), \end{aligned}$$

*hence the continuity constants* $M := \sigma$ *and* $M := 1$ *yield the desired result for* $a(\cdot,\cdot)$ *and* $b(\cdot,\cdot)$*, respectively.*
*(iii) Coercivity of* $a(\cdot,\cdot)$ *is valid since*

$$a(\boldsymbol{u}, \boldsymbol{u}) = |\sigma(\boldsymbol{u}, \boldsymbol{u})| = \sigma \|\boldsymbol{u}\|_0^2, \quad \forall \boldsymbol{u} \in \boldsymbol{L}^2(\boldsymbol{\Omega}),$$

*which gives* $m := \sigma$*.*
*(iv) Inf-sup stability of* $b(\cdot,\cdot)$ *with respect to* $\boldsymbol{V}$ *and* $Q$ *is shown by considering* $\hat{\boldsymbol{v}} := \nabla p \in \boldsymbol{L}^2(\boldsymbol{\Omega})$*, then*

$$\sup_{\boldsymbol{v} \in \boldsymbol{V} \setminus \{\boldsymbol{0}\}} \frac{b(\boldsymbol{v}, q)}{\|\boldsymbol{v}\|_{\boldsymbol{V}}} = \sup_{\boldsymbol{v} \in \boldsymbol{V} \setminus \{\boldsymbol{0}\}} \frac{(\boldsymbol{v}, \nabla p)}{\|\boldsymbol{v}\|_{\boldsymbol{V}}} \geqslant \frac{(\hat{\boldsymbol{v}}, \nabla p)}{\|\hat{\boldsymbol{v}}\|_{\boldsymbol{V}}} = \frac{\|\nabla p\|_0^2}{\|\nabla p\|_0} = \|\nabla p\|_0 \geqslant C \|p\|_{H^1},$$

*where the last inequality is justified by a Poincaré inequality for functions with zero averages (see [Eva10, Sect. 5.8.1]).*

## 3.4 Summary

The central topic of this chapter was the weak formulation of the Brinkman problem. To that end we have introduced the Sobolev spaces, which generalize the classical definition of derivatives to weak derivatives in the sense of $L^2$. Further, a collection of statements from functional analysis was presented, which will also be partly used in the upcoming chapters. We studied abstract frameworks and corresponding well-posedness results, that can be applied to weak problems.

Coming back to the Brinkman problem, the analysis started with a derivation of a dimensionless form and a discussion of appropriate boundary conditions, including their classification into essential and natural boundary conditions. We identified the different character of the Brinkman momentum equation for $\mu_{\text{eff}} > 0$ (second-order differential equation in $\boldsymbol{u}$) and for $\mu_{\text{eff}} = 0$ (0th order differential equation in $u$). In fact, the regularity of the velocity functions and the boundary conditions, which are in either case of essential type, change when $\mu_{\text{eff}}$ approaches zero. This led to a separate analysis of the Stokes case ($\mu_{\text{eff}} > 0$) and the Darcy limit ($\mu_{\text{eff}} = 0$).

Finally, we have provided a proof of the well-posedness of the weak Brinkman problem in the Stokes case for $\boldsymbol{V} \times Q = \boldsymbol{H}_0^1(\boldsymbol{\Omega}) \times L_0^2(\boldsymbol{\Omega})$ and in the Darcy limit, well-posedness was shown for $\boldsymbol{V} \times Q = \boldsymbol{H}_0^{\text{div}}(\boldsymbol{\Omega}) \times L_0^2(\boldsymbol{\Omega})$.

# 4 Aspects about the Finite Element Method for the Brinkman Problem

The goal of this chapter is to discuss the relevant preliminary ingredients of a low-order finite element method for the Brinkman problem, which is robust with respect to the values of the physical parameters, yielding a single well-posed approach that is applicable in the whole physical range (including the Stokes and Darcy limits).

After a few finite element basics, the standard Galerkin formulation of the Brinkman problem will be introduced. On the basis of the separate weak formulations analyzed in Chapter 3, typical finite element pairs for the Stokes (respectively Darcy) problem are discussed, studying the difficulties related to their applicability in the Darcy (respectively Stokes) limit.

From the computational perspective, low-order methods are in general particularly attractive, but might suffer from stability problems. Fortunately several techniques exist which allow to repair these instabilities. Among them, we will focus on continuous equal-order interpolations of the velocity and the pressure and present - along with the original idea of pressure stabilizations - the non-symmetric GLS (Douglas–Wang) method.

Further, pursuing the goal to develop a robust finite element formulation for the Brinkman problem, it remains to appropriately incorporate the essential boundary conditions that differ in the Brinkman limits. Opposed to the strong imposition, an inherent parameter-dependent transition is desired, which can be achieved with the weak imposition of boundary conditions. To this purpose, we will introduce the penalty-free non-symmetric Nitsche method.

## 4.1 The Finite Element Method in a Nutshell

The finite element method aims at approximately solving a system of partial differential equations by seeking the solution in finite dimensional function spaces. The principal components of the method are the discretization of the domain and the definition of appropriate finite-dimensional approximation spaces. As a result one obtains a linear system of equations whose solvability has to be analyzed.

This section is used to briefly present basic aspects of the finite element method. For a detailed introduction to the topic, the books [BBF13], [EG04], [Cia02], [GR86], [BS08], [Joh16], and [BF91] are recommended.

### The Discretization of the Domain and the Discrete Spaces

Let us denote by $\boldsymbol{\Omega} \subset \mathbb{R}^n$, $n = 2, 3$, the computational domain and by $\mathcal{T}_h$ a so-called *admissible triangulation* or *mesh*, consisting of $|I|$ polygonal (for $n = 2$) or polyhedral (for $n = 3$) elements $\boldsymbol{T} \in \mathcal{T}_h$, i.e., for $\boldsymbol{T}_i, \boldsymbol{T}_j \in \mathcal{T}_h$, $i \neq j$, we have

$$\mathring{\boldsymbol{T}}_i \cap \mathring{\boldsymbol{T}}_j = \emptyset \qquad \text{and} \qquad \bigcup_{i \in I} \boldsymbol{T}_i = \overline{\boldsymbol{\Omega}},$$

with $I$ being some index set with cardinality $|I|$, see [Cia02, p. 38].[1] A generic facet ($(n-1)$-dimensional face) of an $n$-dimensional *mesh element* $\boldsymbol{T}$ (also called *mesh cell*) is denoted by $E$. In 2D, $E$ refers to edges, as illustrated in Figure 4.1.

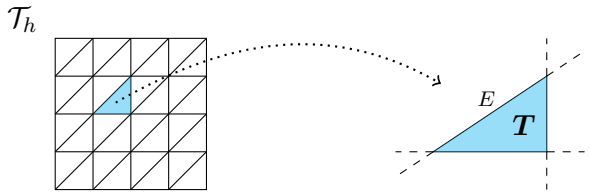



Figure 4.1: A (uniform) triangulation $\mathcal{T}_h$ of a square domain into triangles (left) and a single triangular element $\boldsymbol{T}$ with an edge $E$ (right).

The index $h > 0$ refers to the *mesh-size*, i.e., the maximum element diameter in the considered triangulation

$$h := \max_{\boldsymbol{T} \in \mathcal{T}_h} h_{\boldsymbol{T}}, \qquad \text{with} \qquad h_{\boldsymbol{T}} := \max_{\boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{T}} \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|,$$

with $\|\cdot\|$ being the Euclidean norm (3.1.1) in $\mathbb{R}^n$. Similarly, the diameter of a facet $E$ is defined as

$$h_E := \max_{\boldsymbol{x}_1, \boldsymbol{x}_2 \in E} \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|.$$

The set of boundary facets is referred to as

$$\mathcal{G}_h := \{E \in \mathcal{T}_h \,:\, E \subset \Gamma\}\,. \tag{4.1.1}$$

With the notation above, we introduce the following common terminology.
A family of triangulations $\{\mathcal{T}_h\}_{h>0}$ of $\boldsymbol{\Omega}$ with $h \leqslant \operatorname{diam}(\boldsymbol{\Omega})$ is called:

1. **(shape-)regular/non-degenerate**, if there exists a constant $C_{SR} > 0$ for the family $\{\mathcal{T}_h\}_{h>0}$ such that

$$\forall h > 0,\ \forall \boldsymbol{T} \in \mathcal{T}_h : \quad \frac{h_{\boldsymbol{T}}}{\rho_{\boldsymbol{T}}} \leqslant C_{SR}, \tag{4.1.2}$$

   where $\rho_{\boldsymbol{T}}$ is the radius of the largest inscribed sphere in $\boldsymbol{T}$, see [Cia02, p. 124] or [EG04, Def. 1.107],

[1] This implies that $\overline{\boldsymbol{\Omega}}$ is polygonal (in 2D) or polyhedral (3D).

2. **quasi-uniform**, if there exists a constant $C_{QU} > 0$ such that

$$\forall h > 0 : \quad \frac{h}{\min\limits_{\boldsymbol{T} \in \mathcal{T}_h} \rho_{\boldsymbol{T}}} \leqslant C_{QU},$$

see [BS08, Def. 4.4.13].

Note that quasi-uniformity implies (shape-)regularity with $C_{SR} \leqslant C_{QU}$. Moreover, the constant $C_{SR}$ is an indicator for how much the elements of the triangulation differ from equilateral elements.

On $\boldsymbol{T} \in \mathcal{T}_h$, one can define polynomial spaces

$$\mathbb{P}_k(\boldsymbol{T}) := \{\text{all real polynomials of degree} \leqslant k \text{ on } \boldsymbol{T}\}, \quad k \geqslant 0.$$

These are used to create the finite element spaces which contain functions that are element-wise ($\boldsymbol{T}$-wise) polynomials and possibly continuous or discontinuous across (inner) facets. For a generic infinite-dimensional function space $\boldsymbol{V}$ on $\boldsymbol{\Omega}$, we can define a finite-dimensional approximation space by

$$\boldsymbol{V}_h^k := \left\{ \boldsymbol{v}_h \in \boldsymbol{L}^2(\boldsymbol{\Omega}) \ : \ \boldsymbol{v}_h|_{\boldsymbol{T}} \in [\mathbb{P}_k(\boldsymbol{T})]^n, \quad \forall \boldsymbol{T} \in \mathcal{T}_h \right\} \cap \boldsymbol{V}.$$

Note that the conditions $\boldsymbol{v}_h \in \boldsymbol{C}^0(\overline{\boldsymbol{\Omega}})$ and $\boldsymbol{v}_h|_{\boldsymbol{T}} \in \boldsymbol{H}^1(\boldsymbol{T})$, for all $\boldsymbol{T} \in \mathcal{T}_h$, imply $\boldsymbol{v}_h \in \boldsymbol{H}^1(\boldsymbol{\Omega})$ (see [Cia02, Thm. 2.1.1]), which is often used (without explicit note) for the definition of continuous finite element spaces contained in $\boldsymbol{H}^1(\boldsymbol{\Omega})$.[2]
The dimension $\dim\left(\boldsymbol{V}_h^k\right)$ equals the number of basis functions of $\boldsymbol{V}_h^k$, whose linear combination represents any element of $\boldsymbol{V}_h^k$. In the following discussions we will consider *Lagrange finite element spaces*, where *nodal basis* functions (see, e.g., [EG04, Def. 1.27]) are used and therefore the *degrees of freedom (dofs)* - i.e., the information which uniquely determines a function in $\boldsymbol{V}_h^k$ - are associated with $\dim\left(\boldsymbol{V}_h^k\right)$ point evaluations (at *nodes* of the mesh), see, e.g., [Cia02]. Further, we will abuse the same notation $\mathbb{P}_k(\cdot) = [\mathbb{P}_k(\cdot)]^n$, assuming that the meaning is clear from the context.

One main advantage of polynomial finite element spaces is that valuable results for polynomials can be transfered element-wise to finite element functions: Higher order derivatives can be estimated by lower order derivatives using the *inverse inequalities* or *inverse estimates*.

Under the assumption of shape regularity (and assuming $h \leqslant 1$), there exist constants $c_I, \hat{c}_I > 0$, independent of $h$ and $\boldsymbol{T}$, such that for all $v_h \in \mathbb{P}_k(\boldsymbol{T})$, $k \geqslant 0$, and for all $\boldsymbol{T} \in \mathcal{T}_h$, the following local **inverse inequalities** hold

$$\|\nabla v_h\|_{0,\boldsymbol{T}} \leqslant c_I h_{\boldsymbol{T}}^{-1} \|v_h\|_{0,\boldsymbol{T}}, \tag{4.1.3a}$$

$$\|\Delta v_h\|_{0,\boldsymbol{T}} \leqslant \hat{c}_I h_{\boldsymbol{T}}^{-1} \|\nabla v_h\|_{0,\boldsymbol{T}}. \tag{4.1.3b}$$

[2] Here $\boldsymbol{C}^0(\overline{\boldsymbol{\Omega}})$ denotes continuous function on $\overline{\boldsymbol{\Omega}}$.

The proofs rely on the equivalence of all norms on finite-dimensional spaces and the so-called *scaling arguments*, which refer to mapping the physical cell $\boldsymbol{T}$ (and the function $\boldsymbol{v}_h$) to a fixed reference element and back.[3] For details see, e.g., [EG04, Lem. 1.138], [BZ12, p. 233], [Arn82], [Cia02, Thm. 3.2.6], [FHS93], [Bra07, II §6.8], [QV08, (8.4.7)], [RWG01, Lem. 2.1], [BH06a], [BC09], and [Joh16, Thm. C.30].

Furthermore, the following **discrete trace inequality** holds:

$$\|v\|_{0,\partial\boldsymbol{T}}^2 \leqslant \tilde{c}_{\mathrm{DT}} \left( h_{\boldsymbol{T}}^{-1} \|v\|_{0,\boldsymbol{T}}^2 + h_{\boldsymbol{T}} \|\nabla v\|_{0,\boldsymbol{T}}^2 \right), \quad \forall v \in H^1(\boldsymbol{T}), \tag{4.1.4}$$

where $\tilde{c}_{\mathrm{DT}} > 0$ is a constant, only depending on the shape regularity (4.1.2) of the mesh.

For a facet $E$ of the mesh and an attached mesh cell $\boldsymbol{T}_E$, i.e., $E \subsetneq \partial\boldsymbol{T}_E$, the discrete trace inequality (4.1.4) and shape-regularity (4.1.2) imply an estimate of the form

$$h_E^{-1} \|v\|_{0,E}^2 \leqslant c_{\mathrm{DT}} \left( h_{\boldsymbol{T}_E}^{-2} \|v\|_{0,\boldsymbol{T}_E}^2 + \|\nabla v\|_{0,\boldsymbol{T}_E}^2 \right), \quad \forall v \in H^1(\boldsymbol{T}_E). \tag{4.1.5}$$

Details concerning the derivation of (4.1.4) and (4.1.5) can be found in, e.g., [BH06a], [Tho06, pp. 27], [Arn82, (2.4)], [BC09], [BS08, (10.3.8)], [Clé75, Lem. 4], and [RWG01, (2.5)].

Combining (4.1.3) with (4.1.4) we can also conclude that there exist constants $c_{\mathrm{DTI}}, \widehat{c}_{\mathrm{DTI}}, \widetilde{c}_{\mathrm{DTI}} > 0$ such that, for any element-wise polynomial function $v_h$, i.e., $v_h|_{\boldsymbol{T}} \in \mathbb{P}_k(\boldsymbol{T}), \forall \boldsymbol{T} \in \mathcal{T}_h$, and $E \in \partial\boldsymbol{T}_E$ it holds:

$$\|v_h\|_{0,E}^2 \leqslant c_{\mathrm{DTI}}\, h_{\boldsymbol{T}_E}^{-1} \|v_h\|_{0,\boldsymbol{T}_E}^2, \tag{4.1.6a}$$

$$h_E \|\nabla v_h \cdot \boldsymbol{n}_E\|_{0,E}^2 \leqslant \widehat{c}_{\mathrm{DTI}} \|\nabla v_h\|_{0,\boldsymbol{T}_E}^2, \tag{4.1.6b}$$

$$\sum_{E \in \mathcal{G}_h} h_E \|\nabla v_h \cdot \boldsymbol{n}_E\|_{0,E}^2 \leqslant \widetilde{c}_{\mathrm{DTI}} \|\nabla v_h\|_{0,\boldsymbol{\Omega}}^2, \tag{4.1.6c}$$

see also [JS09, Lem. 3.1], [Tho06, Lem. 2.1], and [Ste95, Lem. 3].

### The Galerkin Method and the Linear System of Equations

Let us write the problem (3.1.2) as:
Find $\boldsymbol{u} \in \boldsymbol{V}$ such that

$$a(\boldsymbol{u}, \boldsymbol{v}) = f(\boldsymbol{v}), \quad \forall \boldsymbol{v} \in \boldsymbol{V}. \tag{4.1.7}$$

The *standard Galerkin approach* consists in substituting the infinite-dimensional space $\boldsymbol{V}$ by a finite-dimensional space $\boldsymbol{V}_h$ and seeking an approximation of $\boldsymbol{u}$ in problem (4.1.7) as the solution of the discrete problem:
Find $\boldsymbol{u}_h \in \boldsymbol{V}_h$ such that

$$a_h(\boldsymbol{u}_h, \boldsymbol{v}_h) = f_h(\boldsymbol{v}_h), \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{4.1.8}$$

[3] Such reference transformations will be discussed in Section 5.1 in more detail.

where $a_h(\cdot,\cdot)$ and $f_h(\cdot)$ are approximations to $a(\cdot,\cdot)$ and $f(\cdot)$, respectively. In case $\boldsymbol{V}_h \subset \boldsymbol{V}$, the resulting method is called *conforming* (and consequently $a_h(\cdot,\cdot)$ and $f_h(\cdot)$ coincide with $a(\cdot,\cdot)$ and $f(\cdot)$, respectively), otherwise *non-conforming*.[4] In general, the terminology *ansatz space* is associated with the space for $\boldsymbol{u}_h$ and *test space* is used for the space associated with $\boldsymbol{v}_h$ (although they currently coincide in (4.1.8)).

The Galerkin method yields a linear system of equations in the following way: Let $\{\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_N\}$ be a basis of $\boldsymbol{V}_h$, then

$$\boldsymbol{u}_h = \sum_{i=1}^{N} U_i \boldsymbol{\psi}_i,$$

for some $U_i \in \mathbb{R}$.[5] Testing the problem (4.1.8) with all basis functions, inserting the basis representation of $\boldsymbol{u}_h$, and using the bilinearity of $a_h(\cdot,\cdot)$, the problem can be rewritten as

$$\sum_{i=1}^{N} U_i \, a_h\left(\boldsymbol{\psi}_i, \boldsymbol{\psi}_j\right) = f_h\left(\boldsymbol{\psi}_j\right), \qquad \text{for } j = 1, \ldots, N.$$

This is a linear system

$$\boldsymbol{A}\boldsymbol{U} = \boldsymbol{F},$$

where

- $\boldsymbol{A} \in \mathbb{R}^{N \times N}$ with $(\boldsymbol{A})_{ji} := a_h\left(\boldsymbol{\psi}_i, \boldsymbol{\psi}_j\right)$, for $i, j = 1, \ldots, N$,
- $\boldsymbol{U} \in \mathbb{R}^N$ with $\boldsymbol{U} := (U_1, \ldots, U_N)^T$,
- $\boldsymbol{F} \in \mathbb{R}^N$ with $F_j := f_h\left(\boldsymbol{\psi}_j\right)$, for $j = 1, \ldots, N$.

In practice, it is essential to choose a basis such that the corresponding matrix $\boldsymbol{A}$ has few nonzero entries, i.e., $\boldsymbol{A}$ is sparse. This is guaranteed by using basis functions with local support (e.g., a nodal basis). Integral expressions in the matrix $\boldsymbol{A}$ and in the vector $\boldsymbol{F}$ can be approximated using numerical integration.

### Well-Posedness

Let us start by introducing two finite-dimensional Hilbert spaces $(\boldsymbol{V}_h, \|\cdot\|_{\boldsymbol{V}_h})$ and $(Q_h, \|\cdot\|_{Q_h})$, two bilinear forms $a_h(\cdot,\cdot) : \boldsymbol{V}_h \times \boldsymbol{V}_h \longrightarrow \mathbb{R}$ and $b_h(\cdot,\cdot) : \boldsymbol{V}_h \times Q_h \longrightarrow \mathbb{R}$, and two linear functionals $f_h(\cdot) : \boldsymbol{V}_h \longrightarrow \mathbb{R}$ and $g_h : Q_h \longrightarrow \mathbb{R}$. Note that linear functionals and bilinear forms on finite-dimensional Hilbert spaces are generally bounded.

In Subsection 3.1.3, we have identified the coercivity and the inf-sup condition as crucial properties for the investigation of the well-posedness in the weak

[4] Common non-conforming approaches are discontinuous Galerkin methods.
[5] The unknowns $U_i$ are the degrees of freedom.

problem setup. A reasonable question is whether certain properties inherit from the weak to the discrete problem (4.1.8).

For conforming methods, $\boldsymbol{V}_h \subset \boldsymbol{V}$, well-posedness of the weak problem in the sense of Lax–Milgram (Thm. 3.1.12), i.e. involving the whole space $\boldsymbol{V}$, implies a well-posed discrete problem (4.1.8). This is true because in particular coercivity remains valid when restricting the underlying space, see, e.g., [BS08, Thm. 2.7.7] and [EG04, Prop. 2.19]. In contrast to that, the inf-sup condition is in general not inherited, even not in the conforming case.

Dealing with non-conforming methods, with problems for which coercivity is difficult to prove, or even with problems which do not satisfy coercivity at all (it is only sufficient for well-posedness, see Subsection 3.1.3), a discrete version of the BNB theorem (Thm. 3.1.13) can be used:

**Theorem 4.1.1** (Discrete Generalized Lax–Milgram/BNB)**.**
*The problem* (4.1.8) *is well-posed, if and only if the discrete inf-sup condition is satisfied, i.e.,*

$$\exists\, \beta_h > 0 : \qquad \inf_{\boldsymbol{u}_h \in \boldsymbol{V}_h \setminus \{\boldsymbol{0}\}} \left( \sup_{\boldsymbol{v}_h \in \boldsymbol{V}_h \setminus \{\boldsymbol{0}\}} \frac{a_h\left(\boldsymbol{u}_h, \boldsymbol{v}_h\right)}{\|\boldsymbol{u}_h\|_{\boldsymbol{V}_h} \|\boldsymbol{v}_h\|_{\boldsymbol{V}_h}} \right) \geqslant \beta_h. \tag{4.1.9}$$

*Proof:* For a proof based on the rank theorem combined with the BNB theorem (Thm. 3.1.13) see [EG04, Sect. 2.2.3] and for a direct argumentation see [Joh16, Lem. B.15]. □

We can also exploit the saddle point structure in the discrete case and thus formulate a discrete analog of Theorem 3.1.17.

**Theorem 4.1.2** (Discrete Version of Brezzi's Splitting)**.**
*If*

*(i)* $a_h\left(\cdot,\cdot\right)$ *is coercive on* $\boldsymbol{V}_{0,h}$, (4.1.10a)
*with* $\boldsymbol{V}_{0,h} := \{\boldsymbol{v}_h \in \boldsymbol{V}_h \,:\, b_h\left(\boldsymbol{v}_h, q_h\right) = 0, \quad \forall q_h \in Q_h\}$,

*(ii)* $b_h\left(\cdot,\cdot\right)$ *satisfies the discrete inf-sup condition*

$$\exists\ \beta_h > 0 : \qquad \inf_{q_h \in Q_h \setminus \{0\}} \left( \sup_{\boldsymbol{v}_h \in \boldsymbol{V}_h \setminus \{\boldsymbol{0}\}} \frac{b_h\left(\boldsymbol{v}_h, q_h\right)}{\|\boldsymbol{v}_h\|_{\boldsymbol{V}_h} \|q_h\|_{Q_h}} \right) \geqslant \beta_h, \tag{4.1.10b}$$

*then the discrete problem:*
*Find* $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h \times Q_h$ *such that*

$$\begin{aligned} a_h\left(\boldsymbol{u}_h, \boldsymbol{v}_h\right) + b_h\left(\boldsymbol{v}_h, p_h\right) &= f_h\left(\boldsymbol{v}_h\right), \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \\ b_h\left(\boldsymbol{u}_h, q_h\right) &= g_h\left(q_h\right), \quad \forall q_h \in Q_h \end{aligned}$$

*is well-posed.*

*Proof:* The proof is a consequence of Brezzi's splitting theorem (Thm. 4.1.2). □

Pairs of finite element spaces $\boldsymbol{V}_h \times Q_h$ that satisfy the inf-sup condition (4.1.9) (respectively (4.1.10b)) are called (inf-sup) *stable finite element pairs*.

**Remark 4.1.3** (On the Discrete Inf-Sup Constant)**.**
*As it will be seen in Section 5.3, the inverse of the discrete inf-sup constant, $\beta_h^{-1}$, is a factor in the a priori finite element error estimate. In order to obtain optimal convergence, the inf-sup constant $\beta_h$ should therefore be bounded away from zero, independent from the mesh-size $h$. For more details see [Joh16, Rem. 3.57].*

**Remark 4.1.4** (On the Discrete Version of Brezzi's Splitting Theorem)**.**
*The inf-sup condition* (4.1.10b) *is a compatibility condition relating the finite element spaces $\boldsymbol{V}_h$ and $Q_h$. An interesting observation is that the conditions* (4.1.10a) *and* (4.1.10b) *indicate in some sense opposite tendencies concerning an appropriate choice of $\boldsymbol{V}_h$ and $Q_h$. The discrete inf-sup condition* (4.1.10b) *can be equivalently written as*

$$\sup_{\boldsymbol{v}_h \in \boldsymbol{V}_h \setminus \{\boldsymbol{0}\}} \frac{b_h(\boldsymbol{v}_h, q_h)}{\|\boldsymbol{v}_h\|_{\boldsymbol{V}_h}} \geqslant \beta_h \|q_h\|_{Q_h}, \quad \forall q_h \in Q_h.$$

*Enlarging the space $\boldsymbol{V}_h$ relative to $Q_h$ can only increase the supremum, thus, this describes a promising strategy to enforce the fulfillment of* (4.1.10b)*. On the other hand, condition* (4.1.10a) *demands for $\boldsymbol{V}_{0,h}$ being small enough which means that $Q_h$ is large enough relative to $\boldsymbol{V}_h$.*

*Apart from appropriately balancing the finite element spaces with respect to each other, there exist possibilities that allow to circumvent the above discrete inf-sup condition as we will see in Subsection 4.3.2.*

## 4.2 The Standard Galerkin Approach for the Brinkman Problem

Let us now apply the Galerkin method in order to formulate a discrete Brinkman problem and discuss its properties. To that end, we denote by $\boldsymbol{V}_h$ a finite-dimensional approximation space for the velocity and by $Q_h$ a finite-dimensional approximation space for the pressure such that the **standard Galerkin method for the Brinkman problem** is given by:
Find $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h \times Q_h$ such that

$$\mathcal{A}_h^{(\pm)}\left[(\boldsymbol{u}_h, p_h); (\boldsymbol{v}_h, q_h)\right] = \mathcal{F}_h^{(\pm)}\left[(\boldsymbol{v}_h, q_h)\right], \quad \forall (\boldsymbol{v}_h, q_h) \in \boldsymbol{V}_h \times Q_h, \tag{4.2.1a}$$

with

$$\mathcal{A}_h^{(\pm)}\left[(\boldsymbol{u}_h, p_h); (\boldsymbol{v}_h, q_h)\right] := a_h(\boldsymbol{u}_h, \boldsymbol{v}_h) - b_h(\boldsymbol{v}_h, p_h) \pm b_h(\boldsymbol{u}_h, q_h), \tag{4.2.1b}$$

$$\mathcal{F}_h^{(\pm)}\left[(\boldsymbol{v}_h, q_h)\right] := (\boldsymbol{f}, \boldsymbol{v}_h) \pm (g, q_h), \tag{4.2.1c}$$

and

$$a_h(\boldsymbol{u}_h, \boldsymbol{v}_h) := \mu_{\text{eff}}(\nabla \boldsymbol{u}_h, \nabla \boldsymbol{v}_h) + \sigma(\boldsymbol{u}_h, \boldsymbol{v}_h), \tag{4.2.1d}$$

$$b_h(\boldsymbol{v}_h, q_h) := (\nabla \cdot \boldsymbol{v}_h, q_h). \tag{4.2.1e}$$

This formulation is obtained as described in Remark 3.1.16. Note that $\boldsymbol{V}_h \times Q_h$ and the data (i.e., $\boldsymbol{f} \in \boldsymbol{L}^2(\boldsymbol{\Omega})$, $g \in L^2(\boldsymbol{\Omega})$) are supposed to yield well-defined integral expressions in (4.2.1).

If $-q_h \in Q_h$, then the sign in the method above can be switched by testing with $-q_h$ instead of $q_h$, such that in theory, the two versions (+ and −) of (4.2.1) become equivalent. In contrast to that, the sign might play an important role in the analysis of stabilization methods (see Subsection 4.3.2). Moreover, in practice, the plus and minus variants differ in the algebraic properties of the corresponding finite element matrix.

Let us consider a basis $\{\boldsymbol{\psi}_i\}_{i=1}^{N_u}$ of $\boldsymbol{V}_h$ and a basis $\{\phi_j\}_{j=1}^{N_p}$ of $Q_h$, i.e., $N_u$ and $N_p$ are the numbers of velocity and pressure dofs, respectively. The solutions exhibit basis representations

$$\boldsymbol{u}_h = \sum_{i=1}^{N_u} U_i \boldsymbol{\psi}_i \qquad \text{and} \qquad p_h = \sum_{j=1}^{N_p} P_j \phi_j,$$

with $U_i, P_j \in \mathbb{R}$ for all $i, j$. Inserting them into (4.2.1) and testing with each basis function yields a linear system of the form

$$\mathcal{A} \begin{pmatrix} \boldsymbol{U} \\ \boldsymbol{P} \end{pmatrix} = \begin{pmatrix} \boldsymbol{F}_1 \\ \pm \boldsymbol{F}_2 \end{pmatrix}, \tag{4.2.2}$$

where

- $\boldsymbol{U} \in \mathbb{R}^{N_u}$ with $\boldsymbol{U} := (U_1, \ldots, U_{N_u})^T$,
- $\boldsymbol{P} \in \mathbb{R}^{N_p}$ with $\boldsymbol{P} := \left(P_1, \ldots, P_{N_p}\right)^T$,
- $\boldsymbol{F}_1 \in \mathbb{R}^{N_u}$ with $(\boldsymbol{F}_1)_i = (\boldsymbol{f}, \boldsymbol{\psi}_i)$, for $i = 1, \ldots, N_u$,
- $\boldsymbol{F}_2 \in \mathbb{R}^{N_p}$ with $(\boldsymbol{F}_2)_j = (g, \phi_j)$, for $j = 1, \ldots, N_p$.

The coefficient matrix in (4.2.2) has the special structure

$$\mathcal{A} := \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B}^T \\ \mp \boldsymbol{B} & \boldsymbol{0} \end{pmatrix} \in \mathbb{R}^{(N_u+N_p)\times(N_u+N_p)}, \tag{4.2.3}$$

with

- $\boldsymbol{A} \in \mathbb{R}^{N_u \times N_u}$ with $(\boldsymbol{A})_{ji} := a_h\left(\boldsymbol{\psi}_i, \boldsymbol{\psi}_j\right)$, for $i, j = 1, \ldots, N_u$,
- $\boldsymbol{B}^T \in \mathbb{R}^{N_u \times N_p}$ with $\left(\boldsymbol{B}^T\right)_{ij} := -b_h\left(\boldsymbol{\psi}_i, \phi_j\right)$, for $i = 1, \ldots, N_u$, $j = 1, \ldots, N_p$,
- $\boldsymbol{B} \in \mathbb{R}^{N_p \times N_u}$.

We observe that, if the matrix $\boldsymbol{A}$ is symmetric and the lower left block matrix is $+\boldsymbol{B}$, then $\mathcal{A}$ is a symmetric matrix.

The crucial question is whether (4.2.2) admits a unique solution, i.e., whether the coefficient matrix is invertible. This question cannot be answered in general without more detailed knowledge of the matrix $\mathcal{A}$, but we can derive a prerequisite.

**Lemma 4.2.1** (A Prerequisite for the Non-Singularity of $\mathcal{A}$)**.**
*If the matrix $\mathcal{A}$ in* (4.2.3) *is invertible, then it necessarily holds*

$$N_p \leqslant N_u.$$

*Proof:* The lemma is proven by contradiction. Let us assume that $N_p > N_u$. A square matrix is invertible if and only if it has full rank. For $\mathcal{A}$ this means $\operatorname{rank}(\mathcal{A}) = \mathrm{N_u} + \mathrm{N_p}$. Consider the two block rows in $\mathcal{A}$ separately, i.e., $\mathcal{A}_1 := \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B}^T \end{pmatrix}$ and $\mathcal{A}_2 := (\mp\boldsymbol{B} \quad \boldsymbol{0})$, and use the inequality $\operatorname{rank}(\mathcal{A}) \leqslant \operatorname{rank}(\mathcal{A}_1) + \operatorname{rank}(\mathcal{A}_2)$, [MS74]. The first block row $\mathcal{A}_1 \in \mathbb{R}^{N_u \times (N_u + N_p)}$ satisfies $\operatorname{rank}(\mathcal{A}_1) \leqslant \min\{\mathrm{N_u}, \mathrm{N_u} + \mathrm{N_p}\} = \mathrm{N_u}$. The second block row $\mathcal{A}_2$ spans a space of dimension $\leqslant \min\{N_u, N_p\}$, since the columns $N_u+1, \ldots, N_u+N_p$ are zero in this block row. By assumption, we then obtain $\operatorname{rank}(\mathcal{A}) \leqslant 2\mathrm{N_u} < \mathrm{N_u} + \mathrm{N_p}$ but this implies that $\mathcal{A}$ does not have full rank, which contradicts the invertibility. Hence, $N_p \leqslant N_u$ has to hold for any non-singular matrix of the form (4.2.3). □

Lemma 4.2.1 demonstrates that the number of pressure dofs should not exceed the number of velocity dofs. This has consequences for the usage of different meshes for the velocity and the pressure and the respective polynomial degrees in the finite element spaces.

**Remark 4.2.2** (Block Row Considerations)**.**
*The system* (4.2.2) *can be considered block row-wise, i.e.,*

$$(4.2.2) \quad \Longleftrightarrow \quad \left\{ \begin{array}{r} \boldsymbol{A}\boldsymbol{U} + \boldsymbol{B}^T\boldsymbol{P} = \boldsymbol{F}_1 \\ \mp\boldsymbol{B}\boldsymbol{U} = \mp\boldsymbol{F}_2 \end{array} \right\}.$$

*Let us assume sufficient regularity of all block matrices, then the first block row can be transformed into*

$$\boldsymbol{U} = \boldsymbol{A}^{-1}\left(\boldsymbol{F}_1 - \boldsymbol{B}^T\boldsymbol{P}\right).$$

*Multiplication by $\boldsymbol{B}$ and inserting the second block row yields*

$$\boldsymbol{B}\boldsymbol{A}^{-1}\boldsymbol{B}^T\boldsymbol{P} = \boldsymbol{B}\boldsymbol{A}^{-1}\boldsymbol{F}_1 - \boldsymbol{F}_2.$$

*This is a problem for the pressure, which has a unique solution if and only if $\boldsymbol{B}\boldsymbol{A}^{-1}\boldsymbol{B}^T$ is non-singular. Therefore, $\boldsymbol{A}$ has to be non-singular,* $\ker\left(\boldsymbol{B}^{\mathrm{T}}\right) = \{\boldsymbol{0}\}$, *and* $\ker(\boldsymbol{B}) \cap \operatorname{Im}\left(\boldsymbol{A}^{-1}\boldsymbol{B}\right) = \{\boldsymbol{0}\}$.

*Although this will not be relevant in the numerical simulations presented in this thesis, note that the iterative solution of* (4.2.2) *requires non-standard approaches, due to the zero block in the diagonal of $\mathcal{A}$, where typical preconditioners are not applicable. The matrix $\boldsymbol{B}\boldsymbol{A}^{-1}\boldsymbol{B}^T$ is called Schur complement matrix and its approximation is a strategy for the construction of preconditioners for saddle point problems, [Ahm+18].*

### 4.2.1 Stable Finite Element Pairs – The Stokes and Darcy Limits

The discussion of finite element pairs $\boldsymbol{V}_h/Q_h$ yielding a well-posed discrete Brinkman problem (4.2.1) will begin with the limit cases, $\sigma = 0$ or $\mu_{\text{eff}} = 0$, separately.[6] Therefore, some of the most common (conforming and nonconforming) inf-sup stable pairs for the Stokes problem and the Darcy problem are summarized first.

Finite element methods for the Stokes problem have been extensively considered in the literature, see e.g., [BS08] or [BF91]. For a survey on inf-sup stable finite element pairs for the Darcy problem, it is referred to the references cited in [MH02].

#### The Stokes Problem

The construction of stable finite element pairs relies on the well-posedness criteria given in Theorem 4.1.1 and Theorem 4.1.2. Let us note that for velocity-conforming approaches, the stability conditions that have to be proven simplify. Based on the analytic result concerning the well-posedness of the weak formulation in Theorem 3.3.1 (in particular the coercivity with respect to the unrestricted velocity space) we can draw the following conclusion. If $\boldsymbol{V}_h \subset \boldsymbol{V} = \boldsymbol{H}_0^1(\boldsymbol{\Omega})$, then the coercivity condition is still satisfied with respect to $\boldsymbol{V}_h$ (independent of $Q_h$). Thus, it remains to fulfill the inf-sup condition (4.1.10b) by appropriately choosing $Q_h$ with respect to $\boldsymbol{V}_h$.

The following finite element pairs have been shown to be stable for the Stokes problem:

- The (conforming) *Taylor–Hood elements* $\mathbb{P}_{k+1}/\mathbb{P}_k$, $k \geq 1$, contain continuous velocities and pressures, where on each element $\boldsymbol{T} \in \mathcal{T}_h$, the velocity space is of exactly one degree higher (degree $k+1$) than the pressure space (degree $k$). These pairs are inf-sup stable for the Stokes problem, see [EG04, Ch. 4].
- The (conforming) *Mini element*, also referred to as $\mathbb{P}_1^{bubble}/\mathbb{P}_1$, is obtained from the continuous linear-linear pair $\mathbb{P}_1/\mathbb{P}_1$ by enlarging the velocity space with so-called bubble functions. This effectively adds one dof per element $\boldsymbol{T} \in \mathcal{T}_h$ (associated with the barycenter of each simplex) to the velocity space and results in an inf-sup stable pair, see [ABF84], [CR73], or [EG04, Ch. 4].
- The *non-conforming Crouzeix–Raviart element* $\mathbb{P}_1^{nc}/\mathbb{P}_0$ is obtained by weakening the inter-element continuity of the discrete velocity functions. The velocity dofs are its mean values over facets or equivalently, since the functions are linear on each facet, the dofs are the values in the barycenters of the facets. This nonconforming pair ($\mathbb{P}_1^{nc} \notin \boldsymbol{H}^1(\boldsymbol{\Omega})$) introduces a non-harmful (concerning the convergence order) consistency error (see [Joh16, Sect. 4.2.3]) and is in fact inf-sup stable, see [CR73] and [Joh16, Thm. 3.151].

[6] Finite element pairs $\boldsymbol{V}_h/Q_h$ are often abbreviated using the element-wise polynomial order, e.g., $\mathbb{P}_{k+1}/\mathbb{P}_k$. If not otherwise specified, these are further assumed to be continuous in the domain.

Note that the dofs are edge-based, such that the number of dofs for fine meshes (with a large number of triangles adjacent to a vertex) might be sincerely larger than that of the continuous version $\mathbb{P}_1$.

- The (conforming) *Scott–Vogelius elements* $\mathbb{P}_k/\mathbb{P}_{k-1}^{disc}, k \geqslant 2$ in 2D ($k \geqslant 3$ in 3D), are so-called higher-order pairs with discontinuous pressure spaces. An advantageous property of this family is the (weak) conservation of mass. Unfortunately, well-posedness is guaranteed only for certain meshes (at least if $k < 4$) and these pairs yield linear systems which are considerably more expensive than the one obtained with, e.g., the Taylor–Hood pair $\mathbb{P}_2/\mathbb{P}_1$, [Joh16, Rem. 3.134-3.136].

Note that the given examples of stable finite element pairs consider the same mesh for the velocity and the pressure. We further observe that, in accordance with Lemma 4.2.1, the polynomial degree of the velocity space is for none of the examples smaller than the degree of the pressure space.

### The Mixed Darcy Problem

We have seen in Subsection 3.3.2 that different functional settings yield a well-posed weak Darcy problem. Each of them can be used to derive a finite element method. We are interested in finite element approaches for the dual (mixed) problem.

In contrast to the Stokes problem (and the mixed primal Darcy problem), conformity does not suffice here to inherit the coercivity condition (4.1.10a). The reason is that coercivity is only guaranteed in $\boldsymbol{V}_0 \subsetneq \boldsymbol{H}_0^{\mathrm{div}}(\boldsymbol{\Omega})$, where we have sufficient control on the $\boldsymbol{L}^2$-norm of the divergence of the velocity, see Theorem 3.3.6. Indeed, $\boldsymbol{V}_h \times Q_h \subset \boldsymbol{V} \times Q \quad \not\Rightarrow \quad \boldsymbol{V}_{0,h} \subset \boldsymbol{V}_0$. However, if we could assure the latter inclusion, i.e., if *discretely divergence-free* velocities (elements in $\boldsymbol{V}_{0,h}$) were *weakly divergence-free* (elements in $\boldsymbol{V}_0$), then coercivity would be satisfied and it would remain to satisfy the inf-sup condition (4.1.10b) in order to guarantee well-posedness.

Stable finite element pairs for the dual Darcy problem include the following:

- The lowest order ($\boldsymbol{H}^{\mathrm{div}}$-conforming) *Raviart–Thomas element* $\boldsymbol{V}_h^{\mathrm{RT}}/\mathbb{P}_0$, where the velocity functions are element-wise polynomials of the form $\boldsymbol{v}_h|_{\boldsymbol{T}} = \boldsymbol{a}_{\boldsymbol{T}} + c_{\boldsymbol{T}}\boldsymbol{x}$, $\boldsymbol{a}_{\boldsymbol{T}} \in \mathbb{R}^n, c_{\boldsymbol{T}} \in \mathbb{R}$, such that $\boldsymbol{v}_h \cdot \boldsymbol{n}$ is continuous across interior element boundaries of the triangulation (which yields $\boldsymbol{V}_h^{\mathrm{RT}} \subset \boldsymbol{H}^{\mathrm{div}}(\boldsymbol{\Omega})$, see [Joh16, Lem. 3.66]).
  By definition we have $\nabla \cdot \boldsymbol{V}_h^{RT} \subset Q_h = \mathbb{P}_0$ (implying $\boldsymbol{V}_{0,h} \subset \boldsymbol{V}_0$) but the complexity of implementation is often mentioned as a drawback, see, e.g., [MH02] and [BC10]. For more details see [KS12] or [Joh16, Example B.45].
- The lowest order ($\boldsymbol{H}^{\mathrm{div}}$-conforming) *Brezzi–Douglas–Marini* (BDM) element $\boldsymbol{V}_h^{\mathrm{BDM}}/\mathbb{P}_0$ differs from the Raviart–Thomas element in that the velocity on each $\boldsymbol{T} \in \mathcal{T}_h$ can be any polynomial in $\mathbb{P}_1(\boldsymbol{T})$ with $\boldsymbol{v}_h \cdot \boldsymbol{n}$ being continuous across inter-element facets. Hence, the BDM element can be considered as a generalization of the Raviart–Thomas element, since $\boldsymbol{V}_h^{\mathrm{RT}} \subset \boldsymbol{V}_h^{\mathrm{BDM}}$. Further information is given in, e.g., [BF91, III.3].

### 4.2.2 Stable Finite Element Pairs – The Brinkman Problem

The development of finite element methods for the Brinkman problem, which are accurate and stable uniformly with respect to the coefficients $\mu_{\text{eff}}$ and $\sigma$, is an involved task and a topic of active research. As we have seen, there are finite element pairs available that yield well-posedness for each limit problem (in standard Galerkin form) separately but a striking observation concerning our survey is that none of them appears in both lists.

In [MTW02], a numerical study in 2D has demonstrated that typical stable Stokes pairs do not behave well for Darcy problems and vice versa. Therefore, the authors considered a scaled version of the Brinkman problem (as in Rem. 3.2.2), defined the velocity energy norm as $\|\boldsymbol{v}_h\|_{t^2}^2 := t^2\|\nabla\boldsymbol{v}_h\|_0^2 + \|\nabla\cdot\boldsymbol{v}_h\|_0^2 + \|\boldsymbol{v}_h\|_0^2$, and computed convergence rates for $t \in [0,1]$, with the following results.

- The $\mathbb{P}_2/\mathbb{P}_0$ element: No convergence for $t^2 = 0$ in the $L^2$-norm of the velocity and the pressure, and the energy-norm of the velocity.
- The non-conforming Crouzeix–Raviart element: No convergence (with respect to the $L^2$-norm of the velocity and the pressure, and the energy-norm of the velocity) if $t^2 = 0$.
- The Mini element: Convergence in the velocity energy-norm deteriorates for $t^2 \to 0$, but if the error is measured in the norm with respect to $\boldsymbol{L}^2 \times H^1 \cap L_0^2$ as in (3.3.13), the Mini element is uniformly stable (see also [JS10] and [HJS11], where in addition the Taylor–Hood element is proven to obey a similar qualitative behavior), however with varying convergence rates for the cases $t = 1$ and $t = 0$.
- The Raviart–Thomas element: No convergence of the $L^2$-error for the velocity and the pressure for $t = 1$.

This opens the question, what kind of criteria assure that stable Stokes elements remain stable in the Darcy limit, and vice versa.

#### Stokes-Based Finite Element Methods for the Brinkman Problem

Starting from a conforming stable Stokes element in the sense of Theorem 4.1.2, the validity of the inf-sup condition (4.1.10b) in the Darcy limit of (4.2.1) can be proven equivalently as in the Stokes case, using Corollary 3.1.4.

For $\mu_{\text{eff}} = 0$, the (sufficient) condition (4.1.10a) is satisfied if there exists a constant $m > 0$, such that

$$a_h(\boldsymbol{v}_h, \boldsymbol{v}_h) = \sigma\|\boldsymbol{v}_h\|_0^2 \geqslant m\|\nabla\cdot\boldsymbol{v}_h\|_0^2, \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_{0,h}, \tag{4.2.4}$$

with $a_h(\cdot,\cdot)$ defined in (4.2.1). Let us mention that with the necessary well-posedness criteria proven in [BF91, Thm. 1.1] (certain inf-sup condition for $a_h(\cdot,\cdot)$), the inequality (4.2.4) is also necessary for the well-posedness.

The condition (4.2.4) is trivially satisfied if discretely divergence-free velocities are weakly divergence-free, i.e.,

$$\boldsymbol{V}_{0,h} \subset \boldsymbol{V}_0 \iff (\boldsymbol{v}_h \in \boldsymbol{V}_h\colon\ (\nabla\cdot\boldsymbol{v}_h, q_h) = 0,\ \forall q_h \in Q_h \implies \nabla\cdot\boldsymbol{v}_h = 0 \text{ a.e.})\,.$$

Unfortunately, this 'recipe' is only of limited value in practice, since standard stable Stokes elements in general do not satisfy the inclusion property $\boldsymbol{V}_{0,h} \subset \boldsymbol{V}_0$, see, e.g., [Joh16, Rem. 4.147].

A well known sufficient criterion for $\boldsymbol{V}_{0,h} \subset \boldsymbol{V}_0$ is $\nabla \cdot \boldsymbol{V}_h \subset Q_h$. This is fulfilled for example by the Scott–Vogelius element, which is stable on special meshes only (in the 'lower'-order cases). For more details see, e.g., [MTW02], [XXX08], and [BFM93].

In [XXX08] the usage of a grad-div stabilization (curing the violation of the condition (4.1.10a) in the Darcy limit) is suggested, such that all classical stable Stokes elements become stable for Darcy problems as well.

#### Darcy-Based ($H^{\mathrm{div}}$-Conforming) Finite Element Methods for the Brinkman Problem

Several $\boldsymbol{H}^{\mathrm{div}}$-conforming finite element approaches have already been investigated for the Stokes problem in several publications, e.g., [CKS07], [KS08], [WWY09], [WY07], and [KR10]. One of the main difficulties is that an $\boldsymbol{H}^{\mathrm{div}}$-conforming discrete velocity space is non-conforming in the Stokes case, since the continuity of the tangential velocity component is not enforced, which causes a harmful (w.r.t. the convergence order) consistency error. Possibilities to repair this deficiency are unfortunately rather complicated and/or expensive.

One approach that allows to appropriately bound the consistency error is to penalize tangential velocity jumps in the Stokes case (e.g., via a symmetric interior penalty approach of Nitsche type, see, [KS11], [KS12]).

Another strategy is to enrich the velocity space appropriately (by adding degrees of freedom), as done in [MTW02], [XXX08], and [GN12]. The modification is realized such that the desired accuracy for the consistency error estimate is obtained. The price for that are additional dofs, in order to in particular control the mean values of the tangential velocity at interior facets.

## 4.3 Equal–Order (Continuous) Approximations and Stabilizations

Subsection 4.2.2 has clarified that constructing low-order finite element methods, which are uniformly (with respect to $\mu_{\mathrm{eff}}$ and $\sigma$) stable and convergent for the Brinkman problem, is still an open challenge. This section discusses the construction of a suitable (nonstandard) finite element method satisfying these criteria.

In practical applications, the computational cost can be a decisive aspect when choosing the proper finite element spaces. Therefore, despite high-order approaches potentially yielding higher accuracies of the approximate solution on the same mesh, in practice, low-order methods are very attractive, in particular low-complexity variants. Unfortunately, these typically suffer from the incompatibility of the finite element spaces, chosen for the velocity and the pressure

(inf-sup instability). Hence, the desire for low-order approaches boils down to the search for modifications which are able to recover stability.

In this section we introduce the family of continuous equal-order finite element pairs for mixed problems. Focusing on the lowest order pair, we discuss stability issues and possibilities of stabilization. This will include the non-symmetric GLS (Douglas–Wang) method. For general information on stabilization methods the books [RST08] and [QV08] are recommended. A recently published review concerning stabilizations for the Stokes problem is given in [JKW19].

### 4.3.1 The Continuous Equal-Order Finite Element Pair

For the (conforming) equal-order finite element pair, the finite element spaces for the velocity and the pressure are chosen to be of the same order $k \geqslant 1$, i.e.,

$$\begin{aligned} \boldsymbol{V}_h^k &:= \left\{ \boldsymbol{v}_h \in \boldsymbol{C}\left(\overline{\boldsymbol{\Omega}}\right) \cap \boldsymbol{V} \,:\, \boldsymbol{v}_h|_{\boldsymbol{T}} \in \mathbb{P}_k\left(\boldsymbol{T}\right), \quad \forall \boldsymbol{T} \in \mathcal{T}_h \right\}, \\ Q_h^k &:= \left\{ q_h \in C\left(\overline{\boldsymbol{\Omega}}\right) \cap Q \,:\, q_h|_{\boldsymbol{T}} \in \mathbb{P}_k\left(\boldsymbol{T}\right), \quad \forall \boldsymbol{T} \in \mathcal{T}_h \right\}. \end{aligned} \tag{4.3.1}$$

A common abbreviation that is also adopted here is to refer to both spaces simply by $\mathbb{P}_k$. For a visualization of the degrees of freedom for $k = 1, 2$ see Figure 4.2.


Figure 4.2: A simplicial element $\boldsymbol{T}$ in 2D with dofs (black circles) for the Lagrange finite elements $\mathbb{P}_1$ (left) and $\mathbb{P}_2$ (right).

A favorable practical aspect of the pair $\mathbb{P}_k/\mathbb{P}_k$ is that the same implementation can be used for both, velocity and pressure.

Among them, the lowest-order member $\mathbb{P}_1/\mathbb{P}_1$ is advantageous in terms of computational effort and memory, and since $\mathbb{P}_1$ is probably already part of any finite element code based on simplicial meshes, [Joh16, Sect. 3.4], [EG04, Ch. 4]. So let us now focus on the case $k = 1$.

#### Instability of the Linear-Linear Pair – $\mathbb{P}_1/\mathbb{P}_1$

The equal-order pair $\boldsymbol{V}_h^1/Q_h^1$ defined in (4.3.1) (and $\boldsymbol{V}_h^k/Q_h^k$ in general, [BF91, §VI.3] and [BBF13, Ch. 8.3]) violates the inf-sup condition (4.1.10b). As a consequence, for some meshes, the numerical results are strongly mesh-dependent and the pressure solution shows an unphysical behavior. Let us exemplify the problem for a uniform, triangular mesh on the unit square $\boldsymbol{\Omega} := (0,1)^2$. Therefore we construct a pressure function $p_h^\star \in Q_h^1$ which is an element of the *space of spurious pressure modes*:

$$p_h^\star \in \left\{ q_h \in Q_h^1 \,:\, (\nabla \cdot \boldsymbol{v}_h, q_h) = 0, \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h^1 \right\}.$$

Whenever this space is nontrivial, the inf-sup constant in (4.1.10b) becomes 0, implying inf-sup instability.

Exploiting the fact that the discrete velocity $\boldsymbol{v}_h \in \boldsymbol{V}_h^1$ is element-wise linear, implying $(\nabla \cdot \boldsymbol{v}_h)\,|_{\boldsymbol{T}} = \textit{constant}$ for all $\boldsymbol{T} \in \mathcal{T}_h$, we deduce that it suffices to construct a function $p_h^\star \in Q_h^1$ with

$$\sum_{i=0}^{2} p_h^\star\left(a_{\boldsymbol{T},i}\right) = 0, \quad \forall \boldsymbol{T} \in \mathcal{T}_h,$$

where $\boldsymbol{T} = \operatorname{conv}\{\mathrm{a}_{\boldsymbol{T},0}, \mathrm{a}_{\boldsymbol{T},1}, \mathrm{a}_{\boldsymbol{T},2}\}$ and $a_{\boldsymbol{T},i}$, $i = 0, 1, 2$, are the vertices of the triangle. Then we obtain

$$\begin{aligned} b_h\left(\boldsymbol{v}_h, p_h^\star\right) = \left(\nabla \cdot \boldsymbol{v}_h, p_h^\star\right) &= \sum_{\boldsymbol{T} \in \mathcal{T}_h} \left(\nabla \cdot \boldsymbol{v}_h, p_h^\star\right)_{\boldsymbol{T}} = \sum_{\boldsymbol{T} \in \mathcal{T}_h} \left(\nabla \cdot \boldsymbol{v}_h\right)|_{\boldsymbol{T}} \left(1, p_h^\star\right)_{\boldsymbol{T}} \\ &= \sum_{\boldsymbol{T} \in \mathcal{T}_h} \left(\nabla \cdot \boldsymbol{v}_h\right)|_{\boldsymbol{T}} \frac{|\boldsymbol{T}|}{3} \left(\sum_{i=0}^{2} p_h^\star\left(a_{\boldsymbol{T},i}\right)\right) = 0, \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h^1, \end{aligned}$$

which contradicts the inf-sup condition. Note that a quadrature formula, which is exact for linear functions, was used in the fourth equality.

The most famous counterexample of this type is a checkerboard-like instability, where the pressure $p_h^\star \in Q_h^1$ is defined to take the values $-1$, 0, and 1 on each triangle $\boldsymbol{T} \in \mathcal{T}_h$, see Figure 4.3.

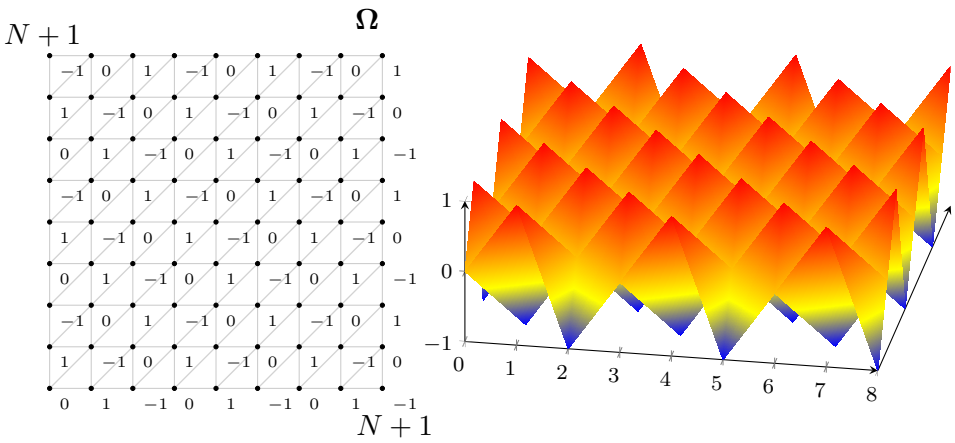



Figure 4.3: Triangulation of a square domain $\boldsymbol{\Omega} \subset \mathbb{R}^2$ with $(N+1)^2 = 81$ nodes in total and $(N-1)^2 = 49$ inner nodes (left). The nodal values $-1$, 0, and 1 correspond to a spurious pressure mode for $\mathbb{P}_1/\mathbb{P}_1$ (left and right).

**Remark 4.3.1** (Low-Order Spaces with Discontinuous Pressures)**.**
*Let us legitimate the choice $\mathbb{P}_1/\mathbb{P}_1$ with respect to alternative low-order pairs. Continuity across elements is not obligatory for $Q_h$, even not if $Q_h \subset Q := L_0^2(\boldsymbol{\Omega})$ (conforming). Therefore, using discontinuous pressures might yield conceivable low-order approaches, which also assure that mass is conserved element-wise*

*in the sense of $L^2$ (and overall if $\nabla \cdot \boldsymbol{V}_h \subset Q_h$, see Subsection 4.2.2). Such low-order pairs are in general not inf-sup stable and usually more expensive. To see that, consider a triangulation of $\boldsymbol{\Omega} \subset \mathbb{R}^2$ into $2N^2$ triangles, $N \geqslant 1$, as visualized in Figure 4.3 (left) for $N = 8$. This yields $(N+1)^2$ nodes in total and $(N-1)^2$ interior nodes. Assuming that the velocity at the boundary nodes is imposed and one pressure dof is fixed to assure $q_h \in L_0^2(\boldsymbol{\Omega})$, we count the velocity and pressure dofs:*

$$\begin{aligned} \dim\left(\boldsymbol{V}_h^1\right) &= 2\left(\#\{\textit{interior nodes}\}\right) = 2\left(N-1\right)^2, \\ \dim\left(Q_h^1\right) &= \#\{\textit{nodes}\} - 1 = \left(N+1\right)^2 - 1. \end{aligned}$$

- *For the linear-constant finite element pair $\boldsymbol{V}_h^1/Q_h^{0,disc}$ (also $\mathbb{P}_1/\mathbb{P}_0$), with $Q_h^{0,disc} := \{q_h \in Q \,:\, q_h|_{\boldsymbol{T}} \in \mathbb{P}_0(\boldsymbol{T}), \quad \forall \boldsymbol{T} \in \mathcal{T}_h\}$, we have*

$$\dim\left(Q_h^{0,disc}\right) = \#\{\textit{elements}\} - 1 = 2N^2 - 1.$$

  *dofs. Hence,*

$$\dim\left(\mathrm{Q}_\mathrm{h}^{0,\mathrm{disc}}\right) - \dim\left(\mathrm{Q}_\mathrm{h}^1\right) = \mathrm{N}^2 - 2\mathrm{N} - 1 > 0, \qquad \textit{for } \mathrm{N} \geqslant 3,$$

  *i.e., for large $N$, the number of extra dofs is essentially of order $N^2$, although $Q_h^1$ has in general better approximation properties than $Q_h^{0,disc}$. Note that also $\boldsymbol{V}_h^1/Q_h^{0,disc}$ is unstable, since $\dim\left(\mathrm{Q}_\mathrm{h}^{0,\mathrm{disc}}\right) > \dim\left(\boldsymbol{V}_\mathrm{h}^1\right)$, see Lemma 4.2.1. A phenomenon called locking might appear, which means that the only discretely mass conservative velocity field is $\boldsymbol{u}_h = \boldsymbol{0}$, see also [EG04, Ch. 4.2.3].*
- *For the pair $\boldsymbol{V}_h^1/Q_h^{1,disc}$ (also $\mathbb{P}_1/\mathbb{P}_1^{disc}$), the situation is similar. With $Q_h^{1,disc} := \{q_h \in Q \,:\, q_h|_{\boldsymbol{T}} \in \mathbb{P}_1(\boldsymbol{T}), \quad \forall \boldsymbol{T} \in \mathcal{T}_h\}$, we have with*

$$\dim\left(Q_h^{1,disc}\right) = 3\left(\#\{\textit{elements}\}\right) - 1 = 3\left(2N^2\right) - 1,$$

$$\dim\left(Q_h^{1,disc}\right) - \dim\left(Q_h^1\right) = 5N^2 - 2N - 1 > 0, \qquad \textit{for } N \geqslant 1.$$

  *Also here $\dim\left(\mathrm{Q}_\mathrm{h}^{1,\mathrm{disc}}\right) > \dim\left(\boldsymbol{V}_\mathrm{h}^1\right)$ yields inf-sup instability according to Lemma 4.2.1.*

*These observations suggest the clear preference of the continuous version* (4.3.1) *concerning memory and computational effort.*

### 4.3.2 Pressure Stabilizations and the Douglas–Wang Method

Inf-sup unstable methods are reported (e.g., in [HFB86]) to perform often well for sufficiently smooth solutions. Nevertheless, it is important to assure the existence of a unique solution, which requires stability.

The instability of $\mathbb{P}_1/\mathbb{P}_1$, demonstrated in Subsection 4.3.1, suggests that the discrete pressure space $Q_h$ is too large with respect to the discrete velocity

space $\boldsymbol{V}_h$. The intuitive conclusion is that either an enlargement of $\boldsymbol{V}_h^1$ or a downsizing of $Q_h^1$ could improve stability.

An explicit extension of the velocity space can be realized by adding degrees of freedom, for example, by increasing the polynomial order and/or by using discontinuous functions. In fact, popular inf-sup stable pairs (see Subsection 4.2.1) can be interpreted as the result of such an ansatz, e.g., the Mini element based on the unstable pair $\mathbb{P}_1/\mathbb{P}_1$ or the (non-conforming) Crouzeix–Raviart element and the Raviart–Thomas element as modifications of the unstable pair $\mathbb{P}_1/\mathbb{P}_0$.

An alternative technique to overcome the (inf-sup) instability is to add appropriate terms to the discrete formulation obtained from the Galerkin method, such that the fulfillment of the discrete inf-sup condition (w.r.t. $b_h(\cdot,\cdot)$ in the standard Galerkin method) is no longer necessary. In the following part, we will discuss stabilizations for $\mathbb{P}_1/\mathbb{P}_1$ based on introducing pressure-pressure couplings (filling the zero block matrix in (4.2.3)) that suppress spurious oscillations.

### The Brezzi–Pitkäranta Method and Residual-Based Variants

In the original version of pressure stabilization, proposed in [BP84] for the Stokes problem and therefore referred to as **Brezzi–Pitkäranta method**, stability is achieved by penalizing pressure gradients via the additional term

$$\alpha \sum_{\boldsymbol{T}\in\mathcal{T}_h} \alpha_{\boldsymbol{T}} \left(\nabla p_h, \nabla q_h\right)_{\boldsymbol{T}} \tag{4.3.2}$$

in the Galerkin formulation. The strength of the penalization can be regulated by a stabilization parameter $\alpha \in \mathbb{R}^+$, while $\alpha_{\boldsymbol{T}}$, which might vary across mesh cells, determines different variants of the method. Following the stability and convergence analysis, a reasonable definition of $\alpha_{\boldsymbol{T}}$ usually contains parameters of the model and a power of the element diameter $h_{\boldsymbol{T}}$. In [BP84], $\alpha = 1$ and $\alpha_{\boldsymbol{T}} = h_{\boldsymbol{T}}^2$ were chosen for $\mathbb{P}_1/\mathbb{P}_1$. Adding terms of the form (4.3.2) enables the additional control of $\sum_{\boldsymbol{T}\in\mathcal{T}_h} \alpha_{\boldsymbol{T}} \|\nabla p_h\|_{0,\boldsymbol{T}}^2$. Consequently, it might be reasonable to consider mesh-dependent norms in the analysis.

Note that this approach is in general inconsistent and $\alpha_{\boldsymbol{T}}$ should be chosen such that the corresponding consistency error does not negatively affect (optimal) convergence. The term (4.3.2) formally belongs to the equation of mass conservation (due to the testing with $\nabla q_h$). Presupposing for simplicity sufficient regularity of the pressure and $\alpha_{\boldsymbol{T}} = 1$, the pressure stabilization can be obtained from the modified continuity equation

$$\nabla \cdot \boldsymbol{u} - \alpha \Delta p = g$$

with a homogeneous natural boundary condition on $p$ such that the integration by parts term $\langle \alpha \nabla p \cdot \boldsymbol{n}, q \rangle$ vanishes. Hence, an unbalanced pressure-pressure coupling of the form (4.3.2) perturbs the conservation of mass (adds *artificial compressibility*) and introduces artificial pressure boundary conditions, see, e.g., [JKW19, Sect. 5] and [QV08, Sect. 9.4].

In order to overcome the inconsistency of the Brezzi–Pitkäranta method, several consistent variants were proposed, originally focusing on the Stokes problem. The idea is to replace the pressure gradient $\nabla p_h$ in (4.3.2) by the strong residual of the momentum balance equation, which contains the crucial term $\nabla p_h$. This yields a so-called **residual-based pressure stabilization** and guarantees consistency due to the presence of an appropriate balancing term.

Focusing on the Brinkman problem, the consistent variants of the Brezzi–Pitkäranta method can be formulated by adding the term:

$$\mathcal{S}_h^\star\left[(\boldsymbol{u}_h, p_h), \zeta_h\right] := \alpha \sum_{\boldsymbol{T}\in\mathcal{T}_h} \alpha_{\boldsymbol{T}} \left(-\mu_{\text{eff}}\Delta \boldsymbol{u}_h + \nabla p_h + \sigma \boldsymbol{u}_h - \boldsymbol{f}, \zeta_h\right)_{\boldsymbol{T}}. \tag{4.3.3}$$

The generic form (4.3.3) can be used to describe different stabilization methods, depending in particular on the form of the used test function $\zeta_h$. Later on we will define (4.3.3) for $\star \in \{\text{GLS, PSPG}\}$. The term (4.3.3) does not negatively affect the consistency of the discrete formulation, since for the exact solution, it is $\mathcal{S}_h^\star\left[(\boldsymbol{u}, p), \zeta_h\right] = 0$, for any $\zeta_h$, by satisfaction of the momentum balance equation. Note that the practical implementation of (4.3.3) requires only little extra effort compared to standard Galerkin code for Lagrange elements.

**Remark 4.3.2** (On Residual-Based Pressure Stabilizations)**.**
*The larger the stabilization parameter $\alpha\alpha_{\boldsymbol{T}}$ gets, the more the residual of the momentum balance on the mesh cell $\boldsymbol{T}$ is weighted.*

*According to [BDG06, Sect. 2.1], the additional stabilization terms (apart from the crucial pressure-pressure coupling) may have a destabilizing effect, such that the parameters have to be chosen carefully in order to control them. In numerical simulations, $\alpha \in [0.01, 1]$ is often a good choice, see e.g., [NS98, Sect.2.3].*

*Assuming that $\alpha$ satisfies a given a priori stability condition, in practice we observe the following: If $\alpha$ is chosen too small, the stabilization might be insufficient to eliminate the problematic pressure modes and then spurious oscillations might still affect the quality of the solution, as will be demonstrated in Section 6.2. Otherwise, if $\alpha$ is too large, this could have a negative impact on the condition number of the finite element matrix (see [Bra+07] and [NS98, Sect.2.3]) and the pressure near the domain boundary might be poorly approximated, [QV08, Sect. 9.4].*

**Remark 4.3.3** (On the Appearance of Second-Order Derivatives)**.**
*The energy norm in the Stokes case ($\mu_{\text{eff}} > 0$) contains only velocity derivatives up to order one. Therefore, the analysis of residual-based stabilizations including a second order velocity term uses estimates of it by lower order derivatives, known as inverse estimate* (4.1.3b)*. This estimation might introduce restrictions on $\alpha$, yielding conditional stability of the method.*

**Remark 4.3.4** (Stabilizations for Discontinuous Pressures)**.**
*A different type of pressure stabilization can be used for $\mathbb{P}_1/\mathbb{P}_0$ elements. In this case, in order to introduce enough coupling between the pressure dofs (avoiding*

*locking and obtaining inf-sup stability), the edge-jump-based term*

$$\eta \sum_{E\in\mathcal{T}_h\backslash\Gamma} h_E \left(\left[\!\left[p_h\right]\!\right], \left[\!\left[q_h\right]\!\right]\right)_E ,$$

*can be used, where* $[\![p_h]\!]$ *is the jump of* $p_h$ *across* $E$*. Such a term is sufficient for the use of* $Q_h = \mathbb{P}_0$*, for higher order pairs (also for discontinuous pressures) an additional pressure stabilization (*$\alpha > 0$*) is required, [HF87]. The parameter* $\eta$ *has to be chosen carefully since asymptotically (*$\eta \to \infty$*), the pressure might tend to be constant, for sure continuous, [NS98].*

### The Non-Symmetric GLS (Douglas–Wang) Method

In 1989, the **non-symmetric GLS (Douglas–Wang) method** was introduced in [DW89]. Applied to the Brinkman problem with continuous pressures it reads as follows:
Find $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h^k \times Q_h^k$ such that

$$\mathcal{A}_h^{(+)}\left[(\boldsymbol{u}_h, p_h); (\boldsymbol{v}_h, q_h)\right] + \mathcal{S}_h^{GLS}\left[(\boldsymbol{u}_h, p_h), \zeta_h\right] = \mathcal{F}_h^{(+)}\left[(\boldsymbol{v}_h, q_h)\right], \tag{4.3.4a}$$

for all $(\boldsymbol{v}_h, q_h) \in \boldsymbol{V}_h^k \times Q_h^k$ with

$$\zeta_h := -\mu_{\text{eff}}\Delta\boldsymbol{v}_h + \nabla q_h + \sigma\boldsymbol{v}_h \tag{4.3.4b}$$

and the expressions $\mathcal{A}_h^{(+)}[(\cdot,\cdot);(\cdot,\cdot)]$ and $\mathcal{F}_h^{(+)}[(\cdot,\cdot)]$ defined as in (4.2.1). Formally, the non-symmetric GLS method (4.3.4a) is obtained by adding (the sign is important for the stability properties) the discrete continuity equation, as well as the stabilization term $\mathcal{S}_h^{GLS}[(\cdot,\cdot),\zeta_h]$, to the discrete momentum balance equation. The resulting formulation is non-symmetric and the choice of the test function (4.3.4b) results in a stabilized bilinear form which contains element-wise *Galerkin least-squares terms*. Therefore, the method is termed here *non-symmetric GLS method*.

Remarkably, this stabilization was introduced as an *absolutely stabilized* method for the Stokes problem ($\mu_{\text{eff}} = 1, \sigma = 0$) meaning that stability was proven for any nonzero choice of stabilization parameter $\alpha > 0$ with $\alpha_{\boldsymbol{T}} := h_{\boldsymbol{T}}^2$, see also [FHS93, Thm. 4.3.3]. The price to pay for this unconditional stability is the non-symmetry of the corresponding linear system matrix.

**Remark 4.3.5** (The Continuity Equation for the Non-Symmetric GLS Method)**.** *The non-symmetric GLS (Douglas–Wang) method is formally associated with a modified continuity equation (and momentum equation), containing a weighted residual of the momentum balance equation. If* $\alpha_{\boldsymbol{T}} = 1$ *and* $\boldsymbol{u}$*,* $p$*, and* $\boldsymbol{f}$ *are sufficiently smooth, then the corresponding continuity equation reads in strong form*

$$\nabla\cdot\boldsymbol{u} - \alpha\nabla\cdot(-\mu_{\text{eff}}\Delta\boldsymbol{u} + \nabla p + \sigma\boldsymbol{u} - \boldsymbol{f}) = g,$$

*where the expression in the brackets is equal to zero, if* $(\boldsymbol{u}, p)$ *solves the momentum equation.*

**Remark 4.3.6** (Other Popular Variants – PSPG and Symmetric GLS)**.**
*In* (4.3.3)*, several different test functions* $\zeta_h$ *depending on* $\nabla q_h$ *can be used. Detailed discussions for the Stokes problem are provided in [FHS93] and [JKW19]. The most popular choices include the following:*

- *The Pressure-Stabilized Petrov–Galerkin (PSPG) method presented in [HFB86] is the most simple and computationally cheapest consistent approach of the form* (4.3.3) *namely one considers* $\zeta_h := \nabla q_h$ *in the stabilized formulation*

$$\mathcal{A}_h^{(\pm)}\left((\boldsymbol{u}_h, p_h); (\boldsymbol{v}_h, q_h)\right) \pm \mathcal{S}_h^{PSPG}\left[(\boldsymbol{u}_h, p_h), \zeta_h\right] = \mathcal{F}_h^{(\pm)}\left(\boldsymbol{v}_h, q_h\right),$$

  *with the definitions in* (4.2.1)*. The resulting problem is strongly non-symmetric.*
- *For the (symmetric) Galerkin least squares method (GLS or GALS) from [HF87], which reads as*

$$\mathcal{A}_h^{(-)}\left((\boldsymbol{u}_h, p_h); (\boldsymbol{v}_h, q_h)\right) - \mathcal{S}_h^{GLS}\left[(\boldsymbol{u}_h, p_h), \zeta_h\right] = \mathcal{F}_h^{(-)}\left(\boldsymbol{v}_h, q_h\right),$$

  *the same* $\zeta_h := -\mu_{\text{eff}}\Delta \boldsymbol{v}_h + \nabla q_h + \sigma \boldsymbol{v}_h$ *as in the Douglas–Wang method is used, however, the formulation differs in the signs which guarantee here the maintenance of symmetry.*

*Both methods were originally proposed for the Stokes problem (* $\mu_{\text{eff}} > 0$*,* $\sigma = 0$*) with* $\alpha_{\boldsymbol{T}} = \frac{h_{\boldsymbol{T}}^2}{2\mu_{\text{eff}}}$*. Stability was proven for stabilization parameters* $\alpha$ *below a specific upper bound depending (inversely) on the constant* $\hat{c}_I$ *in the inverse estimate* (4.1.3b) *(controlling* $\Delta \boldsymbol{u}_h$ *and* $\Delta \boldsymbol{v}_h$*). Notice that in general the bound for the parameter* $\alpha$ *is a priori unknown, which might yield suboptimal choices of* $\alpha$ *in practice. Hence, either additional computational effort (numerical study) or analytical effort for theoretical investigations (see, e.g., [HH92]) is the consequence.*

The non-symmetric GLS method formally differs from the symmetric GLS method (see Rem. 4.3.6) only in terms of signs. However, the analysis suggests that the non-symmetric method is more robust in the sense that the accuracy of the discrete solution is less sensitive to the value of the parameter $\alpha$. This has also been confirmed by numerical studies, see, e.g., [FF92], where the GLS and the Douglas–Wang method were compared. Their tests approved the instability of the symmetric GLS method in case of an inappropriate choice of the stabilization parameter, manifesting as spurious oscillations of the velocity and the pressure, not present for the Douglas–Wang approach.

**Remark 4.3.7** (Special Case – Linear Velocities and the Stokes Limit)**.**
*For the discrete velocity space* $\mathbb{P}_1$*, the residual-based pressure stabilizations simplify, since the second order derivatives vanish on each triangle, i.e.,* $\Delta \boldsymbol{u}_h|_{\boldsymbol{T}} = \boldsymbol{0}$*. In this situation, related upper bounds on* $\alpha$ *(due to inverse estimates) disappear. This is also the case in the Darcy limit (*$\mu_{\text{eff}} = 0$*).*

*Moreover, if* $\boldsymbol{V}_h = \mathbb{P}_1$ *and additionally* $\sigma = 0$ *(Stokes limit), all the residual-based pressure stabilizations discussed in this section coincide (since* $q_h$ *can be replaced by* $-q_h$*). In presence of* $\sigma \boldsymbol{v}_h$*, i.e., for* $\sigma \neq 0$*, this is not the case.*

*Let us observe that changing the sign of $\boldsymbol{v}_h$ or $q_h$ in the formulation (if allowed) does not change the method, but the algebraic equations.*

### Linear Algebra Aspects of the Non-Symmetric GLS (Douglas–Wang) Method

Adding stabilizations of the form (4.3.3) introduces, besides pressure-pressure couplings, also velocity-pressure couplings. The purpose of this paragraph is to discuss the linear system resulting from the non-symmetric GLS (Douglas–Wang) method. In order to see the contributions to the respective matrix blocks, we test the problem (4.3.4) by $(\boldsymbol{v}_h, 0)$ and $(\boldsymbol{0}, q_h)$. Distributing the terms to the matrix blocks with the same strategy and notation as used for (4.2.2) and (4.2.3) yields

$$\underbrace{\begin{pmatrix} \boldsymbol{A}^{\mathrm{nsGLS}} & \boldsymbol{B}_1^{\mathrm{nsGLS}} \\ \boldsymbol{B}_2^{\mathrm{nsGLS}} & \boldsymbol{C}^{\mathrm{nsGLS}} \end{pmatrix}}_{=:\mathcal{A}^{\mathrm{nsGLS}}} \begin{pmatrix} \boldsymbol{U} \\ \boldsymbol{P} \end{pmatrix} = \begin{pmatrix} \boldsymbol{F}_1^{\mathrm{nsGLS}} \\ \boldsymbol{F}_2^{\mathrm{nsGLS}} \end{pmatrix}, \tag{4.3.5}$$

where $\boldsymbol{A}^{\mathrm{nsGLS}} \in \mathbb{R}^{N_u \times N_u}$, $\boldsymbol{B}_1^{\mathrm{nsGLS}} \in \mathbb{R}^{N_u \times N_p}$, $\boldsymbol{B}_2^{\mathrm{nsGLS}} \in \mathbb{R}^{N_p \times N_u}$, $\boldsymbol{C}^{\mathrm{nsGLS}} \in \mathbb{R}^{N_p \times N_p}$, $\boldsymbol{F}_1^{\mathrm{nsGLS}} \in \mathbb{R}^{N_u}$, $\boldsymbol{F}_2^{\mathrm{nsGLS}} \in \mathbb{R}^{N_p}$, and

$$\begin{aligned}
\left(\boldsymbol{A}^{\mathrm{nsGLS}}\right)_{ij} &:= \underbrace{a_h\left(\boldsymbol{\psi}_j, \boldsymbol{\psi}_i\right)}_{=(\boldsymbol{A})_{ij}} + \alpha \sum_{\boldsymbol{T} \in \mathcal{T}_h} \alpha_{\boldsymbol{T}} \left(-\mu_{\mathrm{eff}} \Delta \boldsymbol{\psi}_j + \sigma \boldsymbol{\psi}_j, -\mu_{\mathrm{eff}} \Delta \boldsymbol{\psi}_i + \sigma \boldsymbol{\psi}_i\right)_{\boldsymbol{T}}, \\
\left(\boldsymbol{B}_1^{\mathrm{nsGLS}}\right)_{ij} &:= \underbrace{-b_h\left(\boldsymbol{\psi}_i, \phi_j\right)}_{=\left(\boldsymbol{B}^T\right)_{ij}} + \alpha \sum_{\boldsymbol{T} \in \mathcal{T}_h} \alpha_{\boldsymbol{T}} \left(\nabla \phi_j, -\mu_{\mathrm{eff}} \Delta \boldsymbol{\psi}_i + \sigma \boldsymbol{\psi}_i\right)_{\boldsymbol{T}}, \\
\left(\boldsymbol{B}_2^{\mathrm{nsGLS}}\right)_{ij} &:= \underbrace{b_h\left(\boldsymbol{\psi}_j, \phi_i\right)}_{=-(\boldsymbol{B})_{ij}} + \alpha \sum_{\boldsymbol{T} \in \mathcal{T}_h} \alpha_{\boldsymbol{T}} \left(-\mu_{\mathrm{eff}} \Delta \boldsymbol{\psi}_j + \sigma \boldsymbol{\psi}_j, \nabla \phi_i\right)_{\boldsymbol{T}}, \\
\left(\boldsymbol{C}^{\mathrm{nsGLS}}\right)_{ij} &:= \alpha \sum_{\boldsymbol{T} \in \mathcal{T}_h} \alpha_{\boldsymbol{T}} \left(\nabla \phi_j, \nabla \phi_i\right)_{\boldsymbol{T}}, \\
\left(\boldsymbol{F}_1^{\mathrm{nsGLS}}\right)_i &:= \underbrace{\left(\boldsymbol{f}, \boldsymbol{\psi}_i\right)}_{=(\boldsymbol{F}_1)_i} + \alpha \sum_{\boldsymbol{T} \in \mathcal{T}_h} \alpha_{\boldsymbol{T}} \left(\boldsymbol{f}, -\mu_{\mathrm{eff}} \Delta \boldsymbol{\psi}_i + \sigma \boldsymbol{\psi}_i\right)_{\boldsymbol{T}}, \\
\left(\boldsymbol{F}_2^{\mathrm{nsGLS}}\right)_i &:= \underbrace{\left(g, \phi_i\right)}_{=(\boldsymbol{F}_2)_i} + \alpha \sum_{\boldsymbol{T} \in \mathcal{T}_h} \alpha_{\boldsymbol{T}} \left(\boldsymbol{f}, \nabla \phi_i\right)_{\boldsymbol{T}}.
\end{aligned}$$

Apparently, the matrices $\boldsymbol{A}^{\mathrm{nsGLS}}$ and $\boldsymbol{C}^{\mathrm{nsGLS}}$ are symmetric for any $\alpha \geqslant 0$. However, for $\alpha > 0$, the coefficient matrix $\mathcal{A}^{\mathrm{nsGLS}}$ is not symmetric in general, since then $-\boldsymbol{B}_2^{\mathrm{nsGLS}} \neq \left(\boldsymbol{B}_1^{\mathrm{nsGLS}}\right)^T$. This demonstrates how the Douglas–Wang method destroys the symmetry of the standard Galerkin method. The lack of symmetry should be kept in mind for choosing an appropriate solver for the linear system of equations.

In the special case of linear polynomials, i.e., $\boldsymbol{V}_h = \mathbb{P}_1$, all second order derivatives vanish. If additionally the case $\sigma = 0$ is considered (Stokes limit), then, in accordance with Remark 4.3.7, only the block matrix $\boldsymbol{C}^{\mathrm{nsGLS}}$ and the block vector $\boldsymbol{F}_2^{\mathrm{nsGLS}}$ contain stabilization terms such that symmetry can be recovered by testing the second block row with $-\phi_i$.

Let us mention that Lemma 4.2.1 is in general not valid for stabilized methods which result in a nonzero lower right block matrix, here $\boldsymbol{C}^{\mathrm{nsGLS}}$.

**Remark 4.3.8** (Block Row Considerations)**.**
*Similar to Remark 4.2.2, the system* (4.3.5) *can be considered block row-wise, i.e.,*

$$(4.3.5) \quad \Longleftrightarrow \quad \begin{Bmatrix} \boldsymbol{A}^{\mathrm{nsGLS}}\boldsymbol{U} + \boldsymbol{B}_1^{\mathrm{nsGLS}}\boldsymbol{P} = \boldsymbol{F}_1^{\mathrm{nsGLS}} \\ \boldsymbol{B}_2^{\mathrm{nsGLS}}\boldsymbol{U} + \boldsymbol{C}^{\mathrm{nsGLS}}\boldsymbol{P} = \boldsymbol{F}_2^{\mathrm{nsGLS}} \end{Bmatrix}.$$

*Let us assume that the block matrix $\boldsymbol{A}^{\mathrm{nsGLS}}$ is invertible, then we can transform the first block row into*

$$\boldsymbol{U} = \left(\boldsymbol{A}^{\mathrm{nsGLS}}\right)^{-1}\left(\boldsymbol{F}_1^{\mathrm{nsGLS}} - \boldsymbol{B}_1^{\mathrm{nsGLS}}\boldsymbol{P}\right).$$

*Multiplying by $\boldsymbol{B}_2^{\mathrm{nsGLS}}$ and using the second block row results in*

$$\left(\boldsymbol{B}_2^{\mathrm{nsGLS}}\left(\boldsymbol{A}^{\mathrm{nsGLS}}\right)^{-1}\boldsymbol{B}_1^{\mathrm{nsGLS}} - \boldsymbol{C}^{\mathrm{nsGLS}}\right)\boldsymbol{P} = \boldsymbol{B}_2^{\mathrm{nsGLS}}\left(\boldsymbol{A}^{\mathrm{nsGLS}}\right)^{-1}\boldsymbol{F}_1^{\mathrm{nsGLS}} - \boldsymbol{F}_2^{\mathrm{nsGLS}}.$$

*This equation is a problem for the pressure, which has a unique solution if and only if $\boldsymbol{B}_2^{\mathrm{nsGLS}}\left(\boldsymbol{A}^{\mathrm{nsGLS}}\right)^{-1}\boldsymbol{B}_1^{\mathrm{nsGLS}} - \boldsymbol{C}^{\mathrm{nsGLS}}$ is non-singular. We observe that the lower right block matrix $\boldsymbol{C}^{\mathrm{nsGLS}}$ appears here, in contrast to the standard Galerkin method in Remark 4.2.2.*

#### Stabilized Equal-Order Formulations for the Brinkman Problem

The pressure stabilizations discussed in the current subsection have been originally introduced for Stokes problems, but the idea to use residual-based equal-order stabilizations has been also applied to the Darcy problem in mixed form. In [MH02] a symmetric GLS method (as in Rem. 3.2.2) with $\alpha\alpha_{\boldsymbol{T}} = (2\sigma)^{-1}$ is considered, and stability and convergence with respect to a scaled $\boldsymbol{L}^2 \times H^1$-semi norm are proven. They propose to use a mesh-dependent grad-div stabilization $\frac{\sigma h^2}{2}(\nabla \cdot \boldsymbol{u}_h, \nabla \cdot \boldsymbol{v}_h)$ to increase the control on the velocity. However, the mesh-dependence weakens the resulting norm such that not $\boldsymbol{u}_h$ in $\boldsymbol{H}^{\mathrm{div}}$, but only $h^2\|\nabla \cdot \boldsymbol{u}_h\|_0^2$ is controlled.

In [BV02], a symmetric GLS approach (see Rem. 4.3.6) for equal-order finite elements and a generalized Stokes problem with a stabilization parameter of the form $\alpha_{\boldsymbol{T}} = \frac{c_1 h_{\boldsymbol{T}}^2}{\sigma h_{\boldsymbol{T}}^2 + c_2 \mu_{\mathrm{eff}}}$ is discussed. Some years later, [JS10] and [HJS11] analyze and test a similar symmetric GLS approach with $\alpha_{\boldsymbol{T}} = \frac{h_{\boldsymbol{T}}^2}{t^2 + h_{\boldsymbol{T}}^2}$ for a scaled Brinkman problem (as in Rem. 3.2.2). The stability is here shown for stabilization parameters $\alpha \in \left(0, \min\left\{\frac{1}{2\hat{c}_I^2}, \frac{1}{2}\right\}\right)$ with $\hat{c}_I$ denoting the constant in the inverse estimate (4.1.3b). The authors base their analysis and simulation in the Darcy limit on the mixed primal formulation (see Rem. 3.3.4) with natural boundary conditions, i.e., control on the divergence of the velocity is not considered.

Later on, in [BC09], [BC10], [BC12], and [Cod15], (hybrid) stabilized finite element approaches capable of the Darcy and the Stokes problem are proposed. Thereby, the method referred to ASGS encompasses a grad-div stabilization, a reformulated symmetric GLS stabilization, and a pressure jump stabilization (that can be omitted for continuous pressure approximations). This multi-purpose stabilization allows to use arbitrary polynomial orders under the premise of conformity of the finite element spaces. The authors consider different choices of stabilization parameters that combine $\mu_{\text{eff}}$, $\sigma$, $h_{\boldsymbol{T}}$, and a length scale $\ell_{\boldsymbol{\Omega}}$, e.g., the diameter of the computational domain. They deduce that while the presence of the characteristic length is irrelevant as long as $\mu_{\text{eff}} \neq 0$, the choice is decisive in the Darcy limit and might vary with polynomial orders. Depending on the choice of the length scale (defined as the mesh-size, the domain size, or a combination) the formulation mimics the different functional settings, i.e., stability and convergence of the velocity versus the pressure is regulated. In fact, an appropriate choice allows to control $\boldsymbol{u}_h$ and $\nabla \cdot \boldsymbol{u}_h$ in a dimensionally correct way (i.e., such that both expressions can be incorporated in a single norm), yielding control of the velocity in the $\boldsymbol{H}^{\text{div}}$-norm, [Cod15].

**Remark 4.3.9** (Some Other Stabilizations in the Literature)**.**
*Several stabilizations, which do not fall into the category of pressure stabilizations discussed above, have been proposed with the purpose to construct unified methods for Stokes and Darcy problems:*

- $\mathbb{P}_1/\mathbb{P}_0$ *with an edge stabilization as described in Remark 4.3.4 is used in [BH07b], [DZ09], and [DZ11].*
- *In [Lam13], a modification of* $\mathbb{P}_1/\mathbb{P}_0$ *with bubble-enhanced velocities, and a pressure space on a dual mesh, together with grad-div-type stabilization (local-projection in the Darcy limit) is used.*
- *In [BH06b] the jump of normal gradients of the pressure and divergence jumps with respect to the velocity are stabilized for* $\mathbb{P}_1/\mathbb{P}_1$.
- *[BH05] and [RZ09] use* $\mathbb{P}_1^{nc}/\mathbb{P}_0$ *with a stabilization of normal velocity jumps (controlling the non-conformity in the Darcy limit) across all edges.*
- *In [BS11] a symmetric local projection stabilization for* $\mathbb{P}_k/\mathbb{P}_k$ *is proposed. It uses local (macro-element) projections of both, pressure gradients and the grad-div stabilization, such that local fluctuations are stabilized only. The method is not (strongly) consistent but the consistency error does not destroy the optimal order of convergence.*
- *In [KS11] and [KS12], stabilized* $H^{\text{div}}$*-conforming approaches are considered. In order to control the non-conformity in the Stokes case (discontinuity of the tangential velocity across inner facets) a symmetric interior penalty Galerkin method based on the ideas of Nitsche [Nit71] is proposed.*
- *A discussion of various pressure-projection stabilizations is given in [Bur08].*

*Let us finally note that pressure stabilizations are often used in combination with other stabilizations in order to heal multi-purpose problems. Typical supplements are the grad-div stabilization, a stabilization accommodating discontinuous pressures (pressure jump stabilization), and upwinding (SUPG) for convection-dominated flow regimes, see, e.g., [FF92], [DW89], [BC09]. For*

*further discussions on stabilizations for mixed problems see [QV08, Sect. 9.4], [BB01], [BH06b] [RST08, Part III, Sect. 3], [Lub06], [Bra+07], and [FHS93].*

## 4.4 The Weak Imposition of Essential Boundary Conditions

The purpose of this section is to discuss treatments of essential boundary conditions which allow the definition of a finite element method for the Brinkman problem that is robust with respect to the whole range of physical parameters. Such a formulation then has to be able to comprise the appropriate behavior in case of a possibly strong imbalance between the highest order derivative term $\mu_{\text{eff}}\Delta\boldsymbol{u}$ and the term $\sigma\boldsymbol{u}$, and especially in the limit cases.

The stabilized equal-order finite element approach (4.3.4) is a promising candidate for a robust Brinkman formulation. However, the imposition of the essential boundary conditions, which change in dependence of the physical setting as described in (3.2.8c), remain an open issue. While the normal and tangential velocity are set in the Stokes regime ($\mu_{\text{eff}} > 0$), only the normal velocity is controlled in the Darcy limit ($\mu_{\text{eff}} = 0$). This variation has to be incorporated appropriately into a robust method as well.

### Strong Imposition of Essential Boundary Conditions

In the standard finite element context, the so-called *strong imposition of essential boundary conditions*, which we will also refer to as the *classical approach*, realizes their explicit incorporation into the ansatz space. In other words, the discrete spaces are conforming with respect to the boundary condition.

Concerning the Brinkman problem, the strong imposition of the boundary conditions (3.2.8c) does not allow to devise a single discrete method that is well-posed for the whole range of parameters simultaneously, i.e, for all $\mu_{\text{eff}} > 0$, corresponding to a condition on $\boldsymbol{u}|_\Gamma$, and for $\mu_{\text{eff}} = 0$, corresponding to a condition on $(\boldsymbol{u}\cdot\boldsymbol{n})\,|_\Gamma$ only. Moreover, it is (physically) desirable to allow a smoother, parameter-driven transition between the boundary conditions in the limit cases, in contrast to the abrupt change which is realized when simply switching the solution space (and tied to that the boundary condition) in the limit $\mu_{\text{eff}} = 0$.

Apart from that, in certain situations (e.g., in presence of boundary layers), a procedure that does not strictly enforce the ansatz functions to fulfill the essential boundary conditions is attractive. Relaxing the boundary conditions allows for example to obtain satisfactory solutions without resolving boundary layers.

### Weak Imposition of Essential Boundary Conditions

The resulting request for alternative treatments of essential boundary conditions in the finite element context, in particular concerning our objective with respect

to the Brinkman problem, motivates a discussion about possibilities of *weak imposition*.

We will focus on the Nitsche method but, for the sake of completeness, other popular options are briefly discussed now. In the 1970's Babuska proposed two methods that are appropriate for the weak imposition, the *method of Lagrange multipliers* in [Bab72] and the *penalty method* in [Bab73]. The idea of the first one is to introduce an additional unknown variable, the Lagrange multiplier, which is equipped with its own finite element space on the boundary. This approach constructs a saddle point structure and therefore, in order to assure well posedness, the discrete multiplier space has to be chosen such that it is compatible with the other finite element space(s) in the sense that an inf-sup condition is valid, for details see [Bab72], [BZ12], [Ste95], [BH92], [Ver87], and [Ver91]. Moreover, the increase of the computational effort on the discrete level due to the additional unknown is disadvantageous. The idea of the second alternative, the penalty method, is to penalize (in the sense of $L^2$) deviations from the boundary condition. This approach does not require the definition of additional finite element spaces, but the penalty method is not consistent, yielding a consistency error which then has to taken into account, see, e.g., [BZ12] and [BE86]. The weight of the penalization should be chosen large enough such that the resulting solution satisfies the boundary condition sufficiently well, but, since the penalty parameter has a strong impact on the spectral condition number of the resulting finite element matrix, it should not be too large, see [BF91, Ch. II., §4] and [JS09].

A family of techniques referred to as *Nitsche methods*, diminishes the disadvantages of the penalty and the Lagrange multiplier methods and thus bears good prospects for the goal of this treatise.

As a popular ingredient, Nitsche-type methods are used in *discontinuous Galerkin methods* (DG, [CKS12, pp. 137]) for introducing inter-element conditions, there usually referred to as *interior penalties*, see [KS11], [KS12], [Bur12], and references therein. For the particular purpose of the weak imposition of essential boundary conditions, the Nitsche methods are analyzed in [Ste95] and [JS09]. Current fields of application include contact problems (multiple domains with different materials), domain decomposition approaches, and fictitious domain methods. For related literature see, e.g., [Leh16], [Ste98], [BH07b], [DZ09], [BHS03], and [HH02] (treating interface problems, mortar methods (non-matching grids), and unfitted meshes), and [Sch$^+$16], [BH12], and [Boi$^+$18] (considering immersed/cut finite elements, fictitious domain methods, and moving boundaries).

The rest of this section is dedicated to the most popular variants of the Nitsche method. Following the chronological order, we will finally approach the *penalty-free non-symmetric Nitsche method*, which will be one of the key ingredients of the robust finite element method presented and analyzed in Chapter 5.

### 4.4.1 The Classical Nitsche Method

The **classical** or **symmetric Nitsche method** was originally introduced in 1971 in [Nit71] for the weak imposition of essential boundary conditions for a Poisson problem in two dimensions. The method can be defined within the discrete formulation according to the following two steps:

I) add the weak form of the essential boundary conditions weighted with a parameter $\gamma_N > 0$ (which will be referred to as the *Nitsche penalty*),
II) add consistent boundary terms to recover respectively enforce symmetry of the resulting bilinear form.

Step I) is the penalty method itself, i.e., from that point of view, the classical Nitsche method can be interpreted as an enhanced penalty approach, such that strong consistency is recovered. For the scalar *Poisson problem*

$$\begin{aligned} -\Delta u &= f \quad && \text{in } \boldsymbol{\Omega}, \\ u &= u_e \quad && \text{on } \Gamma, \end{aligned} \tag{4.4.1}$$

Nitsche analyzed the discrete problem:
Find $u_h \in V_h^k$ such that

$$a_{h,N}\left(u_h, v_h\right) = f_{h,N}\left(v_h\right), \quad \forall v_h \in V_h^k,$$

with left-hand side

$$\begin{aligned} a_{h,N}\left(u_h, v_h\right) := \left(\nabla u_h, \nabla v_h\right) &- \langle \nabla u_h \cdot \boldsymbol{n}, v_h \rangle_\Gamma \\ &\underbrace{-\langle \nabla v_h \cdot \boldsymbol{n}, u_h \rangle_\Gamma}_{\text{symmetry term}} + \underbrace{\gamma_N \sum_{E \in \mathcal{G}_h} h_E^{-1} \langle u_h, v_h \rangle_E}_{\text{Nitsche penalty term}}, \end{aligned} \tag{4.4.2}$$

the corresponding consistency terms on the right-hand side

$$f_{h,N}\left(v_h\right) := \left(f, v_h\right) \underbrace{-\langle \nabla v_h \cdot \boldsymbol{n}, u_e \rangle_\Gamma + \gamma_N \sum_{E \in \mathcal{G}_h} h_E^{-1} \langle u_e, v_h \rangle_E}_{\text{consistency terms}},$$

and

$$V_h^k := \left\{ v_h \in H^1\left(\boldsymbol{\Omega}\right) \ : \ v_h|_{\boldsymbol{T}} \in \mathbb{P}_k\left(\boldsymbol{T}\right), \quad \forall \boldsymbol{T} \in \mathcal{T}_h \right\}, k \geqslant 1.$$

This finite element space does not contain any information on the solution at the boundary of the domain. In order to prove that (4.4.2) provides sufficient control on the behavior at the boundary, a norm, which takes into account boundary terms has to be used in the analysis:

$$\|v_h\|_{1,h}^2 := \|\nabla v_h\|_0^2 + \sum_{E \in \mathcal{G}_h} h_E^{-1} \|v_h\|_{0,E}^2.$$

Coercivity of $a_{h,N}\left(\cdot,\cdot\right)$ then can be proven (see Thm. 3.1.12), if the Nitsche penalty $\gamma_N$ is chosen to be sufficiently large. Using the Cauchy–Schwarz inequality (Thm. 3.1.5), the Cauchy–Schwarz inequality for sums (Lem. 3.1.2),

the discrete trace-inverse inequality (4.1.6c), and applying the Young inequality (Thm. 3.1.1) yields

$$\begin{aligned}
-2\langle \nabla u_h \cdot \boldsymbol{n}, u_h \rangle &\geqslant -2 \sum_{E \in \mathcal{G}_h} h_E^{\frac{1}{2}} \|\nabla u_h \cdot \boldsymbol{n}\|_{0,E} h_E^{-\frac{1}{2}} \|u_h\|_{0,E} \\
&\geqslant -2 \left( \sum_{E \in \mathcal{G}_h} h_E \|\nabla u_h \cdot \boldsymbol{n}\|_{0,E}^2 \right)^{\frac{1}{2}} \left( \sum_{E \in \mathcal{G}_h} h_E^{-1} \|u_h\|_{0,E}^2 \right)^{\frac{1}{2}} \\
&\geq -2 \left( \widetilde{c}_{DTI} \|\nabla u_h\|_0^2 \right)^{\frac{1}{2}} \left( \sum_{E \in \mathcal{G}_h} h_E^{-1} \|u_h\|_{0,E}^2 \right)^{\frac{1}{2}} \\
&\geqslant -\left( \frac{1}{\varepsilon} \|\nabla u_h\|_0^2 + \varepsilon\, \widetilde{c}_{\mathrm{DTI}} \sum_{E \in \mathcal{G}_h} h_E^{-1} \|u_h\|_{0,E}^2 \right).
\end{aligned}$$

Testing (4.4.2) with $v_h = u_h$ and using the above estimate, we obtain

$$\begin{aligned}
a_{h,N}(u_h, u_h) &= \|\nabla u_h\|_0^2 - 2\langle \nabla u_h \cdot \boldsymbol{n}, u_h \rangle + \gamma_N \sum_{E \in \mathcal{G}_h} h_E^{-1} \|u_h\|_{0,E}^2 \\
&\geqslant \left(1 - \frac{1}{\varepsilon}\right) \|\nabla u_h\|_0^2 + (\gamma_N - \varepsilon\, \widetilde{c}_{\mathrm{DTI}}) \sum_{E \in \mathcal{G}_h} h_E^{-1} \|u_h\|_{0,E}^2 \\
&\geqslant \min\left\{ 1 - \frac{1}{\varepsilon}, \gamma_N - \varepsilon\, \widetilde{c}_{\mathrm{DTI}} \right\} \|u_h\|_{1,h}^2.
\end{aligned}$$

We can choose an $\varepsilon > 1$ such that the condition

$$\gamma_N > \widetilde{c}_{DTI} \tag{4.4.4}$$

guarantees the existence of a strictly positive coercivity constant yielding well-posedness due to the (discrete) theorem of Lax–Milgram (Thm. 3.1.12).

The constant $\widetilde{c}_{\mathrm{DTI}}$ depends on the regularity of the mesh as well as on the polynomial approximation order, see, e.g., [Han05] and [BHS03]. Without further effort, the value of $\widetilde{c}_{\mathrm{DTI}}$ is in practice a priori unknown. If the Nitsche penalty parameter $\gamma_N$ is chosen too small (violating (4.4.4)), stability is not guaranteed anymore. Too large penalties instead result in degenerated error bounds and a growth of the condition number of the finite element matrix, [Leh16]. In fact, the question is now how to assure, that the theoretical bound on the penalty parameter which ensures stability, is not violated. An estimate for $\gamma_N$ can be obtained by solving certain eigenvalue problems as described in [PLM18] or via a trial and error procedure. The presence of such an influential stabilization parameter is a drawback of the method.

**Remark 4.4.1** (On the Classical Nitsche Method)**.**

- *The classical Nitsche method for the Poisson problem yields optimal convergence in the mesh-dependent energy norm $\|\cdot\|_{1,h}$ and in the $L^2$-norm. If $\gamma_N$ is chosen to obtain a stable method (via (4.4.4)), $u \in H^{k+1}(\boldsymbol{\Omega})$, and*

*$u_h \in V_h^k$, then there exist two constants $C_1, C_2 > 0$ such that the a priori estimates*

$$\|u - u_h\|_{1,h} \leqslant C_1 h^k \|u\|_{k+1} \quad \text{and} \quad \|u - u_h\|_0 \leqslant C_2 h^{k+1} \|u\|_{k+1}$$

*hold (see, e.g., [BZ12]).*

- *In [BH07a], for turbulent flows (high Reynolds number) with unresolved boundary layers, the mean flow in the boundary layer was observed to be more accurately captured using the classical Nitsche method (in particular on coarse meshes) instead of the strong imposition of boundary conditions.*

### 4.4.2 The (Penalty-Free) Non-Symmetric Nitsche Method

In 1995, Freund and Stenberg proposed in [SF95] the **non-symmetric Nitsche method** and applied it to a Poisson problem and a convection-diffusion-reaction problem. Their motivation was mainly the non-symmetric nature of convection-diffusion-reaction problems combined with the desire to weaken the stability requirements concerning the Nitsche penalty (condition (4.4.4)). For the Poisson problem (4.4.1), the non-symmetric Nitsche method reads as follows:
Find $u_h \in V_h^k$ such that

$$a_{h,N}^{ns}(u_h, v_h) = f_{h,N}^{ns}(v_h), \quad \forall v_h \in V_h^k,$$

with

$$a_{h,N}^{ns}(u_h, v_h) := (\nabla u_h, \nabla v_h) - \langle \nabla u_h \cdot \boldsymbol{n}, v_h \rangle_\Gamma \underbrace{+ \langle \nabla v_h \cdot \boldsymbol{n}, u_h \rangle_\Gamma}_{\text{non-symmetry term}} \underbrace{+ \gamma_N \sum_{E \in \mathcal{G}_h} h_E^{-1} \langle u_h, v_h \rangle_E}_{\text{Nitsche penalty term}}$$

and

$$f_{h,N}^{ns}(v_h) := (f, v_h) \underbrace{+ \langle \nabla v_h \cdot \boldsymbol{n}, u_e \rangle_\Gamma + \gamma_N \sum_{E \in \mathcal{G}_h} h_E^{-1} \langle u_e, v_h \rangle_E}_{\text{consistency terms}}.$$

Formally, the only difference between the symmetric and the non-symmetric version is the sign in front of the third term in $a_{h,N}^{ns}(\cdot,\cdot)$ and the corresponding consistency term in $f_{h,N}^{ns}(\cdot)$. This effects the non-symmetry of the bilinear form $a_{h,N}^{ns}(\cdot,\cdot)$ such that unconditional coercivity can be proven. Testing with $v_h = u_h$, the skew-symmetric boundary terms cancel out, yielding

$$a_{h,N}^{ns}(u_h, u_h) = \|\nabla u_h\|_0^2 + \gamma_N \sum_{E \in \mathcal{G}_h} h_E^{-1} \|u_h\|_{0,E}^2 \geqslant \min\{1, \gamma_N\} \|u_h\|_{1,h}^2.$$

Remarkably, this change is enough to yield stability for all strictly positive values of the Nitsche penalty parameter, i.e.,

$$\gamma_N > 0.$$

### The Elimination of the Penalty Term

In a series of numerical tests in the discontinuous Galerkin context in [Hug+00], the idea to use the non-symmetric Nitsche method with Nitsche penalty $\gamma_N = 0$ was introduced (see also [Bur12]). As we have seen in the previous considerations, the proof of coercivity of $a_{h,N}(\cdot,\cdot)$, as well as $a^{ns}_{h,N}(\cdot,\cdot)$, enforces strict positivity of the Nitsche penalty. This suggests that the well-posedness for $\gamma_N = 0$, if given at all, has to rely on a different proof strategy.

In 2012, Burman succeeded in proving the well-posedness of the **penalty-free non-symmetric Nitsche method**, i.e., with $\gamma_N = 0$, for the Poisson and convection-diffusion-reaction problems (on a polygonal domain in 2D). In [Bur12] he showed that, although the (sufficient) coercivity argument used for the variants with penalty fails, inf-sup stability still can be proven for $\gamma_N = 0$.

For the Poisson problem (4.4.1), the penalty-free non-symmetric Nitsche formulation reads:
Find $u_h \in V_h^k$ such that

$$a^{nspf}_{h,N}(u_h, v_h) = f^{nspf}_{h,N}(v_h), \quad \forall v_h \in V_h^k,$$

with

$$a^{nspf}_{h,N}(u_h, v_h) := (\nabla u_h, \nabla v_h) - \langle \nabla u_h \cdot \boldsymbol{n}, v_h\rangle_\Gamma \underbrace{+\langle \nabla v_h \cdot \boldsymbol{n}, u_h\rangle_\Gamma}_{\text{non-symmetry term}},$$

$$f^{nspf}_{h,N}(v_h) := (f, v_h) \underbrace{+\langle \nabla v_h \cdot \boldsymbol{n}, u_e\rangle_\Gamma}_{\text{consistency term}}.$$

In essence, the key of Burman's proof is to regroup the boundary elements into (connected) patches $P_j$ with boundary parts $F_j := \partial P_j \cap \Gamma$ and construct a test function $\varphi^{u_h} \in V_h^1$ that patch-wise satisfies

$$\frac{1}{|F_j|}\int_{F_j} \nabla\varphi^{u_h} \cdot \boldsymbol{n} = \frac{1}{|F_j|h}\int_{F_j} u_h, \quad \forall j,$$

and allows to control the behavior at the boundary via the Nitsche term $\langle \nabla v_h \cdot \boldsymbol{n}, u_h\rangle_\Gamma$.

In [BB16], the well-posedness results are extended to linear elasticity problems (compressible and incompressible).

**Remark 4.4.2** (On the (Penalty-Free) Non-Symmetric Nitsche Method)**.**

- *Also the non-symmetric Nitsche method for the Poisson problem yields optimal convergence in the mesh-dependent energy norm $\|\cdot\|_{1,h}$ for $\gamma_N \geqslant 0$ (see Rem. 4.4.1). Due to the lack of adjoint consistency for the non-symmetric versions, the argument used for the $L^2$-error estimate (the Aubin–Nitsche technique) results in sub-optimality (of half an order), i.e.,*

$$\|u - u_h\|_0 \leqslant Ch^{k+\frac{1}{2}}\|u\|_{k+1},$$

*can be proven with a constant $C > 0$, see, e.g., [BZ12]. However, in [Bur12] it is reported that no example could be found such that the suboptimal*

*convergence order was exhibited for the penalty-free variant and that instead, in all numerical tests, the optimal convergence rate was attained on finer meshes.*

- *The non-symmetric Nitsche method with $\gamma_N \geqslant 0$ yields a non-symmetric system matrix, hence it is in particular well suited for problems which already admit a non-symmetric system (e.g., flow or transport problems, where a convective term is apparent).*
- *The penalty-free non-symmetric Nitsche method shall not be expected to be more accurate than the other variants. According to numerical tests in [Bur12], the presence of a penalty term for the boundary condition yields the smaller error in the $L^2$-norm. The impact of a variation of the penalty parameter in the non-symmetric Nitsche method is also tested with the result that it is negligible.*
- *The penalty-free non-symmetric Nitsche method can be interpreted as a Lagrange multiplier method, with the term $\nabla u_h \cdot \boldsymbol{n}$ playing the role of the Lagrange multiplier, see, e.g., [BB16], [Bur12], and [Ste95]. Hence, the multiplier space consists of normal gradients of the solution space (no additional space).*
- *In case of convection- or reaction-dominated convection-diffusion-reaction problems, where boundary layers that are much thinner than the element size might appear (outflow layers), the non-symmetric Nitsche method was noticed to reduce oscillations, see [SF95], [Bur12], and [BH07a].*

**The Penalty-Free Non-Symmetric Nitsche Method for the Brinkman Problem and Related Linear Algebra Aspects**

To derive the terms corresponding to the penalty-free non-symmetric Nitsche method for the Brinkman problem (3.2.8), let us consider a finite element space for the velocity that does not contain any constraints with respect to the behavior at the boundary. As a result, the boundary integrals, that appear due to integration by parts applied twice in the momentum balance of the Brinkman system (see (3.3.3) and (3.3.4)) are present, i.e., we have

$$-\langle \mu_{\text{eff}} \left( \nabla \boldsymbol{u}_h \cdot \boldsymbol{n} \right), \boldsymbol{v}_h \rangle + \langle p_h, \boldsymbol{v}_h \cdot \boldsymbol{n} \rangle.$$

For the penalty-free non-symmetric Nitsche method, their skew-symmetrizing complements are added in a consistent way as well, i.e.,

$$+\langle \mu_{\text{eff}} \left( \nabla \boldsymbol{v}_h \cdot \boldsymbol{n} \right), \boldsymbol{u}_h - \boldsymbol{0} \rangle - \langle q_h, (\boldsymbol{u}_h - \boldsymbol{0}) \cdot \boldsymbol{n} \rangle.$$

The usage of a Nitsche method necessitates to consider the degrees of freedom at the boundary. Let us assign the boundary terms that are present due to the penalty-free non-symmetric Nitsche method to the matrix blocks of the finite element matrix in the standard Galerkin formulation, (4.2.2) and (4.2.3). Using the same notation, we get the additional skew-symmetric matrix summand

$$\mathcal{A}^{\text{pfns}} := \begin{pmatrix} \boldsymbol{A}^{\text{pfns}} & \boldsymbol{B}_1^{\text{pfns}} \\ \boldsymbol{B}_2^{\text{pfns}} & \boldsymbol{0} \end{pmatrix} \in \mathbb{R}^{(N_u+N_p)\times(N_u+N_p)}$$

with

$$\left(\boldsymbol{A}^{\text{pfns}}\right)_{ij} := -\langle \mu_{\text{eff}} \left(\nabla \boldsymbol{\psi}_j \cdot \boldsymbol{n}\right), \boldsymbol{\psi}_i \rangle + \langle \mu_{\text{eff}} \left(\nabla \boldsymbol{\psi}_i \cdot \boldsymbol{n}\right), \boldsymbol{\psi}_j \rangle,$$
$$\left(\boldsymbol{B}_1^{\text{pfns}}\right)_{ij} := \langle \phi_j, \boldsymbol{\psi}_i \cdot \boldsymbol{n} \rangle,$$
$$\left(\boldsymbol{B}_2^{\text{pfns}}\right)_{ij} := -\langle \phi_i, \boldsymbol{\psi}_j \cdot \boldsymbol{n} \rangle.$$

Note that in case of non-homogeneous essential boundary conditions, the corresponding consistency terms have to be included in the right-hand side of the problem.

**Remark 4.4.3** (Nitsche Methods and Brinkman Type Problems)**.**
*The Nitsche method has been already considered for the imposition of essential boundary conditions in Brinkman-type problems. However, the clear focus of the publications - very few in number - is on the classical Nitsche method (symmetric with penalty). Therefore two penalty terms, one for the velocity in the Stokes regime and one for the normal velocity in the Darcy limit are introduced.*

- *In [HJ09] edge-jump pressure-stabilized (see Rem. 4.3.4) $\mathbb{P}_1/\mathbb{P}_0$ elements with a classical Nitsche method for the boundary conditions are used.*
- *In [DZ09] and [DZ11], a unified method for coupled Stokes–Darcy problems based on $\mathbb{P}_k/\mathbb{P}_{k-1}^{disc}$ with edge-jump stabilization of the pressure (see Rem. 4.3.4) and the classical Nitsche method for the weak imposition of boundary and interface conditions is presented.*
- *In [JS10] and [HJS11], $\mathbb{P}_k/\mathbb{P}_k$ for a scaled Brinkman problem (similar to that in Rem. 3.2.2) with a symmetric GLS stabilization (see Rem. 4.3.6) and the classical Nitsche method for boundary conditions is proposed.*[7] *Since the authors consider the primal (mixed) Darcy limit (see Rem. 3.3.4), the boundary condition on $\boldsymbol{u}_h \cdot \boldsymbol{n}$ is of natural type, thus not imposed by the Nitsche method.*

One of the main contributions of this thesis will be to extend the theoretical field of application of the penalty-free non-symmetric Nitsche method to the Brinkman problem (3.2.8) including its limit cases. To the authors' knowledge, there exists no stability analysis for the dual Darcy problem in this context. The stability proof for the incompressible elasticity problem in [BB16] is not trivial and will not be utilized.

## 4.5 Summary

This chapter was used to introduce the theoretical ingredients of a low-order, low-complexity, robust finite element method for the Brinkman problem.

A short review of finite element basics led us to the well-posedness criteria for mixed finite element methods, which are basically coercivity and/or inf-sup

[7] For some reason, only velocity-velocity terms are incorporated via the Nitsche method and no velocity-pressure couplings.

stability. In particular the latter represents a compatibility condition for the finite element spaces for the velocity and the pressure which is fulfilled by certain pairs only.

The standard Galerkin formulation for the Brinkman problem was stated and, focusing initially on the limit problems, typical stable Stokes and Darcy finite element pairs were presented. We have discussed difficulties that arise for the attempt to use them in the respective opposite limit, which has also revealed strategies for the construction of uniformly stable methods, often used in the literature.

Following the objective to develop a robust, low-order method, stability issues of spaces with low polynomial order were demonstrated and the (linear) equal-order interpolation was selected to be analyzed further. In order to enforce stability of $\mathbb{P}_k/\mathbb{P}_k$ without introducing additional dofs, residual-based pressure stabilizations, in particular the non-symmetric GLS (Douglas–Wang) method, were introduced. This stabilization has then been formulated for the Brinkman problem and its properties were investigated.

As a possibility to relax the typical strong imposition of essential boundary conditions, the idea of their weak imposition has been described. To that end the Nitsche methods were introduced, whose penalty-free non-symmetric variant is the method of choice. This approach potentially enables a parameter-driven transition between the limits of the Brinkman problem and is therefore a promising component of a robust method.

# 5 A Robust and Unconditionally Stable Finite Element Method for the Brinkman Problem using a Stabilized Penalty-Free Non-Symmetric Nitsche Approach

In this chapter we will propose a low-order finite element method for the Brinkman problem (3.2.8) where the crucial ingredients - chosen to yield robustness - will be the following:

(1) In order to keep the computational cost and complexity low, continuous finite element spaces, based on linear polynomials, i.e., $\boldsymbol{V}_h/Q_h := \mathbb{P}_1/\mathbb{P}_1$, will be considered. For the purpose of stabilization, the discrete formulation is modified. Among the different possibilities (see Section 4.3) we will utilize a pressure stabilization method, which is referred to as non-symmetric Galerkin least squares (GLS) method (sometimes also Douglas–Wang method).

(2) A grad-div stabilization is included such that the method admits robust stability with respect to the Darcy limit (in particular allowing control on the divergence of the velocity, see Subsection 4.2.2).

(3) In order to impose the essential boundary conditions, such that one single finite element formulation can be considered for the whole range of Brinkman problems, the penalty-free non-symmetric Nitsche method (see Section 4.4) is applied.

The first goal is to prove the robust well-posedness of the method. The second objective is to derive robust, optimal a priori error estimates in the energy norm.

After introducing preliminary results, the discrete formulation and an appropriate framework will be stated. In the subsequent stability analysis, the inf-sup criterion is investigated, which will be the basis for the error analysis.

The main results of this chapter have been published in [Bla$^+$18].

## 5.1 Tools for the Stability and A Priori Error Analysis

In this section, several results that will be used to proof stability theorems and error estimates will be presented.

### Interpolation Estimates

The approximation quality of a finite element space is restricted by the best approximation. Error bounds for interpolation operators from Sobolev spaces onto finite element spaces are therefore an essential ingredient for the a priori error analysis.

Based on the notation introduced in Section 4.1, let us introduce interpolation operators onto the finite element space

$$\mathcal{P}_h^k := \left\{ \boldsymbol{v}_h \in \boldsymbol{C}\left(\overline{\boldsymbol{\Omega}}\right) \ : \ \boldsymbol{v}_h|_{\boldsymbol{T}} \in \mathbb{P}_k\left(\boldsymbol{T}\right), \quad \forall \boldsymbol{T} \in \mathcal{T}_h \right\},$$

where $\{\mathcal{T}_h\}_{h>0}$ is a family of shape-regular triangulations of $\boldsymbol{\Omega}$.

We denote by $I_{h,k}^{\mathrm{La}} \colon \boldsymbol{C}\left(\overline{\boldsymbol{\Omega}}\right) \longrightarrow \mathcal{P}_h^k$ the *Lagrange (nodal) interpolation operator*. Note that this nodal interpolation operator is also well defined on $\boldsymbol{H}^2\left(\boldsymbol{\Omega}\right)$, since we have Sobolev's imbedding $\boldsymbol{H}^r\left(\boldsymbol{\Omega}\right) \subset \boldsymbol{C}^s\left(\overline{\boldsymbol{\Omega}}\right)$ for $\boldsymbol{\Omega} \subset \mathbb{R}^n$, provided $\frac{n}{2} < r - s$ is valid, see [Cia02, Ch. 3]. Then there exists a constant $c_{\mathrm{La}} > 0$ such that for $1 \leqslant k$ it holds [EG04, Thm. 1.109]:

$$\left\|\boldsymbol{v} - I_{h,k}^{\mathrm{La}}\left(\boldsymbol{v}\right)\right\|_{0,\boldsymbol{\Omega}} + \sum_{m=1}^{2} h^m \left( \sum_{\boldsymbol{T} \in \mathcal{T}_h} \left|\boldsymbol{v} - I_{h,k}^{\mathrm{La}}\left(\boldsymbol{v}\right)\right|_{m,\boldsymbol{T}}^2 \right)^{\frac{1}{2}} \leqslant c_{\mathrm{La}} h^2 \left|\boldsymbol{v}\right|_{2,\boldsymbol{\Omega}}, \quad (5.1.1)$$

for all $\boldsymbol{v} \in \boldsymbol{H}^2\left(\boldsymbol{\Omega}\right)$.

Moreover, let us denote with $I_{h,k}^{\mathrm{SZ}} \colon \boldsymbol{H}^l\left(\boldsymbol{\Omega}\right) \longrightarrow \mathcal{P}_h^k$ the *Scott–Zhang interpolation operator*, which is in particular well-defined for functions in $\boldsymbol{H}^1\left(\boldsymbol{\Omega}\right)$ and preserves essential boundary conditions on $\Gamma$, see e.g., [Ape99, Sect. 3], [EG04, Thm. 1.130], [BS08, pp. 121], and [SZ90]. Then, for $l, m \in \mathbb{N}_0$, with $1 \leqslant l$, there exist two constants $c_{\mathrm{SZ}}, \overline{c}_{\mathrm{SZ}} > 0$, depending on the geometry and on the shape regularity, such that the following interpolation estimates hold:

$$\forall\, 0 \leqslant m \leqslant 1\colon \quad \left\|I_{h,k}^{\mathrm{SZ}}\left(\boldsymbol{v}\right)\right\|_{m,\boldsymbol{\Omega}} \leqslant c_{\mathrm{SZ}} \|\boldsymbol{v}\|_{l,\boldsymbol{\Omega}}, \quad \forall \boldsymbol{v} \in \boldsymbol{H}^l\left(\boldsymbol{\Omega}\right), \forall h, \quad (5.1.2a)$$

$$\forall\, 0 \leqslant m \leqslant l \leqslant k+1\colon \quad \left|\boldsymbol{v} - I_{h,k}^{\mathrm{SZ}}\left(\boldsymbol{v}\right)\right|_{m,\boldsymbol{T}} \leqslant \overline{c}_{\mathrm{SZ}} h_{\boldsymbol{T}}^{l-m} \left|\boldsymbol{v}\right|_{l,S(\boldsymbol{T})}, \quad (5.1.2b)$$
$$\forall \boldsymbol{v} \in \boldsymbol{H}^l\left(S\left(\boldsymbol{T}\right)\right), \forall h, \forall \boldsymbol{T} \in \mathcal{T}_h,$$

where $S\left(\boldsymbol{T}\right)$ denotes the union of all cells in $\mathcal{T}_h$ that have a vertex in common with $\boldsymbol{T}$.

For a function $p \in L_0^2\left(\boldsymbol{\Omega}\right)$ we can consider the *modified Scott–Zhang interpolation operator* $I_{h,k}^{\mathrm{SZ}^{\int 0}} \colon H^1\left(\boldsymbol{\Omega}\right) \cap L_0^2\left(\boldsymbol{\Omega}\right) \longrightarrow \mathcal{P}_h^k \cap L_0^2\left(\boldsymbol{\Omega}\right)$ given by

$$I_{h,k}^{\mathrm{SZ}^{\int 0}}(p) := I_{h,k}^{\mathrm{SZ}}\left(p\right) - c \quad \text{with} \quad c := \frac{1}{|\boldsymbol{\Omega}|} \int_{\boldsymbol{\Omega}} I_{h,k}^{\mathrm{SZ}}\left(p\right) \ \in \mathbb{R}.$$

This definition guarantees that the average property is preserved, namely $\int_{\boldsymbol{\Omega}} I_{h,k}^{\mathrm{SZ}^{\int 0}}(p) = 0$. Using the triangle inequality, the definition of $c \in \mathbb{R}$ (in

particular the fact that its derivatives vanish), and the Cauchy–Schwarz inequality (Thm. 3.1.5), we obtain similar estimates (as for $I_{h,k}^{\mathrm{SZ}}$) for the error

$$\begin{aligned}\|p - I_{h,k}^{\mathrm{SZ}^{\int 0}}(p)\|_{m,\boldsymbol{\Omega}} &= \left\|p - \left(I_{h,k}^{\mathrm{SZ}}(p) - c\right)\right\|_{m,\boldsymbol{\Omega}} \leqslant \left\|p - I_{h,k}^{\mathrm{SZ}}(p)\right\|_{m,\boldsymbol{\Omega}} + \|c\|_{m,\boldsymbol{\Omega}} \\ &= \left\|p - I_{h,k}^{\mathrm{SZ}}(p)\right\|_{m,\boldsymbol{\Omega}} + |c|\,\|1\|_{0,\boldsymbol{\Omega}} \\ &= \left\|p - I_{h,k}^{\mathrm{SZ}}(p)\right\|_{m,\boldsymbol{\Omega}} + |\boldsymbol{\Omega}|^{-\frac{1}{2}}\left|\left(I_{h,k}^{\mathrm{SZ}}(p) - p, 1\right)\right| \\ &\leqslant \left\|p - I_{h,k}^{\mathrm{SZ}}(p)\right\|_{m,\boldsymbol{\Omega}} + \left\|I_{h,k}^{\mathrm{SZ}}(p) - p\right\|_{0,\boldsymbol{\Omega}} \\ &\leqslant 2\left\|p - I_{h,k}^{\mathrm{SZ}}(p)\right\|_{m,\boldsymbol{\Omega}}.\end{aligned}$$

Note that in the third line we have subtracted $\int_{\boldsymbol{\Omega}} p$, which is zero by assumption $p \in L_0^2(\boldsymbol{\Omega})$.

### Quadrature on an Interval

**Theorem 5.1.1** (Simpson's Rule)**.**
*The integral of a cubic polynomial $v\colon [a,b] \longrightarrow \mathbb{R}$, on an interval $[a,b] \subset \mathbb{R}$ can be computed exactly using the quadrature formula called Simpson's rule:*

$$\int_{[a,b]} v(x)\,\mathrm{d}x = \frac{(b-a)}{6}\left(v(a) + 4v\left(\frac{a+b}{2}\right) + v(b)\right). \tag{5.1.3}$$

*Proof:* The equality is straightforward by simply using a general definition of a polynomial of degree three. □

### Scaling Arguments - Coordinate Transformations, Linear Polynomials, and Triangles

For the analysis it will be useful to employ several (in)equalities based on so-called *scaling arguments*. These concern the integrals of finite element functions and the integrals of their derivatives, and can be derived via coordinate transformations.

Let us define the *reference triangle* $\widehat{\boldsymbol{T}} \subset \mathbb{R}^2$ by

$$\begin{aligned}\widehat{\boldsymbol{T}} &:= \left\{(\xi,\eta)^T \in \mathbb{R}^2 \,:\, 0 \leqslant \xi \leqslant 1, 0 \leqslant \eta \leqslant 1-\xi\right\} \\ &= \mathrm{conv}\left\{(0,0)^T, (1,0)^T, (0,1)^T\right\},\end{aligned}$$

see Figure 5.1.

For any non-degenerate triangle $\boldsymbol{T} := \mathrm{conv}\{\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2\} \subset \mathbb{R}^2$, with $\boldsymbol{x} = (x,y)^T$, we can introduce an affine linear (bijective) map that satisfies w.l.o.g. $\mathcal{F}_{\boldsymbol{T}}(0,0) = (x_0,y_0)^T$, $\mathcal{F}_{\boldsymbol{T}}(1,0) = (x_1,y_1)^T$, and $\mathcal{F}_{\boldsymbol{T}}(0,1) = (x_2,y_2)^T$:

$$\begin{aligned}\mathcal{F}_{\boldsymbol{T}}\colon \quad \widehat{\boldsymbol{T}} &\longrightarrow \boldsymbol{T}, \\ (\xi,\eta)^T \mapsto\ (x,y)^T &= \mathcal{F}_{\boldsymbol{T}}(\xi,\eta) := \xi\boldsymbol{x}_1 + \eta\boldsymbol{x}_2 + (1-\xi-\eta)\,\boldsymbol{x}_0 \\ &= \boldsymbol{B}_{\boldsymbol{T}}\begin{pmatrix}\xi\\ \eta\end{pmatrix} + \begin{pmatrix}x_0\\ y_0\end{pmatrix},\end{aligned} \tag{5.1.4a}$$

where

$$\boldsymbol{B_T} := \begin{pmatrix} x_1 - x_0 & x_2 - x_0 \\ y_1 - y_0 & y_2 - y_0 \end{pmatrix}, \tag{5.1.4b}$$

see Figure 5.1. In fact, the transformation matrix $\boldsymbol{B_T}$ is non-singular, i.e., $\det(\boldsymbol{B_T}) \neq 0$.

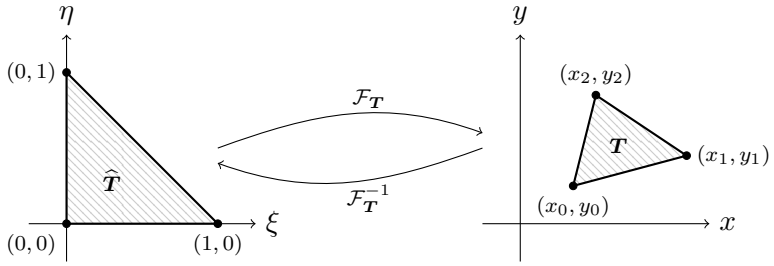



Figure 5.1: Reference map $\mathcal{F}_{\boldsymbol{T}}$ and its inverse $\mathcal{F}_{\boldsymbol{T}}^{-1}$.

Using $\mathcal{F}_{\boldsymbol{T}}$, we can apply the *integral transform* (see, e.g., [Joh16, Sect. B.5], also referred to as *integration by substitution* or *change of variables*)

$$\int_{\mathcal{F}_{\boldsymbol{T}}(\widehat{\boldsymbol{T}})} v(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\widehat{\boldsymbol{T}}} \widehat{v}(\widehat{\boldsymbol{x}})\,|\det(J\mathcal{F}_{\boldsymbol{T}}(\widehat{\boldsymbol{x}}))|\,\mathrm{d}\widehat{\boldsymbol{x}}, \tag{5.1.5}$$

with $\widehat{\boldsymbol{x}} := (\xi, \eta)^T$, $\widehat{v}(\widehat{\boldsymbol{x}}) := v(\mathcal{F}_{\boldsymbol{T}}(\widehat{\boldsymbol{x}}))$, $J$ denoting the Jacobian, and $v\colon \boldsymbol{T} \longrightarrow \mathbb{R}$ integrable. Via (5.1.4), we can deduce

$$J\mathcal{F}_{\boldsymbol{T}}(\widehat{\boldsymbol{x}}) := \begin{pmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial \eta} \\ \frac{\partial y}{\partial \xi} & \frac{\partial y}{\partial \eta} \end{pmatrix} = \boldsymbol{B_T}, \tag{5.1.6}$$

which is constant in space. With $v = 1$ in (5.1.5) and using the fact $\left|\widehat{\boldsymbol{T}}\right| = \frac{1}{2}$, we obtain

$$\int_{\boldsymbol{T}} \mathrm{d}\boldsymbol{x} = |\det(\boldsymbol{B_T})| \int_{\widehat{\boldsymbol{T}}} \mathrm{d}\widehat{\boldsymbol{x}} \iff |\det(J\mathcal{F}_{\boldsymbol{T}})| = |\det(\boldsymbol{B_T})| = \frac{|\boldsymbol{T}|}{\left|\widehat{\boldsymbol{T}}\right|} = 2\,|\boldsymbol{T}|\,. \tag{5.1.7}$$

This shows that $B_{\boldsymbol{T}}$ is non-singular for non-degenerate triangles ($|\boldsymbol{T}| \neq 0$). Combining (5.1.6) and (5.1.7) yields

$$|\det(J\mathcal{F}_{\boldsymbol{T}})| = 2|\boldsymbol{T}|. \tag{5.1.8}$$

Let us consider now an affine linear function $\phi\colon \boldsymbol{T} \longrightarrow \mathbb{R}$ which has the general form

$$\phi(x, y) = ax + by + c, \tag{5.1.9}$$

with $a, b, c \in \mathbb{R}$. In fact, the coefficients are uniquely determined by the values of $\phi$ at the vertices $\boldsymbol{x}_i$ of $\boldsymbol{T}$:

$$\phi(\boldsymbol{x}_i) = ax_i + by_i + c, \quad \text{for } i = 0, 1, 2. \tag{5.1.10}$$

**Lemma 5.1.2** (Norm Estimates for Linear Functions on Triangles and Edges)**.** *Let $\phi\colon \boldsymbol{T} \longrightarrow \mathbb{R}$ be a linear function (as described in* (5.1.9) *and* (5.1.10)*) on $\boldsymbol{T} := \operatorname{conv}\{\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2\}$. Using the notation $\phi_i := \phi(\boldsymbol{x}_i)$, for $i = 0, 1, 2$, and $E_{ij} := \overline{\boldsymbol{x}_i \boldsymbol{x}_j}$, for an edge between $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, $i \neq j$, we have*

$$(i) \quad \|\phi\|_{0,\boldsymbol{T}}^2 = \frac{|\boldsymbol{T}|}{6}\left(\phi_0^2 + \phi_1^2 + \phi_2^2 + \phi_0\phi_1 + \phi_0\phi_2 + \phi_1\phi_2\right), \tag{5.1.11a}$$

$$(ii) \quad \|\nabla\phi\|_{0,\boldsymbol{T}}^2 \leqslant c\left(\left(\phi_1 - \phi_0\right)^2 + \left(\phi_2 - \phi_0\right)^2\right), \tag{5.1.11b}$$

$$(iii) \quad \|\phi\|_{0,E_{ij}}^2 = \frac{h_{E_{ij}}}{3}\left(\phi_i^2 + \phi_i\phi_j + \phi_j^2\right), \tag{5.1.11c}$$

*where the constant $c > 0$ depends only on the mesh regularity constant given in* (4.1.2)*.*

*Proof:* (i): Applying the transform of integrals (5.1.5) together with (5.1.8) and (5.1.9), we can rewrite

$$\begin{aligned}
\frac{\|\phi\|_{0,\boldsymbol{T}}^2}{2|\boldsymbol{T}|} &= \frac{1}{2|\boldsymbol{T}|}\int_{\boldsymbol{T}} |\phi(\boldsymbol{x})|^2 \,\mathrm{d}\boldsymbol{x} = \frac{1}{2|\boldsymbol{T}|}\int_{\mathcal{F}_{\boldsymbol{T}}(\widehat{\boldsymbol{T}})} |\phi(x,y)|^2 \,\mathrm{d}(x,y) \\
&= \frac{1}{2|\boldsymbol{T}|}\int_{\widehat{\boldsymbol{T}}} \left|\phi\left(\mathcal{F}_{\boldsymbol{T}}(\xi,\eta)\right)\right|^2 \left|\det\left(J\mathcal{F}_{\boldsymbol{T}}(\xi,\eta)\right)\right| \mathrm{d}(\xi,\eta) \\
&= \int_{\widehat{\boldsymbol{T}}} \left|\phi\left(\xi x_1 + \eta x_2 + (1-\xi-\eta)\, x_0, \xi y_1 + \eta y_2 + (1-\xi-\eta)\, y_0\right)\right|^2 \mathrm{d}(\xi,\eta) \\
&= \int_{\widehat{\boldsymbol{T}}} \Big|a\left(\xi x_1 + \eta x_2 + (1-\xi-\eta)\, x_0\right) \\
&\qquad + b\left(\xi y_1 + \eta y_2 + (1-\xi-\eta)\, y_0\right) + c\Big|^2 \mathrm{d}(\xi,\eta) \\
&= \int_{\widehat{\boldsymbol{T}}} \left|\phi_1\xi + \phi_2\eta + \phi_0\left(1-\xi-\eta\right)\right|^2 \mathrm{d}(\xi,\eta) \\
&= \int_0^1 \int_0^{1-\xi} \left|\phi_1\xi + \phi_2\eta + \phi_0\left(1-\xi-\eta\right)\right|^2 \mathrm{d}\eta\mathrm{d}\xi \\
&= \frac{1}{12}\left(\phi_0^2 + \phi_1^2 + \phi_2^2 + \phi_0\phi_1 + \phi_0\phi_2 + \phi_1\phi_2\right).
\end{aligned}$$

(ii): The linearity of $\phi$ implies that its gradient is constant on $\boldsymbol{T}$. The transformation of derivatives according to the chain rule is given by

$$\begin{aligned}
\frac{\partial}{\partial\xi}\widehat{\phi}(\widehat{\boldsymbol{x}}) &= \frac{\partial}{\partial\xi}\phi\left(\mathcal{F}_{\boldsymbol{T}}(\xi,\eta)\right) = \frac{\partial}{\partial x}\phi(\boldsymbol{x})\frac{\partial x}{\partial\xi} + \frac{\partial}{\partial y}\phi(\boldsymbol{x})\frac{\partial y}{\partial\xi}, \\
\frac{\partial}{\partial\eta}\widehat{\phi}(\widehat{\boldsymbol{x}}) &= \frac{\partial}{\partial\eta}\phi\left(\mathcal{F}_{\boldsymbol{T}}(\xi,\eta)\right) = \frac{\partial}{\partial x}\phi(\boldsymbol{x})\frac{\partial x}{\partial\eta} + \frac{\partial}{\partial y}\phi(\boldsymbol{x})\frac{\partial y}{\partial\eta},
\end{aligned}$$

resulting in,

$$\widehat{\nabla}\widehat{\phi}(\widehat{\boldsymbol{x}}) = \widehat{\nabla}\phi\left(\boldsymbol{x}(\xi,\eta)\right) = \left(J\mathcal{F}_{\boldsymbol{T}}(\widehat{\boldsymbol{x}})\right)^T \nabla\phi(\boldsymbol{x}) = \left(\boldsymbol{B}_{\boldsymbol{T}}\right)^T \nabla\phi(\boldsymbol{x}), \tag{5.1.12}$$

where $\boldsymbol{x} := \mathcal{F}_{\boldsymbol{T}}(\xi,\eta)$ and $\widehat{\boldsymbol{x}} = (\xi,\eta)^T$. Further it holds

$$\left(\boldsymbol{B}_{\boldsymbol{T}}^T\right)^{-1} = \left(\det\left(\boldsymbol{B}_{\boldsymbol{T}}^{\mathrm{T}}\right)\right)^{-1}\left(\boldsymbol{B}_{\boldsymbol{T}}^T\right)^{adj} \xRightarrow{\det\left(\boldsymbol{B}_{\boldsymbol{T}}^{\mathrm{T}}\right)=\det\left(\boldsymbol{B}_{\boldsymbol{T}}\right)} \boldsymbol{B}_{\boldsymbol{T}}^{-T} = \frac{|\widehat{\boldsymbol{T}}|}{|\boldsymbol{T}|}\left(\boldsymbol{B}_{\boldsymbol{T}}^T\right)^{adj},$$

with $(\cdot)^{adj}$ denoting the adjunct. Using the reference map (5.1.4), the integral transform (5.1.5), the transformation of the gradient (5.1.12), Corollary 3.1.3, $(a-b)^2 \leqslant 2(a^2+b^2)$, for $a,b \in \mathbb{R}$, and $a^2+b^2 \leqslant (a+b)^2$, for $a,b \in \mathbb{R}_{\geqslant 0}$ we can estimate

$$
\begin{aligned}
\|\nabla\phi\|_{0,\boldsymbol{T}}^2 &= \int_{\boldsymbol{T}} |\nabla\phi(\boldsymbol{x})|^2 \,\mathrm{d}\boldsymbol{x} = \int_{\mathcal{F}_{\boldsymbol{T}}(\widehat{\boldsymbol{T}})} |\nabla\phi(x,y)|^2 \,\mathrm{d}(x,y) \\
&= \int_{\widehat{\boldsymbol{T}}} \left| \left(\boldsymbol{B}_{\boldsymbol{T}}^{-T}\right) \widehat{\nabla}\phi\left(\mathcal{F}_{\boldsymbol{T}}(\xi,\eta)\right) \right|^2 \left| \det\left(J\mathcal{F}_{\boldsymbol{T}}(\xi,\eta)\right) \right| \mathrm{d}(\xi,\eta) \\
&= |\det(\boldsymbol{B}_{\boldsymbol{T}})| \int_{\widehat{\boldsymbol{T}}} \left| \left(\boldsymbol{B}_{\boldsymbol{T}}^{-T}\right) \widehat{\nabla}\left(\phi_1\xi + \phi_2\eta + \phi_0(1-\xi-\eta)\right) \right|^2 \mathrm{d}(\xi,\eta) \\
&= \frac{|\boldsymbol{T}|}{|\widehat{\boldsymbol{T}}|} \int_{\widehat{\boldsymbol{T}}} \underbrace{\left| \left(\boldsymbol{B}_{\boldsymbol{T}}^{-T}\right) \begin{pmatrix} \phi_1-\phi_0 \\ \phi_2-\phi_0 \end{pmatrix} \right|^2}_{const.} \mathrm{d}(\xi,\eta) = |\boldsymbol{T}| \left| \left(\boldsymbol{B}_{\boldsymbol{T}}^{-T}\right) \begin{pmatrix} \phi_1-\phi_0 \\ \phi_2-\phi_0 \end{pmatrix} \right|^2 \\
&= |\boldsymbol{T}| \left( \frac{|\widehat{\boldsymbol{T}}|}{|\boldsymbol{T}|} \right)^2 \left| \left(\boldsymbol{B}_{\boldsymbol{T}}^{T}\right)^{adj} \begin{pmatrix} \phi_1-\phi_0 \\ \phi_2-\phi_0 \end{pmatrix} \right|^2 \\
&= \frac{1}{4|\boldsymbol{T}|} \left| \begin{pmatrix} y_2-y_0 & -(y_1-y_0) \\ -(x_2-x_0) & x_1-x_0 \end{pmatrix} \begin{pmatrix} \phi_1-\phi_0 \\ \phi_2-\phi_0 \end{pmatrix} \right|^2 \\
&\leqslant \frac{1}{2|\boldsymbol{T}|} \left( \underbrace{\|\boldsymbol{x}_2-\boldsymbol{x}_0\|^2}_{=|E_{02}|^2} (\phi_1-\phi_0)^2 + \underbrace{\|\boldsymbol{x}_1-\boldsymbol{x}_0\|^2}_{=|E_{01}|^2} (\phi_2-\phi_0)^2 \right) \\
&\leqslant \underbrace{\frac{\left(|E_{01}|+|E_{02}|\right)^2}{2|\boldsymbol{T}|}}_{=:c_{\boldsymbol{T}}} \left( (\phi_1-\phi_0)^2 + (\phi_2-\phi_0)^2 \right),
\end{aligned}
$$

where $|E_{ij}| := \|\boldsymbol{x}_i - \boldsymbol{x}_j\|$ denotes the length (Euclidean norm) of the edge $E_{ij}$. Since $|\boldsymbol{T}| \leqslant \frac{1}{2}h_{\boldsymbol{T}}^2$ and the radius of the largest inscribed circle in $\boldsymbol{T}$ is given by $\rho_{\boldsymbol{T}} = \frac{2|\boldsymbol{T}|}{|E_{01}|+|E_{02}|+|E_{12}|}$, we can use the shape-regularity assumption (4.1.2) in order to further estimate the coefficient as

$$
c_{\boldsymbol{T}} \leqslant \frac{\left(|E_{01}|+|E_{02}|+|E_{12}|\right)^2}{2|\boldsymbol{T}|} \leqslant \frac{\left(|E_{01}|+|E_{02}|+|E_{12}|\right)^2}{2|\boldsymbol{T}|} \frac{h_{\boldsymbol{T}}^2}{2|\boldsymbol{T}|} = \frac{h_{\boldsymbol{T}}^2}{\rho_{\boldsymbol{T}}^2} \leqslant C_{SR}^2,
$$

which gives the desired result with $c := C_{SR}^2$.
(iii): Given an edge $E_{ij} := \overline{\boldsymbol{x}_i\boldsymbol{x}_j}$ of the triangulation (with length $|E_{ij}| := \|\boldsymbol{x}_i - \boldsymbol{x}_j\|$), consider a linear mapping from the reference edge $\widehat{E} := [0,1]$ onto $E_{ij}$

$$
\begin{aligned}
\mathcal{F}_{E_{ij}} : [0,1] &\longrightarrow E_{ij}, \\
\xi &\mapsto \mathcal{F}_{E_{ij}}(\xi) := (\boldsymbol{x}_j - \boldsymbol{x}_i)\,\xi + \boldsymbol{x}_i.
\end{aligned}
$$

By the linearity of $\phi$ on $E_{ij}$, one can rewrite

$$\begin{aligned}\int_{E_{ij}} |\phi(\boldsymbol{x})|^2 \mathrm{d}s &= \int_{\mathcal{F}_{E_{ij}}(\widehat{E})} |\phi(\boldsymbol{x})|^2 \mathrm{d}s = \int_0^1 \left|\phi\left(\mathcal{F}_{E_{ij}}(\xi)\right)\right|^2 \left|\det\left(\mathcal{J}\mathcal{F}_{E_{ij}}(\xi)\right)\right| \mathrm{d}\xi \\ &= |E_{ij}| \int_0^1 \left|\phi(\boldsymbol{x}_i)(1-\xi) + \phi(\boldsymbol{x}_j)\,\xi\right|^2 \mathrm{d}\xi \\ &= |E_{ij}| \left(\frac{\phi_i^2}{3} + \frac{\phi_i\phi_j}{3} + \frac{\phi_j^2}{3}\right).\end{aligned}$$

Note that this is exactly what we get from Simpson's rule (5.1.3) with $v(\boldsymbol{x}) = \phi(\boldsymbol{x})^2$. □

## 5.2 The Discrete Formulation

The purpose of this section is to introduce the stabilized penalty-free non-symmetric Nitsche method for the Brinkman problem.

### 5.2.1 The Mesh and Related Notation

For a Lipschitz domain $\boldsymbol{\Omega} \in \mathbb{R}^2$, assumed to be in particular polygonal and connected, we consider a shape-regular (satisfying the condition (4.1.2)) family of boundary conforming[1], simplicial triangulations $\{\mathcal{T}_h\}_{h>0}$. Recalling the notation introduced in Section 4.1, an element $\boldsymbol{T} \in \mathcal{T}_h$ refers to a generic triangular mesh cell and $E$ denotes a generic edge.

Moreover, define the set of *corner nodes* of $\mathcal{T}_h$ by

$$\mathcal{C} := \left\{\boldsymbol{x} \in \Gamma \,:\, \exists\, E, E' \in \mathcal{G}_h \text{ such that } \boldsymbol{x} = E \cap E' \text{ and } \boldsymbol{n}_E \neq \boldsymbol{n}_{E'}\right\}, \quad (5.2.1)$$

where $\mathcal{G}_h$ is defined in (4.1.1). For a visualization of the notation, see Figure 5.2.

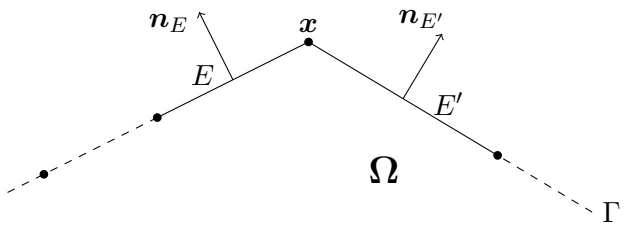



Figure 5.2: Sketch of a corner node $\boldsymbol{x} \in \mathcal{C}$.

Three further properties of the mesh will be important ingredients in the upcoming analysis:

[1] This is also termed *fitted mesh* and allows to exclude errors due to the discretization of the domain from the following analysis.

($\mathbf{M}_1$) Firstly we assume that the length ratio of neighboring boundary edges is bounded from above, i.e.,

$$\forall E, E' \in \mathcal{G}_h \text{ with } E \cap E' \neq \emptyset\colon \quad \frac{h_E}{h_{E'}} \leqslant \eta_0 < 7 + 4\sqrt{3}, \tag{5.2.2}$$

where the particular value of the bound will become clear in the analysis (proof of Lem. 5.3.10).

We observe that this assumption is weaker than quasi-uniformity of the boundary mesh, as it only restricts the ratio between the lengths of adjacent boundary edges. In 2D, this condition assures that the distance of neighboring vertices along the boundary mesh does not differ too much, see Figure 5.3.


Figure 5.3: Illustration of edge relations that respect (green dashed line) and violate (green dashed line extended into the red dashed line) the condition (5.2.2).

($\mathbf{M}_2$) Secondly, collecting the boundary elements, i.e., those mesh cells which have at least one vertex on the boundary, in

$$\boldsymbol{B}_h := \bigcup_{\substack{\boldsymbol{T}\in\mathcal{T}_h \\ \boldsymbol{T}\cap\Gamma\neq\emptyset}} \boldsymbol{T}, \tag{5.2.3}$$

we presuppose that there exists $\omega < 1$ such that

$$\forall h\colon \quad \frac{|\boldsymbol{B}_h|}{|\boldsymbol{\Omega}|} \leqslant \omega. \tag{5.2.4}$$

This means that the triangulation comprises elements which are strictly contained in the interior of the domain, independent of the mesh-size $h$, see Figure 5.4. This condition can always be satisfied by carrying out appropriate mesh refinements.

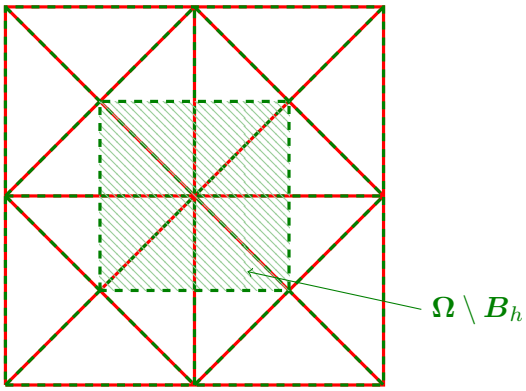



Figure 5.4: Illustration of a mesh that respects (green, dashed) and violates (red) the condition (5.2.4).

($\mathbf{M}_3$) Thirdly, we assume that there are no triangles whose vertices are all contained in the boundary.

If present, such a triangle can be split by introducing a vertex on the edge which is not contained in the boundary, resulting in two triangle splits (for the triangle itself and its neighbor).

Inspired by [BC09], we define an *auxiliary viscosity* via:

$$\hat{\mu} := \mu_{\text{eff}} + \sigma \ell_{\boldsymbol{\Omega}}^2. \tag{5.2.5}$$

To be precise, $\hat{\mu}$ is a linear combination of the effective viscosity and the fluid viscosity, where the latter is weighted by the dimensionless quantity $l_{\boldsymbol{\Omega}}^2 \mathbb{K}^{-1}$ (recall the definition $\sigma := \mu \, \mathbb{K}^{-1}$). Here, $l_{\boldsymbol{\Omega}} > 0$ represents a typical length scale of the problem and it is assumed to satisfy the condition

$$\ell_{\boldsymbol{\Omega}} \geqslant h_{\boldsymbol{T}}, \quad \forall \boldsymbol{T} \in \{\mathcal{T}_h\}_{h>0} \, . \tag{5.2.6}$$

Apart from the condition (5.2.6), the argumentation in the analysis will be independent from the choice of $\ell_{\boldsymbol{\Omega}}$.

An important property of the auxiliary viscosity in (5.2.5) is that $\hat{\mu} \neq 0$ as long as $\mu_{\text{eff}}, \sigma \geqslant 0$ and $\mu_{\text{eff}} + \sigma > 0$. Thus division by $\hat{\mu}$ is well-defined, even in the Stokes and Darcy limits.

Combining (5.2.5) and (5.2.6), the following inequalities hold:

$$\frac{\mu_{\text{eff}}}{\hat{\mu}} \leqslant 1 \qquad \text{and} \qquad \frac{\sigma h_{\boldsymbol{T}}^2}{\hat{\mu}} \leqslant \frac{\sigma \ell_{\boldsymbol{\Omega}}^2}{\hat{\mu}} \leqslant 1, \quad \forall \boldsymbol{T} \in \{\mathcal{T}_h\}_{h>0} \, . \tag{5.2.7}$$

Furthermore, we consider a dimensionally consistent version of the Sobolev norms taking into account the characteristic length:

$$\| \cdot \|_k^2 := \sum_{i=0}^{k} \ell_{\boldsymbol{\Omega}}^{2i} \, | \cdot |_i^2, \tag{5.2.8}$$

The definition (5.2.8) will be used throughout this chapter in place of the standard definitions (using the same notation).

### 5.2.2 A Robust Finite Element Method for the Brinkman Problem

In the following we will focus on the linear-linear finite element pair

$$\boldsymbol{V}_h^1 := \{\boldsymbol{v}_h \in \boldsymbol{H}^1(\boldsymbol{\Omega}) \cap \boldsymbol{C}^0\left(\overline{\boldsymbol{\Omega}}\right) \ : \ \boldsymbol{v}_h|_{\boldsymbol{T}} \in \mathbb{P}_1(\boldsymbol{T}), \quad \forall \boldsymbol{T} \in \mathcal{T}_h\}, \tag{5.2.9a}$$

$$Q_h^1 := \{q_h \in L_0^2(\boldsymbol{\Omega}) \cap C^0\left(\overline{\boldsymbol{\Omega}}\right) \ : \ q_h|_{\boldsymbol{T}} \in \mathbb{P}_1(\boldsymbol{T}), \quad \forall \boldsymbol{T} \in \mathcal{T}_h\} \tag{5.2.9b}$$

and assume, for simplicity of notation, that $\boldsymbol{f} \in \boldsymbol{L}^2(\boldsymbol{\Omega})$ and $g \in L^2(\boldsymbol{\Omega})$.

The **penalty-free non-symmetric Nitsche/nsGLS/grad-div formulation of the Brinkman problem** (3.2.8) is given by:
Find $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h^1 \times Q_h^1$ such that

$$\mathcal{A}_h\left[(\boldsymbol{u}_h, p_h); (\boldsymbol{v}_h, q_h)\right] = \mathcal{L}_h\left[(\boldsymbol{v}_h, q_h)\right], \quad \forall (\boldsymbol{v}_h, q_h) \in \boldsymbol{V}_h^1 \times Q_h^1 \tag{5.2.10a}$$

with

$$\begin{aligned}\mathcal{A}_h\left[(\boldsymbol{u}_h,p_h)\,;(\boldsymbol{v}_h,q_h)\right] := \mathcal{A}_h^{(+)}&\left[(\boldsymbol{u}_h,p_h)\,;(\boldsymbol{v}_h,q_h)\right]\\ &+\mathcal{S}_{h,\alpha}^{GLS,lhs}\left[(\boldsymbol{u}_h,p_h)\,;(\boldsymbol{v}_h,q_h)\right]\\ &-\left\langle(\mu_{\text{eff}}\nabla\boldsymbol{u}_h-p_h\mathbb{I})\cdot\boldsymbol{n},\boldsymbol{v}_h\right\rangle\\ &+\left\langle(\mu_{\text{eff}}\nabla\boldsymbol{v}_h-q_h\mathbb{I})\cdot\boldsymbol{n},\boldsymbol{u}_h\right\rangle\\ &+S_{h,\delta}^{GD,lhs}\left[(\boldsymbol{u}_h,\boldsymbol{v}_h)\right]+S_{h,\rho}^{C,lhs}\left[(\boldsymbol{u}_h,\boldsymbol{v}_h)\right],\end{aligned}\tag{5.2.10b}$$

and

$$\mathcal{L}_h\left[(\boldsymbol{v}_h,q_h)\right] := \mathcal{F}_h^{(+)}\left[(\boldsymbol{v}_h,q_h)\right]+\mathcal{S}_{h,\alpha}^{GLS,rhs}\left[(\boldsymbol{v}_h,q_h)\right]+S_{h,\delta}^{GD,rhs}\left[\boldsymbol{v}_h\right],\tag{5.2.10c}$$

where $\mathcal{A}_h^{(+)}\left[(\cdot,\cdot)\,;(\cdot,\cdot)\right]$ and $\mathcal{F}_h^{(+)}\left[(\cdot,\cdot)\right]$ are defined as in (4.2.1).[2] Further, there are terms belonging to the stabilizations, containing non-negative dimensionless stabilization parameters $\alpha$, $\delta$, $\rho \geqslant 0$, which are independent from the mesh size and constant in space. The expressions corresponding to the *non-symmetric GLS* method are defined as

$$\begin{aligned}S_{h,\alpha}^{GLS,lhs}\left[(\boldsymbol{u}_h,p_h)\,;(\boldsymbol{v}_h,q_h)\right] &:= \alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\left(\zeta_h\left(\boldsymbol{u}_h,p_h\right),\zeta_h\left(\boldsymbol{v}_h,q_h\right)\right)_{\boldsymbol{T}},\\ S_{h,\alpha}^{GLS,rhs}\left[(\boldsymbol{v}_h,q_h)\right] &:= \alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\left(\boldsymbol{f},\zeta_h\left(\boldsymbol{v}_h,q_h\right)\right)_{\boldsymbol{T}},\end{aligned}\tag{5.2.10d}$$

and

$$\zeta_h\left(\boldsymbol{v}_h,q_h\right) := -\mu_{\text{eff}}\Delta\boldsymbol{v}_h+\nabla q_h+\sigma\boldsymbol{v}_h.\tag{5.2.10e}$$

The *grad-div stabilization* terms are represented by

$$\begin{aligned}S_{h,\delta}^{GD,lhs}\left[(\boldsymbol{u}_h,\boldsymbol{v}_h)\right] &:= \delta\hat{\mu}\left(\nabla\cdot\boldsymbol{u}_h,\nabla\cdot\boldsymbol{v}_h\right),\\ S_{h,\delta}^{GD,rhs}\left[\boldsymbol{v}_h\right] &:= \delta\hat{\mu}\left(g,\nabla\cdot\boldsymbol{v}_h\right),\end{aligned}\tag{5.2.10f}$$

and the *corner stabilization*, which aims at stabilizing jumps of the normal velocity across corners $\boldsymbol{x}\in\mathcal{C}$ of the discrete boundary, is given by

$$S_{h,\rho}^{\mathcal{C},lhs}\left[(\boldsymbol{u}_h,\boldsymbol{v}_h)\right] := \rho\hat{\mu}\sum_{\boldsymbol{x}\in\mathcal{C}}\left[\!\left[\boldsymbol{u}_h\cdot\boldsymbol{n}\right]\!\right](\boldsymbol{x})\cdot\left[\!\left[\boldsymbol{v}_h\cdot\boldsymbol{n}\right]\!\right](\boldsymbol{x}).\tag{5.2.10g}$$

Here,

$$\begin{aligned}\left[\!\left[\boldsymbol{u}_h\cdot\boldsymbol{n}\right]\!\right](\boldsymbol{x}) &:= \boldsymbol{u}_h(\boldsymbol{x})\cdot\boldsymbol{n}_E-\boldsymbol{u}_h(\boldsymbol{x})\cdot\boldsymbol{n}_{E'}\\ &= \boldsymbol{u}_h(\boldsymbol{x})\cdot(\boldsymbol{n}_E-\boldsymbol{n}_{E'}) = \boldsymbol{u}_h(\boldsymbol{x})\cdot\left[\!\left[\boldsymbol{n}\right]\!\right]\end{aligned}\tag{5.2.10h}$$

denotes the jump of $\boldsymbol{u}\cdot\boldsymbol{n}$ at a corner node $\boldsymbol{x}\in\mathcal{C}$.

[2] The indices *lhs* and *rhs* indicate the affiliation with the left-hand side and right-hand side, respectively.

**Remark 5.2.1** (On Non-Homogeneous Essential Boundary Conditions)**.**
*In the case of non-homogeneous boundary conditions, the right-hand sides, consistent with the additional boundary term associated with the Nitsche method and with the corner stabilization* (5.2.10g)*, have to be taken into account. If* $\boldsymbol{u}|_\Gamma = \boldsymbol{u}_e$ *for* $\mu_{\text{eff}} > 0$*, respectively* $(\boldsymbol{u}\cdot\boldsymbol{n})|_\Gamma = \boldsymbol{u}_e\cdot\boldsymbol{n}$ *for* $\mu_{\text{eff}} = 0$*, then for consistency, the right-hand side* (5.2.10c) *has to be replaced by*

$$\mathcal{L}_h\left[(\boldsymbol{v}_h, q_h)\right] + \langle(\mu_{\text{eff}}\nabla\boldsymbol{v}_h - q_h\mathbb{I})\cdot\boldsymbol{n}, \boldsymbol{u}_e\rangle + S_{h,\rho}^{\mathcal{C},rhs}\left[\boldsymbol{v}_h\right],$$

*with* $S_{h,\rho}^{\mathcal{C},rhs}\left[\boldsymbol{v}_h\right] := \rho\hat{\mu}\sum_{\boldsymbol{x}\in\mathcal{C}}[\![\boldsymbol{u}_e\cdot\boldsymbol{n}]\!](\boldsymbol{x})\cdot[\![\boldsymbol{v}_h\cdot\boldsymbol{n}]\!](\boldsymbol{x})$.

**Remark 5.2.2** (On the Non-Symmetric GLS Method)**.**
*Note that since the velocity is element-wise approximated by a linear polynomial, it holds* $(\Delta\boldsymbol{u}_h)|_{\boldsymbol{T}} = \boldsymbol{0}$.

*For the special case* $\sigma = 0$ *(Stokes limit) we have recognized in 4.3.7, that the non-symmetric GLS method coincides with the symmetric GLS method and the PSPG method (for* $\boldsymbol{V}_h^1$*). However, there were no boundary terms present in the discrete formulation. Formally replacing* $q_h$ *by* $-q_h$ *in the method* (5.2.10a)*, has here also an influence on the Nitsche term* $\langle(\mu_{\text{eff}}\nabla\boldsymbol{v}_h - q_h\mathbb{I})\cdot\boldsymbol{n}, \boldsymbol{u}_h\rangle$*. For the test function* $-q_h$*, the mixed boundary integrals in* (5.2.10b) *become symmetric, while the boundary terms that couple velocities stay non-symmetric.*

*In presence of the Darcy term (*$\sigma \neq 0$*), the transformation* $q_h \mapsto -q_h$ *does not yield equivalence of the methods, even not in absence of the Nitsche terms.*

**Remark 5.2.3** (On the Grad-Div Stabilization)**.**
*The usage of a grad-div stabilization* (5.2.10f)*, originally proposed in [FH88] (see also, e.g., [Joh+17] for further detailed, more recent discussions), is motivated here by the desire to control the* $L^2$*-norm of the divergence of the velocity in the Darcy limit (see Subsection 4.2.2).*

*In addition, this term is also exploited in the following analysis (see Lem. 5.3.11 and Rem. 5.3.13) in order to obtain stability of the normal velocity on the boundary in the Darcy limit.*

**Remark 5.2.4** (On the Corner Stabilization)**.**
*The corner stabilization* (5.2.10g) *is an important ingredient of the proposed method, since it is used in order to prove robust stability, in particular in the Darcy limit (*$\mu_{\text{eff}} = 0$*), in absence of a Nitsche penalty term (of the form* $\gamma_N\sum_{E\in\mathcal{G}_h} h_E^{-1}\langle\boldsymbol{u}_h\cdot\boldsymbol{n}, \boldsymbol{v}_h\cdot\boldsymbol{n}\rangle_E$*). More precisely, control on the normal velocity at the boundary in the Darcy limit will be proven using the Nitsche term* $\langle q_h\boldsymbol{n}, \boldsymbol{u}_h\rangle$.

*The corner stabilization contains jump terms* (5.2.10h) *which are proportional to the jump of the normal vector at the boundary in corner points. In particular, for a continuous pressure space (together with a continuous velocity space), the control of the normal velocity at the boundary in corner points via* $\langle q_h\boldsymbol{n}, \boldsymbol{u}_h\rangle$ *is nontrivial (see Rem. 5.3.15).*

*Note that the order of the corner edges* $E$ *and* $E'$ *does not have an impact on the value of the corner stabilization (i.e., the sign of* (5.2.10h)*) since*

$$(\boldsymbol{a}\cdot(\boldsymbol{n}_E - \boldsymbol{n}_{E'}))(\boldsymbol{b}\cdot(\boldsymbol{n}_E - \boldsymbol{n}_{E'})) = (\boldsymbol{a}\cdot(\boldsymbol{n}_{E'} - \boldsymbol{n}_E))(\boldsymbol{b}\cdot(\boldsymbol{n}_{E'} - \boldsymbol{n}_E)).$$

**Remark 5.2.5** (On the Discrete Setting)**.**
*The non-symmetric GLS, the grad-div, as well as the corner stabilizations are motivated by our choice of the discrete spaces ($\boldsymbol{V}_h^1/Q_h^1$). It is worth noticing that the stability estimates which will be proven for the penalty-free non-symmetric Nitsche method and are based on the usage of linear finite elements for velocity and pressure, do not rely on the particular choice of the pressure stabilization, and therefore can be straightforwardly extended to other approaches (e.g., symmetric GLS or PSPG, see Rem. 4.3.6).*

### 5.2.3 The Norm for the Analysis

The discrete velocity space - although $\boldsymbol{V}_h^1 \subset \boldsymbol{H}^1(\boldsymbol{\Omega})$ - is not a subset of the weak velocity space, $\boldsymbol{V}_h^1 \not\subset \boldsymbol{V}$, neither in the Stokes limit nor in the Darcy limit (see Section 3.3). In fact, the weak imposition of essential boundary conditions results in non-conformity of the discrete velocity space which however only concerns the boundary. Therefore, a proper norm shall be considered that is able to explicitly control the velocity on the boundary.

We define the following mesh-dependent norm

$$\begin{aligned}|||(\boldsymbol{u},p)|||_h^2 := |||\boldsymbol{u}|||^2 &+ \sum_{E\in\mathcal{G}_h} \theta\,\frac{\mu_{\text{eff}}}{h_E}\|\boldsymbol{u}\|_{0,E}^2 + \sum_{E\in\mathcal{G}_h}\frac{\hat{\mu}}{h_E}\|\boldsymbol{u}\cdot\boldsymbol{n}_E\|_{0,E}^2 \\ &+ \rho\,\hat{\mu}\sum_{\boldsymbol{x}\in\mathcal{C}}|[\![\boldsymbol{u}\cdot\boldsymbol{n}]\!](\boldsymbol{x})|^2 + \frac{\|p\|_0^2}{\hat{\mu}} + \alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\|\nabla p\|_{0,\boldsymbol{T}}^2,\end{aligned} \tag{5.2.11a}$$

with

$$\theta := \frac{\mu_{\text{eff}}}{\hat{\mu}} \in [0,1] \tag{5.2.11b}$$

and

$$|||\boldsymbol{u}|||^2 := \mu_{\text{eff}}\|\nabla\boldsymbol{u}\|_0^2 + \sigma\|\boldsymbol{u}\|_0^2 + \delta\,\hat{\mu}\|\nabla\cdot\boldsymbol{u}\|_0^2. \tag{5.2.11c}$$

Later, we will refer to $|||(\cdot,\cdot)|||_h^2$ as the *energy norm*.

**Lemma 5.2.6.** *For $\mu_{\text{eff}} + \sigma > 0$, the expression* (5.2.11a) *defines a norm on $\boldsymbol{V}_h^1/Q_h^1$.*

*Proof:* The proof is valid for the more general case of any polynomial order $k \geqslant 1$, i.e., for $\boldsymbol{V}_h^k \times Q_h^k$. Since all summands in the norm (5.2.11a) are semi-norms, it remains to prove definiteness, i.e., for $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h^k \times Q_h^k$ it holds

$$|||(\boldsymbol{u}_h,p_h)|||_h^2 = 0 \quad \overset{!}{\Longrightarrow} \quad (\boldsymbol{u}_h,p_h) = (\boldsymbol{0},0)\,.$$

Let us assume that the norm is zero, then indeed all summands have to be zero due to their non-negativity. By the definition of $\boldsymbol{V}_h^k$ we have $\boldsymbol{u}_h|_{\boldsymbol{T}} \in \mathbb{P}_k(\boldsymbol{T})$.

If $\mu_{\text{eff}} \neq 0$, then

$$0 = \|\nabla\boldsymbol{u}_h\|_0^2 = \sum_{\boldsymbol{T}\in\mathcal{T}_h}\|\nabla\boldsymbol{u}_h\|_{0,\boldsymbol{T}}^2 = \sum_{\boldsymbol{T}\in\mathcal{T}_h}\int_{\boldsymbol{T}}|\nabla\boldsymbol{u}_h|^2 \implies (\nabla\boldsymbol{u}_h)|_{\boldsymbol{T}} = \boldsymbol{0}, \quad \forall\boldsymbol{T}\in\mathcal{T}_h,$$

which states that $\boldsymbol{u}_h|_{\boldsymbol{T}}$ is constant on each $\boldsymbol{T} \in \mathcal{T}_h$. Moreover,

$$0 = \sum_{E \in \mathcal{G}_h} \|\boldsymbol{u}_h\|_{0,E}^2 = \sum_{E \in \mathcal{G}_h} \int_E |\boldsymbol{u}_h|^2 \implies \boldsymbol{u}_h|_E = \boldsymbol{0}, \quad \forall E \in \mathcal{G}_h,$$

which, together with the continuity of $\boldsymbol{u}_h \in \boldsymbol{V}_h^k$ implies $\boldsymbol{u}_h = \boldsymbol{0}$.

In case of $\sigma \neq 0$, we have

$$0 = \|\boldsymbol{u}_h\|_0^2 = \sum_{\boldsymbol{T} \in \mathcal{T}_h} \|\boldsymbol{u}_h\|_{0,\boldsymbol{T}}^2 = \sum_{\boldsymbol{T} \in \mathcal{T}_h} \int_{\boldsymbol{T}} |\boldsymbol{u}_h|^2 \implies \boldsymbol{u}_h|_{\boldsymbol{T}} = \boldsymbol{0}, \quad \forall \boldsymbol{T} \in \mathcal{T}_h,$$

which implies $\boldsymbol{u}_h = \boldsymbol{0}$.

Similarly, for $\mu_{\text{eff}} + \sigma > 0$ we deduce $\|p\|_0^2 = 0 \implies p = 0$. □

As it will be shown in the stability analysis, the scaling by $\theta$ is important in order to obtain robust estimates (see Rem. 5.3.9), which are also valid in the Darcy limit ($\mu_{\text{eff}} = 0$). Note, that the terms scaled by $\theta$ vanish for $\mu_{\text{eff}} = 0$ (Darcy limit), and we obtain $\theta = 1$, if $\sigma = 0$ (Stokes limit).

**Remark 5.2.7** (On the Limit Cases of the Norm (5.2.11a))**.**
*In order to gain insight into the behavior of the weighted norm* (5.2.11a)*, let us discuss the limit cases.*
*(i) The Darcy limit: For* $\mu_{\text{eff}} = 0$*, the norm* (5.2.11a) *reduces to*

$$|||(\boldsymbol{u},p)|||_h^2 = \sigma \left( \|\boldsymbol{u}\|_0^2 + \ell_{\boldsymbol{\Omega}}^2 \left( \delta \|\nabla \cdot \boldsymbol{u}\|_0^2 + \sum_{E \in \mathcal{G}_h} \frac{1}{h_E} \|\boldsymbol{u} \cdot \boldsymbol{n}_E\|_{0,E}^2 \right.\right.$$
$$\left.\left. + \rho \sum_{\boldsymbol{x} \in \mathcal{C}} |[\![\boldsymbol{u} \cdot \boldsymbol{n}]\!](\boldsymbol{x})|^2 \right) \right) + \frac{1}{\sigma \ell_{\boldsymbol{\Omega}}^2} \left( \|p\|_0^2 + \alpha \sum_{\boldsymbol{T} \in \mathcal{T}_h} h_{\boldsymbol{T}}^2 \|\nabla p\|_{0,\boldsymbol{T}}^2 \right).$$

*Hence, control of the gradient of the velocity and control of the tangential velocity at the boundary are lost which is in line with* $\boldsymbol{u} \in \boldsymbol{H}^{div}$*.*

*Control on the divergence of the velocity is provided by the grad-div stabilization,* $\delta \neq 0$*.*

*We observe that the norms of the pressure scale with* $\sigma^{-1}$ *and the norms of the velocity are weighted with* $\sigma$*, i.e., they are scaled inversely compared to each other. Except for the* $\boldsymbol{L}^2$ *norm of the velocity, a similar behavior is given with respect to* $\ell_{\boldsymbol{\Omega}}^2$*.*
*(ii) The Stokes limit: For* $\sigma = 0$*, the corresponding norm* (5.2.11a) *reduces to*

$$|||(\boldsymbol{u},p)|||_h^2 = \mu_{\text{eff}} \left( \|\nabla \boldsymbol{u}\|_0^2 + \delta \|\nabla \cdot \boldsymbol{u}\|_0^2 + \sum_{E \in \mathcal{G}_h} \frac{1}{h_E} \left( \|\boldsymbol{u}\|_{0,E}^2 + \|\boldsymbol{u} \cdot \boldsymbol{n}_E\|_{0,E}^2 \right) \right.$$
$$\left. + \rho \sum_{\boldsymbol{x} \in \mathcal{C}} |[\![\boldsymbol{u} \cdot \boldsymbol{n}]\!](\boldsymbol{x})|^2 \right) + \frac{1}{\mu_{\text{eff}}} \left( \|p\|_0^2 + \alpha \sum_{\boldsymbol{T} \in \mathcal{T}_h} h_{\boldsymbol{T}}^2 \|\nabla p\|_{0,\boldsymbol{T}}^2 \right).$$

*Also in this case, the parts of the norm corresponding to the velocity and the parts corresponding to the pressure are scaled inversely by the physical coefficient* $\mu_{\text{eff}}$*. Note that the length scale* $\ell_{\boldsymbol{\Omega}}$ *has disappeared.*

## 5.3 Stability and Convergence

This section is dedicated to the stability and convergence of the proposed stabilized penalty-free non-symmetric Nitsche method (5.2.10). Before discussing the technical proofs in detail, let us state the main results that will be shown.

**(1) Inf-sup stability** for $\alpha, \delta, \rho > 0$, and $\mu_{\text{eff}}, \sigma \geqslant 0$ with $\mu_{\text{eff}} + \sigma > 0$, i.e., the existence of a constant $\beta_h > 0$, independent of $h$ and of the physical parameters, such that

$$\inf_{(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h^1 \times Q_h^1 \setminus \{(\mathbf{0},0)\}} \left( \sup_{(\boldsymbol{v}_h, q_h) \in \boldsymbol{V}_h^1 \times Q_h^1 \setminus \{(\mathbf{0},0)\}} \left( \frac{\mathcal{A}_h \left[ (\boldsymbol{u}_h, p_h) ; (\boldsymbol{v}_h, q_h) \right]}{|||(\boldsymbol{u}_h, p_h)|||_h |||(\boldsymbol{v}_h, q_h)|||_h} \right) \right) \geqslant \beta_h.$$

The inf-sup constant $\beta_h$ depends on the stabilization parameters and on the shape regularity of the mesh.

This statement assesses unconditional stability with respect to the physical parameters, including the limit cases $\mu_{\text{eff}} = 0$ or $\sigma = 0$. In particular, for small values of the stabilization parameters, it holds $\beta_h^{-1} = O\left(\alpha^{-1}\left(\alpha^{-1}\left(\rho + \delta\right)^{-1}\right)\right)$.

**(2) A priori error estimates** for $\alpha, \delta, \rho > 0$ and $\mu_{\text{eff}}, \sigma \geqslant 0$ with $\mu_{\text{eff}} + \sigma > 0$, i.e.,

$$|||(\boldsymbol{u} - \boldsymbol{u}_h, p - p_h)|||_h \leqslant h \left( C_{\boldsymbol{u}} \left| \boldsymbol{u} \right|_2 + C_p \left| p \right|_1 \right),$$

for $(\boldsymbol{u}, p) \in \boldsymbol{H}^2(\boldsymbol{\Omega}) \times H^1(\boldsymbol{\Omega})$ solving (3.2.8) and $(\boldsymbol{u}_h, p_h)$ solving (5.2.10). The constants $C_{\boldsymbol{u}}$ and $C_p$ are independent of $h$ and for small respectively moderate values of stabilization parameters, it holds

$$C_{\boldsymbol{u}} = O\left( \frac{\hat{\mu}^{\frac{1}{2}}}{\beta_h} \right), \qquad C_p = O\left( \frac{1}{\hat{\mu}^{\frac{1}{2}} \delta^{\frac{1}{2}} \beta_h} \right).$$

This a priori estimate is optimal with respect to the chosen norm.

**Remark 5.3.1** (On (1) and (2))**.**
*The above results imply that the penalty-free Nitsche formulation possesses a convergence and stability behavior that is comparable to that of the classical formulation (strong imposition of essential boundary conditions, see, e.g., [BC09]) and to that of the classical Nitsche method (where a penalty is present, see, e.g., [HJ09] and [JS10]).*

*The scaling of the stabilization terms with the auxiliary viscosity $\hat{\mu}$ defined in* (5.2.5) *is an important requirement in order to obtain the robust convergence estimate. Alternative scalings were analyzed, e.g., in [BC09], for the Brinkman problem with strong imposition of essential boundary conditions. There, it was shown that stability and optimal error estimates can also be obtained by scaling the stabilization of the Darcy terms with respect to the mesh, replacing $\hat{\mu}$ by $\hat{\mu}_{\boldsymbol{T}} := \mu_{\text{eff}} + \sigma h_{\boldsymbol{T}}^2$ on each triangle $\boldsymbol{T} \in \mathcal{T}_h$. A similar scaling has been analyzed in [JS10] in the context of a rescaled Brinkman problem with the classical Nitsche*

*method (limited to the case $\sigma > 0$). However, as it will be shown in the next subsections, the scaling with* (5.2.5) *is crucial for the given proof of uniform control on velocity at the boundary for $\mu_{\text{eff}}, \sigma \geqslant 0$ (see also Rem. 5.3.14).*

### 5.3.1 Well-Posedness

The goal in this subsection is to show that the problem (5.2.10) has exactly one solution $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h^1 \times Q_h^1$. As discussed in Section 4.1, the condition we have to prove is the inf-sup stability of the discrete bilinear form (5.2.10b) with respect to the mesh-dependent norm (5.2.11a). In fact, we will prove unconditional stability with respect to the physical parameters, including the limit cases $\mu_{\text{eff}} = 0$ or $\sigma = 0$.

The proof will pursue the following strategy: Find a pair of test functions $(\boldsymbol{v}_h^\star, q_h^\star) \in \boldsymbol{V}_h^1 \times Q_h^1 \setminus \{(\boldsymbol{0}, 0)\}$ such that there exist two constants $\beta_{h,1} > 0$ and $\beta_{h,2} > 0$ satisfying the two inequalities:

$$\begin{aligned} (I) \quad & \mathcal{A}_h\left[(\boldsymbol{u}_h, p_h)\,;(\boldsymbol{v}_h^\star, q_h^\star)\right] \geqslant \beta_{h,1} |\!|\!| (\boldsymbol{u}_h, p_h) |\!|\!|_h^2, \\ (II) \quad & |\!|\!| (\boldsymbol{u}_h, p_h) |\!|\!|_h \geqslant \beta_{h,2} |\!|\!| (\boldsymbol{v}_h^\star, q_h^\star) |\!|\!|_h. \end{aligned}$$

Combining the two estimates $(I)$ and $(II)$ we get

$$\begin{aligned} \sup_{(\boldsymbol{v}_h, q_h) \in \boldsymbol{V}_h^1 \times Q_h^1 \setminus \{(\boldsymbol{0},0)\}} \frac{\mathcal{A}_h\left[(\boldsymbol{u}_h, p_h)\,;(\boldsymbol{v}_h, q_h)\right]}{|\!|\!| (\boldsymbol{v}_h, q_h) |\!|\!|_h} &\geqslant \frac{\mathcal{A}_h\left[(\boldsymbol{u}_h, p_h)\,;(\boldsymbol{v}_h^\star, q_h^\star)\right]}{|\!|\!| (\boldsymbol{v}_h^\star, q_h^\star) |\!|\!|_h} \\ &\geqslant \beta_h\, |\!|\!| (\boldsymbol{u}_h, p_h) |\!|\!|_h, \end{aligned}$$

for all $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h^1 \times Q_h^1 \setminus \{(\boldsymbol{0}, 0)\}$, which implies the inf-sup condition with $\beta_h := \beta_{h,1}\beta_{h,2}$.

The construction of an appropriate test function pair satisfying conditions $(I)$ and $(II)$ will be split into four steps, each involving its own pair of test functions. Control of the bulk norms will be shown in Parts I and II, whereas control of the boundary norms in Parts III and IV. Subsequently, the four test pairs will be linearly combined such that the result additionally satisfies condition $(II)$.

Throughout the proofs, the introduced constants depending on the physical parameters or on the stabilization parameters will be discussed in detail, in order to allow the reader to follow the derivation and, eventually, to clearly assess the role of the physical parameters within the derived estimates (especially in the limit cases).

#### Bulk Control

The first result concerns the coercivity of the bilinear form (5.2.10b) in a norm which is weaker than (5.2.11a) and yields control on all symmetric parts of the bilinear form $\mathcal{A}_h\left[(\cdot,\cdot)\,,(\cdot,\cdot)\right]$.

**Lemma 5.3.2** (Part I – Coercivity in a Weaker Norm)**.**
*Let $\alpha, \delta, \rho \geqslant 0$ and $\mu_{\text{eff}}, \sigma \geqslant 0$ with $\mu_{\text{eff}} + \sigma > 0$. Then there exists a constant $C_0 = C_0(\alpha) > 0$, independent of the physical parameters and $h$, such that*

$$\mathcal{A}_h\left[(\boldsymbol{v}_h, q_h)\,;(\boldsymbol{v}_h, q_h)\right] \geqslant C_0 \left( |||\boldsymbol{v}_h|||^2 + \rho\hat{\mu} \sum_{\boldsymbol{x}\in\mathcal{C}} |[\![\boldsymbol{v}_h \cdot \boldsymbol{n}]\!](\boldsymbol{x})|^2 + \alpha \sum_{\boldsymbol{T}\in\mathcal{T}_h} \frac{h_{\boldsymbol{T}}^2}{\hat{\mu}} \|\nabla q_h\|_{0,\boldsymbol{T}}^2 \right),$$

*for all $(\boldsymbol{v}_h, q_h) \in \boldsymbol{V}_h^1 \times Q_h^1$.*

*Proof:* Let $(\boldsymbol{v}_h, q_h) \in \boldsymbol{V}_h^1 \times Q_h^1$, then it is

$$\begin{aligned}
\mathcal{A}_h\left[(\boldsymbol{v}_h, q_h)\,;(\boldsymbol{v}_h, q_h)\right] = \mu_{\text{eff}}\|\nabla \boldsymbol{v}_h\|_0^2 + \sigma\|\boldsymbol{v}_h\|_0^2 + \alpha \sum_{\boldsymbol{T}\in\mathcal{T}_h} \frac{h_{\boldsymbol{T}}^2}{\hat{\mu}} \|\nabla q_h\|_{0,\boldsymbol{T}}^2 \\
+ \alpha\sigma^2 \sum_{\boldsymbol{T}\in\mathcal{T}_h} \frac{h_{\boldsymbol{T}}^2}{\hat{\mu}} \|\boldsymbol{v}_h\|_{0,\boldsymbol{T}}^2 + 2\alpha\sigma \sum_{\boldsymbol{T}\in\mathcal{T}_h} \frac{h_{\boldsymbol{T}}^2}{\hat{\mu}} (\nabla q_h, \boldsymbol{v}_h)_{\boldsymbol{T}} \\
+ \delta\,\hat{\mu}\|\nabla \cdot \boldsymbol{v}_h\|_0^2 + \rho\,\hat{\mu} \sum_{\boldsymbol{x}\in\mathcal{C}} |[\![\boldsymbol{v}_h \cdot \boldsymbol{n}]\!](\boldsymbol{x})|^2 ,
\end{aligned}$$

where the boundary integrals and the terms corresponding to the bilinear form $b_h(\cdot,\cdot)$ (defined in (4.2.1)) have canceled. If $\alpha = 0$, we are already done with any $C_0 \leqslant 1$. For $\alpha > 0$ we can use the Cauchy–Schwarz and Young inequalities (Theorems 3.1.5 and 3.1.1) to obtain

$$2\alpha\sigma \sum_{\boldsymbol{T}\in\mathcal{T}_h} \frac{h_{\boldsymbol{T}}^2}{\hat{\mu}} (\nabla q_h, \boldsymbol{v}_h)_{\boldsymbol{T}} \geqslant -\varepsilon\alpha \sum_{\boldsymbol{T}\in\mathcal{T}_h} \frac{h_{\boldsymbol{T}}^2}{\hat{\mu}} \|\nabla q_h\|_{0,\boldsymbol{T}}^2 - \sigma\frac{\alpha}{\varepsilon} \sum_{\boldsymbol{T}\in\mathcal{T}_h} \frac{\sigma h_{\boldsymbol{T}}^2}{\hat{\mu}} \|\boldsymbol{v}_h\|_{0,\boldsymbol{T}}^2$$

and further

$$\begin{aligned}
\mathcal{A}_h\left[(\boldsymbol{v}_h, q_h)\,;(\boldsymbol{v}_h, q_h)\right] \geqslant \mu_{\text{eff}}\|\nabla \boldsymbol{v}_h\|_0^2 + \sigma\|\boldsymbol{v}_h\|_0^2 + \delta\,\hat{\mu}\|\nabla \cdot \boldsymbol{v}_h\|_0^2 \\
+ \rho\,\hat{\mu} \sum_{\boldsymbol{x}\in\mathcal{C}} |[\![\boldsymbol{v}_h \cdot \boldsymbol{n}]\!](\boldsymbol{x})|^2 + \sigma\left(\alpha - \frac{\alpha}{\varepsilon}\right) \sum_{\boldsymbol{T}\in\mathcal{T}_h} \frac{\sigma h_{\boldsymbol{T}}^2}{\hat{\mu}} \|\boldsymbol{v}_h\|_{0,\boldsymbol{T}}^2 \\
+ (1-\varepsilon)\,\alpha \sum_{\boldsymbol{T}\in\mathcal{T}_h} \frac{h_{\boldsymbol{T}}^2}{\hat{\mu}} \|\nabla q_h\|_{0,\boldsymbol{T}}^2,
\end{aligned}$$

for any $\varepsilon > 0$. The coefficient in front of the last term, $(1-\varepsilon)$, is strictly positive for any $\varepsilon < 1$, which implies $\alpha - \frac{\alpha}{\varepsilon} < 0$, such that with (5.2.7) we can estimate

$$\begin{aligned}
\mathcal{A}_h\left[(\boldsymbol{v}_h, q_h)\,;(\boldsymbol{v}_h, q_h)\right] \geqslant \mu_{\text{eff}}\,\|\nabla \boldsymbol{v}_h\|_0^2 + \left(1 + \alpha - \frac{\alpha}{\varepsilon}\right)\sigma\|\boldsymbol{v}_h\|_0^2 + \delta\,\hat{\mu}\|\nabla \cdot \boldsymbol{v}_h\|_0^2 \\
+ \rho\,\hat{\mu} \sum_{\boldsymbol{x}\in\mathcal{C}} |[\![\boldsymbol{v}_h \cdot \boldsymbol{n}]\!](\boldsymbol{x})|^2 + (1-\varepsilon)\,\alpha \sum_{\boldsymbol{T}\in\mathcal{T}_h} \frac{h_{\boldsymbol{T}}^2}{\hat{\mu}} \|\nabla q_h\|_{0,\boldsymbol{T}}^2.
\end{aligned}$$

In order to guarantee strict positivity of the coefficient in front of the $L^2$-norm of the velocity, namely $\left(1 + \alpha - \frac{\alpha}{\varepsilon}\right)$, we further have to choose $\varepsilon$ such that $\frac{\alpha}{\alpha+1} < \varepsilon$ is satisfied. Taking $\varepsilon^\star := \sqrt{\frac{\alpha^2}{4} + \alpha} - \frac{\alpha}{2}$ is an appropriate choice (since $\alpha > 0$),

which moreover implicates that the two $\varepsilon$-dependent coefficients coincide. Thus, we obtain

$$\mathcal{A}_h\left[(\boldsymbol{v}_h, q_h)\,;(\boldsymbol{v}_h, q_h)\right] \geqslant \mu_{\text{eff}}\|\nabla \boldsymbol{v}_h\|_0^2 + \delta\,\hat{\mu}\|\nabla\cdot\boldsymbol{v}_h\|_0^2 + \rho\,\hat{\mu}\sum_{\boldsymbol{x}\in\mathcal{C}}\left|[\![\boldsymbol{v}_h\cdot\boldsymbol{n}]\!](\boldsymbol{x})\right|^2 + (1-\varepsilon^\star)\left(\sigma\|\boldsymbol{v}_h\|_0^2 + \alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\|\nabla q_h\|_{0,\boldsymbol{T}}^2\right).$$

The proof is concluded defining

$$C_0 := 1-\varepsilon^\star = 1 - \sqrt{\frac{\alpha^2}{4}+\alpha} + \frac{\alpha}{2} \tag{5.3.1}$$

and using the fact that $C_0 \leqslant 1$. □

**Remark 5.3.3** (On the Behavior of $C_0$)**.**
*Notice that the constant $C_0 = C_0(\alpha)$ introduced in* (5.3.1) *is a decreasing function of $\alpha$ satisfying $C_0(0) = 1$ and $C_0(\alpha) \in (0,1)$ for $\alpha > 0$. In particular, $C_0 = O(1)$ for small and moderate values of $\alpha$. Note that the estimate in Lemma 5.3.2 holds also for $\alpha = 0$, i.e., without GLS stabilization.*

The following lemma provides stability of the pressure in the $L^2$-norm.

**Lemma 5.3.4** (Part II – Pressure Control)**.**
*Let $\alpha > 0$, $\delta, \rho \geqslant 0$, and $\mu_{\text{eff}}, \sigma \geqslant 0$ with $\mu_{\text{eff}} + \sigma > 0$. Then, there exists a constant $C_1 = C_1(\alpha, \delta) > 0$, independent of the physical parameters and $h$, such that, for all $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h^1 \times Q_h^1$, we can find a function $\boldsymbol{v}_h \in \boldsymbol{V}_h^1$ that satisfies*

$$\mathcal{A}_h\left[(\boldsymbol{u}_h, p_h)\,;(\boldsymbol{v}_h, 0)\right] \geqslant \frac{1}{2}\frac{\|p_h\|_0^2}{\hat{\mu}} - C_1\left(|||\boldsymbol{u}_h|||^2 + \alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\|\nabla p_h\|_{0,\boldsymbol{T}}^2 + \sum_{E\in\mathcal{G}_h}\theta\,\frac{\mu_{\text{eff}}}{h_E}\|\boldsymbol{u}_h\|_{0,E}^2\right).$$

*Proof:* Let $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h^1 \times Q_h^1$. Since $p_h \in Q_h^1 \subset L_0^2(\boldsymbol{\Omega})$ (due to conformity), according to Theorem 3.1.10 there exists a function $\boldsymbol{v}_{p_h} \in \boldsymbol{H}_0^1(\boldsymbol{\Omega})$ and a dimensionless constant $\hat{c}_{\boldsymbol{\Omega}}$ (that only depends on $\boldsymbol{\Omega}$) such that

$$\nabla\cdot\boldsymbol{v}_{p_h} = -\frac{1}{\hat{\mu}}\,p_h, \tag{5.3.2a}$$

$$\|\nabla\boldsymbol{v}_{p_h}\|_0 \leqslant \frac{\hat{c}_{\boldsymbol{\Omega}}}{\hat{\mu}}\|p_h\|_0. \tag{5.3.2b}$$

Let now

$$\boldsymbol{v}_h := I_{h,1}^{\text{SZ}}(\boldsymbol{v}_{p_h}) \in \boldsymbol{V}_h^1 \tag{5.3.3}$$

be the Scott–Zhang interpolation of the function $\boldsymbol{v}_{p_h}$. Due to the $H^1$-stability of the Scott–Zhang interpolation operator (5.1.2a) (with $l = 1$), (5.2.8), the Poincaré inequality (Thm. 3.1.8 with the Poncaré constant decomposed into

a non-dimensional and a dimensional part, i.e. $C_P = \overline{C}_P \ell_{\boldsymbol{\Omega}}$), and property (5.3.2b), there also holds

$$\|\nabla \boldsymbol{v}_h\|_0 \leqslant \frac{c_{\boldsymbol{\Omega}}}{\hat{\mu}} \|p_h\|_0 \qquad \text{and} \qquad \|\boldsymbol{v}_h\|_0 \leqslant \frac{\ell_{\boldsymbol{\Omega}} c_{\boldsymbol{\Omega}}}{\hat{\mu}} \|p_h\|_0, \tag{5.3.4}$$

with a dimensionless constant $c_{\boldsymbol{\Omega}} := c_{\mathrm{SZ}} \left( \overline{C}_P^2 + 1 \right)^{\frac{1}{2}} \hat{c}_{\boldsymbol{\Omega}}$ that only depends on the domain and on the (shape-)regularity of the mesh. Moreover, according to (5.1.2b) (with $l = 1$ and $m = 0$), it holds

$$\sum_{\boldsymbol{T} \in \mathcal{T}_h} \frac{1}{h_{\boldsymbol{T}}^2} \|\boldsymbol{v}_{p_h} - \boldsymbol{v}_h\|_{0,\boldsymbol{T}}^2 \leqslant \sum_{\boldsymbol{T} \in \mathcal{T}_h} \frac{1}{h_{\boldsymbol{T}}^2} \overline{c}_{\mathrm{SZ}}^2 h_{\boldsymbol{T}}^2 \|\nabla \boldsymbol{v}_{p_h}\|_{0,S(\boldsymbol{T})}^2 \leqslant \widetilde{c_{\mathrm{SZ}}}^2 \|\nabla \boldsymbol{v}_{p_h}\|_0^2, \tag{5.3.5}$$

with $\widetilde{c_{\mathrm{SZ}}}^2 := \overline{c}_{\mathrm{SZ}}^2 \left( \max_{\boldsymbol{T} \in \mathcal{T}_h} \{ \# S(\boldsymbol{T}) \} \right)$. Here, $\# S(\boldsymbol{T})$ denotes the number of triangles contained in $S(\boldsymbol{T})$ which depends on the regularity of the mesh. Since the Scott–Zhang interpolation operator preserves essential boundary conditions, it holds $\boldsymbol{v}_h \in \boldsymbol{H}_0^1(\boldsymbol{\Omega}) \cap \boldsymbol{V}_h^1$ such that the boundary terms involving $\boldsymbol{v}_h$ vanish. Using the decomposition $\boldsymbol{v}_h = \boldsymbol{v}_{p_h} - (\boldsymbol{v}_{p_h} - \boldsymbol{v}_h)$ and integration by parts for the term involving $(\boldsymbol{v}_{p_h} - \boldsymbol{v}_h) \in \boldsymbol{H}_0^1(\boldsymbol{\Omega})$ we get

$$\begin{aligned}
\mathcal{A}_h \left[ (\boldsymbol{u}_h, p_h) ; (\boldsymbol{v}_h, 0) \right] &= \mu_{\mathrm{eff}} \left( \nabla \boldsymbol{u}_h, \nabla \boldsymbol{v}_h \right) + \sigma \left( \boldsymbol{u}_h, \boldsymbol{v}_h \right) - \left( p_h, \nabla \cdot \boldsymbol{v}_h \right) \\
&\quad + \alpha \sum_{\boldsymbol{T} \in \mathcal{T}_h} \frac{h_{\boldsymbol{T}}^2}{\hat{\mu}} \left( \sigma \boldsymbol{u}_h + \nabla p_h, \sigma \boldsymbol{v}_h \right)_{\boldsymbol{T}} + \delta \, \hat{\mu} \left( \nabla \cdot \boldsymbol{u}_h, \nabla \cdot \boldsymbol{v}_h \right) \\
&\quad + \left\langle \mu_{\mathrm{eff}} \nabla \boldsymbol{v}_h \cdot \boldsymbol{n}, \boldsymbol{u}_h \right\rangle \\
&= \mu_{\mathrm{eff}} \left( \nabla \boldsymbol{u}_h, \nabla \boldsymbol{v}_h \right) + \sigma \left( \boldsymbol{u}_h, \boldsymbol{v}_h \right) - \left( p_h, \nabla \cdot \boldsymbol{v}_{p_h} \right) \\
&\quad - \left( \nabla p_h, \boldsymbol{v}_{p_h} - \boldsymbol{v}_h \right) + \alpha \sum_{\boldsymbol{T} \in \mathcal{T}_h} \frac{h_{\boldsymbol{T}}^2}{\hat{\mu}} \left( \sigma \boldsymbol{u}_h + \nabla p_h, \sigma \boldsymbol{v}_h \right)_{\boldsymbol{T}} \\
&\quad + \delta \, \hat{\mu} \left( \nabla \cdot \boldsymbol{u}_h, \nabla \cdot \boldsymbol{v}_h \right) + \left\langle \mu_{\mathrm{eff}} \nabla \boldsymbol{v}_h \cdot \boldsymbol{n}, \boldsymbol{u}_h \right\rangle .
\end{aligned}$$

Using the Cauchy–Schwarz inequality (Thm. 3.1.5), the equality (5.3.2a), the Cauchy–Schwarz inequality for sums (Lem. 3.1.2), the inequality (5.3.5), and Corollary 3.1.4, we obtain

$$\begin{aligned}
\mathcal{A}_h \left[ (\boldsymbol{u}_h, p_h) ; (\boldsymbol{v}_h, 0) \right] &\geqslant -\mu_{\mathrm{eff}}^{\frac{1}{2}} \left( \mu_{\mathrm{eff}}^{\frac{1}{2}} \|\nabla \boldsymbol{u}_h\|_0 \right) \|\nabla \boldsymbol{v}_h\|_0 - \sigma^{\frac{1}{2}} \left( \sigma^{\frac{1}{2}} \|\boldsymbol{u}_h\|_0 \right) \|\boldsymbol{v}_h\|_0 \\
&\quad + \frac{\|p_h\|_0^2}{\hat{\mu}} - \widetilde{c_{\mathrm{SZ}}} \left( \sum_{\boldsymbol{T} \in \mathcal{T}_h} h_{\boldsymbol{T}}^2 \|\nabla p_h\|_{0,\boldsymbol{T}}^2 \right)^{\frac{1}{2}} \|\nabla \boldsymbol{v}_{p_h}\|_0 \\
&\quad \underbrace{- \alpha \sigma \sum_{\boldsymbol{T} \in \mathcal{T}_h} \frac{\sigma h_{\boldsymbol{T}}^2}{\hat{\mu}} \|\boldsymbol{u}_h\|_{0,\boldsymbol{T}} \|\boldsymbol{v}_h\|_{0,\boldsymbol{T}}}_{=: \mathcal{T}_1} \\
&\quad \underbrace{- \alpha \sum_{\boldsymbol{T} \in \mathcal{T}_h} \frac{\sigma h_{\boldsymbol{T}}^2}{\hat{\mu}} \|\nabla p_h\|_{0,\boldsymbol{T}} \|\boldsymbol{v}_h\|_{0,\boldsymbol{T}}}_{=: \mathcal{T}_2} \\
&\quad - \delta \, \hat{\mu} \|\nabla \cdot \boldsymbol{u}_h\|_0 N^{\frac{1}{2}} \|\nabla \boldsymbol{v}_h\|_0 + \left\langle \mu_{\mathrm{eff}} \nabla \boldsymbol{v}_h \cdot \boldsymbol{n}, \boldsymbol{u}_h \right\rangle .
\end{aligned} \tag{5.3.6}$$

The terms $\mathcal{T}_1$ and $\mathcal{T}_2$, introduced above, can be estimated using (5.2.7), the Cauchy–Schwarz inequality for sums (Lem. 3.1.2), and the second inequality of (5.3.4), yielding

$$\mathcal{T}_1 \leqslant \alpha\sigma^{\frac{1}{2}}\|\boldsymbol{u}_h\|_0\,\sigma^{\frac{1}{2}}\|\boldsymbol{v}_h\|_0 \leqslant c_{\boldsymbol{\Omega}}\,\alpha\sigma^{\frac{1}{2}}\|\boldsymbol{u}_h\|_0\frac{\|p_h\|_0}{\hat{\mu}^{\frac{1}{2}}} \tag{5.3.7}$$

and

$$\begin{aligned}\mathcal{T}_2 &= \alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\sigma^{\frac{1}{2}}\left(\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\right)^{\frac{1}{2}}\left(\frac{\sigma h_{\boldsymbol{T}}^2}{\hat{\mu}}\right)^{\frac{1}{2}}\|\nabla p_h\|_{0,\boldsymbol{T}}\|\boldsymbol{v}_h\|_{0,\boldsymbol{T}}\\ &\leqslant \alpha\left(\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\|\nabla p_h\|_{0,\boldsymbol{T}}^2\right)^{\frac{1}{2}}\sigma^{\frac{1}{2}}\|\boldsymbol{v}_h\|_0\\ &\leqslant c_{\boldsymbol{\Omega}}\,\alpha\left(\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\|\nabla p_h\|_{0,\boldsymbol{T}}^2\right)^{\frac{1}{2}}\frac{\|p_h\|_0}{\hat{\mu}^{\frac{1}{2}}}.\end{aligned} \tag{5.3.8}$$

For the boundary term we apply the Cauchy–Schwarz inequality (Thm. 3.1.5), the Cauchy–Schwarz inequality for sums (Lem. 3.1.2), the inequality (4.1.6c), and the first estimate in (5.3.4) to derive

$$\begin{aligned}\langle\mu_{\text{eff}}\nabla\boldsymbol{v}_h\cdot\boldsymbol{n},\boldsymbol{u}_h\rangle &\leqslant \left(\mu_{\text{eff}}\sum_{E\in\mathcal{G}_h}h_E\|\nabla\boldsymbol{v}_h\cdot\boldsymbol{n}_E\|_{0,E}^2\right)^{\frac{1}{2}}\left(\sum_{E\in\mathcal{G}_h}\frac{\mu_{\text{eff}}}{h_E}\|\boldsymbol{u}_h\|_{0,E}^2\right)^{\frac{1}{2}}\\ &\leqslant \left(\mu_{\text{eff}}\widetilde{c}_{\text{DTI}}\|\nabla\boldsymbol{v}_h\|_{0,\boldsymbol{\Omega}}^2\right)^{\frac{1}{2}}\left(\sum_{E\in\mathcal{G}_h}\frac{\mu_{\text{eff}}}{h_E}\|\boldsymbol{u}_h\|_{0,E}^2\right)^{\frac{1}{2}}\\ &\leqslant c_{\boldsymbol{\Omega}}\widetilde{c}_{\text{DTI}}^{\frac{1}{2}}\frac{\|p_h\|_0}{\hat{\mu}^{\frac{1}{2}}}\underbrace{\left(\frac{\mu_{\text{eff}}}{\hat{\mu}}\right)^{\frac{1}{2}}}_{=\theta^{\frac{1}{2}}}\left(\sum_{E\in\mathcal{G}_h}\frac{\mu_{\text{eff}}}{h_E}\|\boldsymbol{u}_h\|_{0,E}^2\right)^{\frac{1}{2}}.\end{aligned} \tag{5.3.9}$$

Inserting (5.3.7), (5.3.8), and (5.3.9) into (5.3.6), using the estimates (5.3.4), (5.3.2b), (5.2.7), and rearranging the terms one obtains

$$\begin{aligned}\mathcal{A}_h\left[(\boldsymbol{u}_h,p_h);(\boldsymbol{v}_h,0)\right] \geqslant\ &\frac{\|p_h\|_0^2}{\hat{\mu}} - c_{\boldsymbol{\Omega}}\,(N\delta)^{\frac{1}{2}}\left(\delta\,\hat{\mu}\|\nabla\cdot\boldsymbol{u}_h\|_0^2\right)^{\frac{1}{2}}\frac{\|p_h\|_0}{\hat{\mu}^{\frac{1}{2}}}\\ &- c_{\boldsymbol{\Omega}}\left(\mu_{\text{eff}}^{\frac{1}{2}}\|\nabla\boldsymbol{u}_h\|_0\right)\frac{\|p_h\|_0}{\hat{\mu}^{\frac{1}{2}}} - c_{\boldsymbol{\Omega}}\,(1+\alpha)\,\sigma^{\frac{1}{2}}\|\boldsymbol{u}_h\|_0\frac{\|p_h\|_0}{\hat{\mu}^{\frac{1}{2}}}\\ &- \left(\widetilde{c_{\text{SZ}}}\hat{c}_{\boldsymbol{\Omega}}\alpha^{-\frac{1}{2}} + c_{\boldsymbol{\Omega}}\alpha^{\frac{1}{2}}\right)\left(\alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\|\nabla p_h\|_{0,\boldsymbol{T}}^2\right)^{\frac{1}{2}}\frac{\|p_h\|_0}{\hat{\mu}^{\frac{1}{2}}}\\ &- c_{\boldsymbol{\Omega}}\widetilde{c}_{\text{DTI}}^{\frac{1}{2}}\left(\sum_{E\in\mathcal{G}_h}\theta\frac{\mu_{\text{eff}}}{h_E}\|\boldsymbol{u}_h\|_{0,E}^2\right)^{\frac{1}{2}}\frac{\|p_h\|_0}{\hat{\mu}^{\frac{1}{2}}}.\end{aligned}$$

We now define

$$C_1' := \max\left\{ c_{\boldsymbol{\Omega}}^2 \left(1+\alpha\right)^2,\ \frac{\left(\widetilde{c_{\mathrm{SZ}}}\widehat{c}_{\boldsymbol{\Omega}} + c_{\boldsymbol{\Omega}}\alpha\right)^2}{\alpha},\ c_{\boldsymbol{\Omega}}^2 N\delta,\ c_{\boldsymbol{\Omega}}^2 \widetilde{c}_{\mathrm{DTI}} \right\}$$

and using the Young inequality (Thm. 3.1.1 (ii)) with $\varepsilon = 5$, we obtain the estimate

$$\begin{aligned}\mathcal{A}_h\left[(\boldsymbol{u}_h, p_h);(\boldsymbol{v}_h, 0)\right] \geqslant \frac{1}{2}\frac{\|p_h\|_0^2}{\hat{\mu}} \\ - C_1\left(|||\boldsymbol{u}_h|||^2 + \alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\|\nabla p_h\|_{0,\boldsymbol{T}}^2 + \sum_{E\in\mathcal{G}_h}\theta\frac{\mu_{\mathrm{eff}}}{h_E}\|\boldsymbol{u}_h\|_{0,E}^2\right),\end{aligned}$$

with $C_1 := 3C_1'$. □

**Remark 5.3.5** (On the Behavior of $C_1$)**.**
*The constant $C_1$ in Lemma 5.3.4 depends only on the stabilization parameters $\alpha$ and $\delta$, and on the domain $\boldsymbol{\Omega}$ and its discretization (through the constants $N, c_{\boldsymbol{\Omega}}$, $\widehat{c}_{\boldsymbol{\Omega}}$, $\widetilde{c_{\mathrm{SZ}}}$, and $\widetilde{c}_{\mathrm{DTI}}$). In particular $C_1 \sim \frac{1}{\alpha}$ for $\alpha \ll 1$.*

### Boundary Control

The next steps concern the stability of the proposed formulation with respect to the velocity at the boundary. Therefore, we will construct two test functions, which can be used to show that the Nitsche terms in (5.2.10b) together with the corner stabilization term, provide sufficient control on the boundary norms of the velocity.

The construction of the first test function and its main properties are stated in the following lemma.

**Lemma 5.3.6.** *Let us assume that the family of triangulations $\{\mathcal{T}_h\}_{h>0}$ satisfies assumption $\mathbf{M}_3$. For any $\boldsymbol{u}_h \in \boldsymbol{V}_h^1$ we define $\boldsymbol{w}_h^{\boldsymbol{u}_h} \in \boldsymbol{V}_h^1$ such that, for each mesh node $\boldsymbol{x}$, it holds*

$$\boldsymbol{w}_h^{\boldsymbol{u}_h}(\boldsymbol{x}) := \begin{cases} \boldsymbol{u}_h(\boldsymbol{x}), & \text{for } \boldsymbol{x}\in\Gamma, \\ \boldsymbol{0}, & \text{for } \boldsymbol{x}\in\overline{\boldsymbol{\Omega}}\setminus\Gamma. \end{cases} \tag{5.3.10}$$

*Then the function $\boldsymbol{w}_h^{\boldsymbol{u}_h}$ satisfies the following properties:*

*(1) There exist two positive constants $c_0$ and $c_1$, depending only on the regularity of the mesh, such that*

$$\left\langle \mu_{\mathrm{eff}}\nabla\boldsymbol{w}_h^{\boldsymbol{u}_h}\cdot\boldsymbol{n}, \boldsymbol{u}_h\right\rangle \geqslant c_0\sum_{E\in\mathcal{G}_h}\frac{\mu_{\mathrm{eff}}}{h_E}\|\boldsymbol{u}_h\|_{0,E}^2 - c_1\mu_{\mathrm{eff}}\|\nabla\boldsymbol{u}_h\|_0^2. \tag{5.3.11}$$

*(2) There exists a constant $c_2 > 0$, depending only on the regularity of the mesh, such that*

$$\mu_{\mathrm{eff}}\left\|\nabla\boldsymbol{w}_h^{\boldsymbol{u}_h}\right\|_0^2 \leqslant c_2\sum_{E\in\mathcal{G}_h}\frac{\mu_{\mathrm{eff}}}{h_E}\|\boldsymbol{u}_h\|_{0,E}^2. \tag{5.3.12}$$

*(3) There exists a constant $c_3 > 0$, depending on the mesh regularity, such that*

$$\left\|\boldsymbol{w}_h^{\boldsymbol{u}_h}\right\|_0 \leqslant c_3 \|\boldsymbol{u}_h\|_0. \tag{5.3.13}$$

*(4) There exists a constant $\widetilde{c_3} > 0$, depending on the mesh regularity, such that*

$$\left\|\boldsymbol{w}_h^{\boldsymbol{u}_h}\right\|_{0,\boldsymbol{T}}^2 \leqslant \widetilde{c_3} h_{\boldsymbol{T}}^2 \left\|\nabla \boldsymbol{w}_h^{\boldsymbol{u}_h}\right\|_{0,\boldsymbol{T}}^2, \quad \forall \boldsymbol{T} \in \mathcal{T}_h\,. \tag{5.3.14}$$

*Proof:* Let us consider $\boldsymbol{u}_h \in \boldsymbol{V}_h^1$ and let $\boldsymbol{w}_h^{\boldsymbol{u}_h} \in \boldsymbol{V}_h^1$ be defined as in (5.3.10) and sketched in Figure 5.5.
(1): In order to prove (5.3.11), let us introduce the following notation. For an edge $E \in \mathcal{G}_h$ with vertices $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ we will denote the (unique) attached triangle by $\boldsymbol{T}_E = \operatorname{conv}\{\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2\}$ and introduce $\boldsymbol{w}_E \colon \mathbb{R}^2 \to \mathbb{R}^2$ as the linear function that coincides with $\boldsymbol{w}_h^{\boldsymbol{u}_h}$ in $\boldsymbol{T}_E$ and extends it everywhere in $\mathbb{R}^2$.

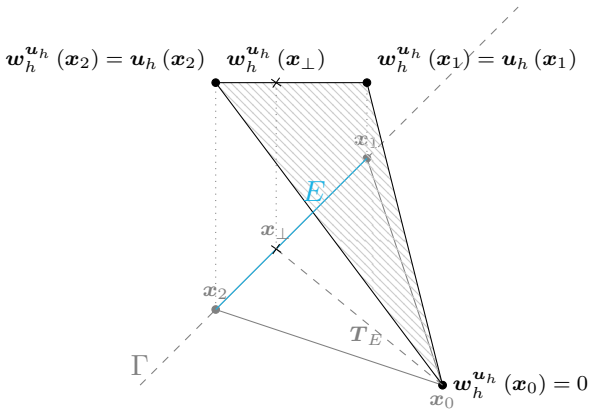



Figure 5.5: Illustration of the function $\boldsymbol{w}_h^{\boldsymbol{u}_h}$ on $\boldsymbol{T}_E$ with $\boldsymbol{x}_0 \in \overline{\boldsymbol{\Omega}} \setminus \Gamma$ and $\boldsymbol{x}_\perp \in E$.

By assumption $\mathbf{M}_3$ (see Subsection 5.2.1) and without loss of generality, let $\boldsymbol{x}_0 \in \overline{\boldsymbol{\Omega}} \setminus \Gamma$ such that by definition (5.3.10) it is $\boldsymbol{w}_h^{\boldsymbol{u}_h}(\boldsymbol{x}_0) = \mathbf{0}$. Then it holds

$$\left(\nabla \boldsymbol{w}_h^{\boldsymbol{u}_h} \cdot \boldsymbol{n}_E\right)\big|_E = \frac{\boldsymbol{w}_E(\boldsymbol{x}_\perp)}{h_{E,\perp}}, \tag{5.3.15}$$

where $\boldsymbol{x}_\perp$ is the perpendicular foot of the vertex $\boldsymbol{x}_0$ and $h_{E,\perp}$ is the height of the triangle $\boldsymbol{T}_E$ with respect to the edge $E$. Depending on the shape of $\boldsymbol{T}_E$, $\boldsymbol{x}_\perp$ might fall inside or outside the edge $E$. Formally, there exists an $a \in \mathbb{R}$, such that

$$\boldsymbol{x}_\perp = \boldsymbol{x}_2 + a(\boldsymbol{x}_1 - \boldsymbol{x}_2) = a\boldsymbol{x}_1 + (1-a)\boldsymbol{x}_2, \quad |a| + |1-a| \leqslant M, \tag{5.3.16}$$

where $M > 0$ depends only on the mesh regularity constant in (4.1.2). Hence, by adding and subtracting $\boldsymbol{u}_h$ and using (5.3.15), we can reformulate

$$\begin{aligned}
\left\langle \nabla \boldsymbol{w}_h^{\boldsymbol{u}_h} \cdot \boldsymbol{n}, \boldsymbol{u}_h \right\rangle_E &= \frac{1}{h_{E,\perp}} \left\langle \boldsymbol{w}_E(\boldsymbol{x}_\perp), \boldsymbol{u}_h \right\rangle_E \\
&= \frac{1}{h_{E,\perp}} \left( \langle \boldsymbol{u}_h, \boldsymbol{u}_h \rangle_E - \langle \boldsymbol{u}_h - \boldsymbol{w}_E(\boldsymbol{x}_\perp), \boldsymbol{u}_h \rangle_E \right) \\
&= \frac{h_E}{h_{E,\perp}} \left( \frac{\|\boldsymbol{u}_h\|_{0,E}^2}{h_E} \right) - \frac{1}{h_{E,\perp}} \left\langle \boldsymbol{u}_h - \boldsymbol{w}_E(\boldsymbol{x}_\perp), \boldsymbol{u}_h \right\rangle_E .
\end{aligned} \tag{5.3.17}$$

Exploiting (5.3.16), the linearity of $\boldsymbol{w}_E$, and the fact that $\boldsymbol{w}_E$ coincides with $\boldsymbol{u}_h$ on $E$ (by definition), we get with $1 = a + (1-a)$ and the triangle inequality, for all $\boldsymbol{x} \in E$,

$$\begin{aligned}
|\boldsymbol{u}_h(\boldsymbol{x}) - \boldsymbol{w}_E(\boldsymbol{x}_\perp)| &= |\boldsymbol{u}_h(\boldsymbol{x}) - (a\boldsymbol{u}_h(\boldsymbol{x}_1) + (1-a)\,\boldsymbol{u}_h(\boldsymbol{x}_2))| \\
&\leqslant |a|\,|\boldsymbol{u}_h(\boldsymbol{x}) - \boldsymbol{u}_h(\boldsymbol{x}_1)| + |1-a|\,|\boldsymbol{u}_h(\boldsymbol{x}) - \boldsymbol{u}_h(\boldsymbol{x}_2)| \\
&\leqslant M\left(|\boldsymbol{u}_h(\boldsymbol{x}) - \boldsymbol{u}_h(\boldsymbol{x}_1)| + |\boldsymbol{u}_h(\boldsymbol{x}) - \boldsymbol{u}_h(\boldsymbol{x}_2)|\right) \\
&\leqslant 2M\, h_E \left|(\nabla \boldsymbol{u}_h)\,|_{\boldsymbol{T}_E}\right|,
\end{aligned}$$

where $|\cdot|$ stands for the Euclidean norm (3.1.1). Since $\nabla \boldsymbol{u}_h$ is constant on $\boldsymbol{T}_E$, it also holds

$$\|\nabla \boldsymbol{u}_h\|_{0,\boldsymbol{T}_E} = |\boldsymbol{T}_E|^{\frac{1}{2}} \left|(\nabla \boldsymbol{u}_h)\,|_{\boldsymbol{T}_E}\right|,$$

from which we deduce

$$\left|(\nabla \boldsymbol{u}_h)\,|_{\boldsymbol{T}_E}\right| \leqslant c h_E^{-1} \|\nabla \boldsymbol{u}_h\|_{0,\boldsymbol{T}_E},$$

where the constant $c > 0$ only depends on the regularity of the mesh. The above arguments allow to conclude

$$\begin{aligned}
\|\boldsymbol{u}_h - \boldsymbol{w}_E(\boldsymbol{x}_\perp)\|_{0,E} &\leqslant h_E^{\frac{1}{2}} \max_{\boldsymbol{x}\in E} |\boldsymbol{u}_h(\boldsymbol{x}) - \boldsymbol{w}_E(\boldsymbol{x}_\perp)| \leqslant 2M h_E^{\frac{3}{2}} \left|(\nabla \boldsymbol{u}_h)\,|_{\boldsymbol{T}_E}\right| \\
&\leqslant c_\Gamma\, h_E^{\frac{1}{2}} \|\nabla \boldsymbol{u}_h\|_{0,\boldsymbol{T}_E}, \qquad (5.3.18)
\end{aligned}$$

with $c_\Gamma := 2Mc$. Thus, applying the Cauchy–Schwarz inequality (Thm. 3.1.5), the inequality (5.3.18), and the Young inequality (Thm. 3.1.1 (ii)) with $\varepsilon = 1$ yields

$$\langle \boldsymbol{u}_h - \boldsymbol{w}_E(\boldsymbol{x}_\perp), \boldsymbol{u}_h \rangle_E \leqslant c_\Gamma\, h_E^{\frac{1}{2}} \|\nabla \boldsymbol{u}_h\|_{0,\boldsymbol{T}_E} \|\boldsymbol{u}_h\|_{0,E} \leqslant \frac{1}{2}\|\boldsymbol{u}_h\|_{0,E}^2 + c_\Gamma^2 \frac{h_E}{2} \|\nabla \boldsymbol{u}_h\|_{0,\boldsymbol{T}_E}^2.$$

Combining this inequality with (5.3.17) leads to

$$\left\langle \mu_{\text{eff}} \nabla \boldsymbol{w}_h^{\boldsymbol{u}_h} \cdot \boldsymbol{n}, \boldsymbol{u}_h \right\rangle_E \geqslant \frac{1}{2} \frac{h_E}{h_{E,\perp}} \left( \frac{\mu_{\text{eff}}}{h_E} \|\boldsymbol{u}_h\|_{0,E}^2 \right) - c_\Gamma^2 \frac{h_E}{2h_{E,\perp}} \left( \mu_{\text{eff}} \|\nabla \boldsymbol{u}_h\|_{0,\boldsymbol{T}_E}^2 \right).$$

The proof is concluded by taking the sum over all boundary edges and defining

$$c_0 := \frac{1}{2} \min_{E\in\mathcal{G}_h} \left\{ \frac{h_E}{h_{E,\perp}} \right\}, \qquad c_1 := \frac{c_\Gamma^2}{2} \max_{E\in\mathcal{G}_h} \left\{ \frac{h_E}{h_{E,\perp}} \right\},$$

which are only dependent on the shape regularity of the mesh.
(2): Let us consider a triangle $\boldsymbol{T} = \operatorname{conv}\{\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2\}$ such that by assumption $\mathbf{M}_3$ (see Subsection 5.2.1) and without loss of generality, $\boldsymbol{x}_0 \notin \Gamma$, i.e., $\boldsymbol{w}_h^{\boldsymbol{u}_h}(\boldsymbol{x}_0) = \mathbf{0}$. Then we can show as an immediate consequence of the scaling argument (5.1.11b) and the definition of $\boldsymbol{w}_h^{\boldsymbol{u}_h}$ in (5.3.6) (in particular $\boldsymbol{w}_h^{\boldsymbol{u}_h}(\boldsymbol{x}_0) = \mathbf{0}$), that there exists a constant $c > 0$ depending only on the mesh regularity, such that

$$\left\|\nabla \boldsymbol{w}_h^{\boldsymbol{u}_h}\right\|_{0,\boldsymbol{T}}^2 \leqslant \begin{cases} c \sum_{i=1}^{N} |\boldsymbol{u}_h(\boldsymbol{x}_i)|^2, & \text{if } \boldsymbol{T} \cap \Gamma \neq \emptyset, \\ 0, & \text{otherwise,} \end{cases}$$

where $N \in \{1,2\}$ is the number of vertices of $\boldsymbol{T}$ which are on the boundary. Hence, denoting by $N_\Gamma$ the total number of boundary nodes, by $c_{\mathrm{NB}}$ the maximum number of triangles adjacent to a boundary node (which is bounded depending on the smallest angle of the triangulation $\mathcal{T}_h$), using the inequality $a^2+b^2 \leqslant 2(a^2+b^2+ab)$, for $a,b \in \mathbb{R}$, and (5.1.11c) component-wise, one can write

$$\mu_{\mathrm{eff}} \left\| \nabla \boldsymbol{w}_h^{\boldsymbol{u}_h} \right\|_0^2 \leqslant \mu_{\mathrm{eff}} c_{\mathrm{NB}} c \sum_{i=1}^{N_\Gamma} \left| \boldsymbol{u}_h(\boldsymbol{x}_i) \right|^2 \leqslant 2\mu_{\mathrm{eff}} c_{\mathrm{NB}} c \sum_{E \in \mathcal{G}_h} \frac{3}{h_E} \|\boldsymbol{u}_h\|_{0,E}^2.$$

Defining $c_2 := 6 c_{\mathrm{NB}} c$, which depends on the regularity of the mesh, gives the desired estimate.
(3): The inequality (5.3.13) can be proven using scaling arguments (Lem. 5.1.2) similar to the previous ones. If all three vertices of a triangle $\boldsymbol{T} = \operatorname{conv}\{\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2\}$ lie in the interior, then by the definition of $\boldsymbol{w}_h^{\boldsymbol{u}_h}$ (5.3.6) we have $\left\| \boldsymbol{w}_h^{\boldsymbol{u}_h} \right\|_{0,\boldsymbol{T}} = 0$. Otherwise, by assumption $\mathbf{M}_3$ and without loss of generality, let $\boldsymbol{x}_0$ be an interior node, i.e., $\boldsymbol{w}_h^{\boldsymbol{u}_h}(\boldsymbol{x}_0) = \mathbf{0}$. Then (5.1.11a) and the definition of $\boldsymbol{w}_h^{\boldsymbol{u}_h}$ in (5.3.6) imply for each component $w_h^{\boldsymbol{u}_h,i}$ of $\boldsymbol{w}_h^{\boldsymbol{u}_h}$, $i = 1, \ldots, n$,

$$\begin{aligned} \left\| w_h^{\boldsymbol{u}_h,i} \right\|_{0,\boldsymbol{T}}^2 &= \frac{|\boldsymbol{T}|}{6} \left( w_h^{\boldsymbol{u}_h,i}(\boldsymbol{x}_1)^2 + w_h^{\boldsymbol{u}_h,i}(\boldsymbol{x}_2)^2 + w_h^{\boldsymbol{u}_h,i}(\boldsymbol{x}_1)\, w_h^{\boldsymbol{u}_h,i}(\boldsymbol{x}_2) \right) \\ &= \frac{|\boldsymbol{T}|}{6} \left( u_h^i(\boldsymbol{x}_1)^2 + u_h^i(\boldsymbol{x}_2)^2 + u_h^i(\boldsymbol{x}_1)\, u_h^i(\boldsymbol{x}_2) \right), \end{aligned} \tag{5.3.19}$$

where $u_h^i$ denotes the $i$th component of $\boldsymbol{u}_h$. Let us consider the polynomial

$$p(x,y,z) := (x+y)^2 + (x+z)^2 + (y+z)^2, \quad x,y,z \in \mathbb{R},$$

which is non-negative and quadratic, thus has a local minimum. Let us calculate $z$ such that $\frac{\partial p}{\partial z} = 0$:

$$0 = \frac{\partial p}{\partial z} = 2(x+z) + 2(y+z) \qquad \Longleftrightarrow \qquad z^\star = -\frac{1}{2}(x+y).$$

Then,

$$\frac{1}{2} p(x,y,z) \geqslant \frac{1}{2} p(x,y,z^\star) = \frac{1}{2} \left( \frac{3}{2} x^2 + xy + \frac{3}{2} y^2 \right) \geqslant \frac{1}{2} \left( x^2 + xy + y^2 \right).$$

Using (5.1.11a), the above result with $x \equiv u_h^i(\boldsymbol{x}_1)$, $y \equiv u_h^i(\boldsymbol{x}_2)$, and $z \equiv u_h^i(\boldsymbol{x}_0)$, and (5.3.19), we can estimate component-wise

$$\begin{aligned} \left\| u_h^i \right\|_{0,\boldsymbol{T}}^2 &= \frac{|\boldsymbol{T}|}{6} \frac{1}{2} \left[ \left( u_h^i(\boldsymbol{x}_2) + u_h^i(\boldsymbol{x}_1) \right)^2 + \left( u_h^i(\boldsymbol{x}_2) + u_h^i(\boldsymbol{x}_0) \right)^2 + \left( u_h^i(\boldsymbol{x}_1) + u_h^i(\boldsymbol{x}_0) \right)^2 \right] \\ &\geqslant \frac{|\boldsymbol{T}|}{6} \frac{1}{2} \left[ u_h^i(\boldsymbol{x}_1)^2 + u_h^i(\boldsymbol{x}_1)\, u_h^i(\boldsymbol{x}_2) + u_h^i(\boldsymbol{x}_2)^2 \right] \\ &= \frac{1}{2} \left\| w_h^{\boldsymbol{u}_h,i} \right\|_{0,\boldsymbol{T}}^2. \end{aligned}$$

If only one vertex of $\boldsymbol{T}$ is on $\Gamma$ we can further use $a^2 \leqslant 2\left(a^2+b^2+ab\right)$ and obtain then, after summation over all $\boldsymbol{T} \in \mathcal{T}_h$, the desired result with $c_3 := 4$.

(4): Let us recall that for a real, symmetric matrix $M \in \mathbb{R}^{n\times n}$, the *Rayleigh quotient* for $\boldsymbol{x} \in \mathbb{R}^n$, defined as the ratio

$$R_M(\boldsymbol{x}) := \frac{\boldsymbol{x}^T M \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}},$$

satisfies the property (known as theorem of Courant–Fischer/Min-Max principle [CL02, Thm. 5.2])

$$\lambda_{\min}(M) \leqslant R_M(\boldsymbol{x}) \leqslant \lambda_{\max}(M), \quad \forall \boldsymbol{x} \in \mathbb{R}^n,$$

which is equivalent to

$$\lambda_{\min}(M)\, \boldsymbol{x}^T \boldsymbol{x} \leqslant \boldsymbol{x}^T M \boldsymbol{x} \leqslant \lambda_{\max}(M)\, \boldsymbol{x}^T \boldsymbol{x}. \tag{5.3.20}$$

Here $\lambda_{\min}(M)$ and $\lambda_{\max}(M)$ denote the smallest and largest eigenvalue of $M$, respectively. Note that if $M$ is symmetric and positive semi-definite, then all its eigenvalues are real and non-negative. Let us consider, w.l.o.g., the case $\boldsymbol{w}_h^{\boldsymbol{u}_h}(\boldsymbol{x}_0) = \boldsymbol{0}$ (we can argue analogous for two vertices in the interior) and a component $w_h^{\boldsymbol{u}_h,i}$ of $\boldsymbol{w}_h^{\boldsymbol{u}_h}$, $i = 1,\dots,n$, with the notation $w_k := w_h^{\boldsymbol{u}_h,i}(\boldsymbol{x}_k)$, i.e., $w_0 = 0$. As in the proof of Lemma 5.1.2 and with (5.3.20), we have

$$\begin{aligned}
\left\|\nabla w_h^{\boldsymbol{u}_h,i}\right\|_{0,\boldsymbol{T}}^2 &= |\boldsymbol{T}| \left| \left(\boldsymbol{B}_{\boldsymbol{T}}^{-T}\right) \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \right|^2 = |\boldsymbol{T}| \left( \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}^T \underbrace{\left(\boldsymbol{B}_{\boldsymbol{T}}^{-T}\right)^T \left(\boldsymbol{B}_{\boldsymbol{T}}^{-T}\right)}_{=:M} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \right) \\
&\geqslant |\boldsymbol{T}| \left( \lambda_{\min}(M) \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}^T \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \right) = |\boldsymbol{T}| \lambda_{\min}(M) \left(w_1^2 + w_2^2\right) \\
&\geqslant |\boldsymbol{T}| \lambda_{\min}(M)\, \frac{2}{3} \underbrace{\left(w_1^2 + w_2^2 + w_1 w_2\right)}_{= \frac{6}{|\boldsymbol{T}|} \left\| w_h^{\boldsymbol{u}_h,i} \right\|_{0,\boldsymbol{T}}^2} \\
&= \frac{4}{\lambda_{\max}(M^{-1})} \left\| w_h^{\boldsymbol{u}_h,i} \right\|_{0,\boldsymbol{T}}^2,
\end{aligned} \tag{5.3.21}$$

where we have used $a^2 + b^2 \geqslant \frac{2}{3}\left(a^2 + b^2 + ab\right)$ and that $M$ symmetric, positive semi-definite, and moreover invertible (since $B_{\boldsymbol{T}}$ is invertible). By definition $M^{-1} = B_{\boldsymbol{T}}^T B_{\boldsymbol{T}}$ we can use (5.1.4) to obtain

$$M^{-1} = \begin{pmatrix} \|\boldsymbol{x}_1 - \boldsymbol{x}_0\|^2 & (\boldsymbol{x}_2 - \boldsymbol{x}_0)^T (\boldsymbol{x}_1 - \boldsymbol{x}_0) \\ (\boldsymbol{x}_2 - \boldsymbol{x}_0)^T (\boldsymbol{x}_1 - \boldsymbol{x}_0) & \|\boldsymbol{x}_2 - \boldsymbol{x}_0\|^2 \end{pmatrix} =: \begin{pmatrix} \boldsymbol{m}_1^T \boldsymbol{m}_1 & \boldsymbol{m}_2^T \boldsymbol{m}_1 \\ \boldsymbol{m}_2^T \boldsymbol{m}_1 & \boldsymbol{m}_2^T \boldsymbol{m}_2 \end{pmatrix}$$

with $\boldsymbol{m}_1 := \boldsymbol{x}_1 - \boldsymbol{x}_0$ and $\boldsymbol{m}_2 := \boldsymbol{x}_2 - \boldsymbol{x}_0$. Therefore, the maximum eigenvalue of $M^{-1}$ is given by

$$\begin{aligned}
\lambda_{\max}\left(M^{-1}\right) &= \frac{\|\boldsymbol{m}_1\|^2 + \|\boldsymbol{m}_2\|^2}{2} + \sqrt{\frac{\left(\|\boldsymbol{m}_1\|^2 + \|\boldsymbol{m}_2\|^2\right)^2}{4} - \Big(\|\boldsymbol{m}_1\|^2 \|\boldsymbol{m}_2\|^2 - (\underbrace{\boldsymbol{m}_2^T \boldsymbol{m}_1}_{\|\boldsymbol{m}_1\|\|\boldsymbol{m}_2\| \cos(\angle(\boldsymbol{m}_1, \boldsymbol{m}_2)) =})^2\Big)} \\
&\leqslant \frac{\|\boldsymbol{m}_1\|^2 + \|\boldsymbol{m}_2\|^2}{2} + \sqrt{\frac{\|\boldsymbol{m}_1\|^4 - 2\|\boldsymbol{m}_1\|^2\|\boldsymbol{m}_2\|^2 + \|\boldsymbol{m}_2\|^4 + 4\|\boldsymbol{m}_1\|^2\|\boldsymbol{m}_2\|^2}{4}} \\
&= \|\boldsymbol{m}_1\|^2 + \|\boldsymbol{m}_2\|^2 \leqslant 2h_{\boldsymbol{T}}^2,
\end{aligned} \tag{5.3.22}$$

where we have used $\cos\left(\angle\left(\boldsymbol{m}_1, \boldsymbol{m}_2\right)\right)^2 \in (0,1)$ (for non-degenerate triangles) and the fact that $\|\boldsymbol{x}_i - \boldsymbol{x}_j\| \leqslant h_{\boldsymbol{T}}$ for $i,j = 0,1,2$, $i \neq j$. Inserting (5.3.22) into (5.3.21) and summing up the components yields the desired result with $\widetilde{c}_3 := \frac{1}{2}$. □

Utilizing the function $\boldsymbol{w}_h^{\boldsymbol{u}_h}$ defined in (5.3.10), the next lemma states stability of the boundary velocity.

**Lemma 5.3.7** (Part III – Boundary Control I)**.**
*Let $\alpha, \delta, \rho, \mu_{\text{eff}}, \sigma \geqslant 0$ with $\mu_{\text{eff}} + \sigma > 0$. For any $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h^1 \times Q_h^1$, there exist a function $\boldsymbol{w}_h \in \boldsymbol{V}_h^1$ and a constant $C_2 = C_2(\alpha, \delta) > 0$ which is independent of the physical parameters, of $\boldsymbol{u}_h$, and of $h$, such that*

$$\mathcal{A}_h\left[(\boldsymbol{u}_h, p_h); (\boldsymbol{w}_h, 0)\right] \geqslant \frac{c_0}{4} \sum_{E \in \mathcal{G}_h} \theta \, \frac{\mu_{\text{eff}}}{h_E} \|\boldsymbol{u}_h\|_{0,E}^2 - C_2 \left( |||\boldsymbol{u}_h|||^2 + \alpha \sum_{\boldsymbol{T} \in \mathcal{T}_h} \frac{h_{\boldsymbol{T}}^2}{\hat{\mu}} \|\nabla p_h\|_{0,\boldsymbol{T}}^2 \right),$$

*where $c_0$ is the constant defined in Lemma 5.3.6 and $\theta$ is given in* (5.2.11b).

*Proof:* For a given pair $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h^1 \times Q_h^1$, let $\boldsymbol{w}_h := \theta \boldsymbol{w}_h^{\boldsymbol{u}_h}$, where $\boldsymbol{w}_h^{\boldsymbol{u}_h}$ is the function defined in Lemma 5.3.6. Then, we get

$$\begin{aligned}
\mathcal{A}_h\left[(\boldsymbol{u}_h, p_h); (\boldsymbol{w}_h, 0)\right] &= \theta \mu_{\text{eff}} \left(\nabla \boldsymbol{u}_h, \nabla \boldsymbol{w}_h^{\boldsymbol{u}_h}\right) + \theta \left(\sigma \boldsymbol{u}_h, \boldsymbol{w}_h^{\boldsymbol{u}_h}\right) \\
&\quad - \theta \left\langle \mu_{\text{eff}} \nabla \boldsymbol{u}_h \cdot \boldsymbol{n}, \boldsymbol{w}_h^{\boldsymbol{u}_h} \right\rangle + \theta \left\langle \mu_{\text{eff}} \nabla \boldsymbol{w}_h^{\boldsymbol{u}_h} \cdot \boldsymbol{n}, \boldsymbol{u}_h \right\rangle \\
&\quad - \theta \left(p_h, \nabla \cdot \boldsymbol{w}_h^{\boldsymbol{u}_h}\right) + \theta \left\langle p_h \boldsymbol{n}, \boldsymbol{w}_h^{\boldsymbol{u}_h} \right\rangle \\
&\quad + \alpha\theta \sum_{\boldsymbol{T} \in \mathcal{T}_h} \frac{h_{\boldsymbol{T}}^2}{\hat{\mu}} \left(\sigma \boldsymbol{u}_h + \nabla p_h, \sigma \boldsymbol{w}_h^{\boldsymbol{u}_h}\right)_{\boldsymbol{T}} \\
&\quad + \delta\, \hat{\mu}\theta \left(\nabla \cdot \boldsymbol{u}_h, \nabla \cdot \boldsymbol{w}_h^{\boldsymbol{u}_h}\right) + \rho\, \hat{\mu}\theta \sum_{\boldsymbol{x} \in \mathcal{C}} \left|[\![\boldsymbol{u}_h \cdot \boldsymbol{n}]\!](\boldsymbol{x})\right|^2 .
\end{aligned}$$

Observing that the last term (corresponding to the corner stabilization) is always non-negative, that $\theta \leqslant 1$, using the Cauchy–Schwarz inequality (Thm. 3.1.5), and the inequalities (5.3.11), (5.3.12), and (5.3.13) leads to

$$\begin{aligned}
\mathcal{A}_h\left[(\boldsymbol{u}_h, p_h); (\boldsymbol{w}_h, 0)\right] \geqslant\; & -c_2^{\frac{1}{2}} \mu_{\text{eff}}^{\frac{1}{2}} \|\nabla \boldsymbol{u}_h\|_0 \underbrace{\theta^{\frac{1}{2}}}_{\leqslant 1} \left( \sum_{E \in \mathcal{G}_h} \theta \, \frac{\mu_{\text{eff}}}{h_E} \|\boldsymbol{u}_h\|_{0,E}^2 \right)^{\frac{1}{2}} - \sigma c_3 \|\boldsymbol{u}_h\|_0^2 \\
& \underbrace{- \theta \left\langle \mu_{\text{eff}} \nabla \boldsymbol{u}_h \cdot \boldsymbol{n}, \boldsymbol{w}_h^{\boldsymbol{u}_h} \right\rangle}_{=:Q_0} + c_0 \sum_{E \in \mathcal{G}_h} \theta \, \frac{\mu_{\text{eff}}}{h_E} \|\boldsymbol{u}_h\|_{0,E}^2 \\
& - c_1 \mu_{\text{eff}} \|\nabla \boldsymbol{u}_h\|_0^2 \underbrace{- \theta \left(p_h, \nabla \cdot \boldsymbol{w}_h^{\boldsymbol{u}_h}\right) + \theta \left\langle p_h \boldsymbol{n}, \boldsymbol{w}_h^{\boldsymbol{u}_h} \right\rangle}_{=:Q_1} \\
& \underbrace{+ \alpha\theta \sum_{\boldsymbol{T} \in \mathcal{T}_h} \frac{h_{\boldsymbol{T}}^2}{\hat{\mu}} \left(\sigma \boldsymbol{u}_h + \nabla p_h, \sigma \boldsymbol{w}_h^{\boldsymbol{u}_h}\right)_{\boldsymbol{T}}}_{=:Q_2} \\
& \underbrace{+ \delta\, \hat{\mu}\theta \left(\nabla \cdot \boldsymbol{u}_h, \nabla \cdot \boldsymbol{w}_h^{\boldsymbol{u}_h}\right)}_{=:Q_3} . \qquad (5.3.23)
\end{aligned}$$

Combining the Cauchy–Schwarz inequality (Thm. 3.1.5), the Cauchy–Schwarz inequality for sums (Lem. 3.1.2), the trace inequality (4.1.6c), and the fact $\left\|\boldsymbol{w}_h^{\boldsymbol{u}_h}\right\|_{0,E} = \|\boldsymbol{u}_h\|_{0,E}$ we obtain

$$\begin{aligned} Q_0 &\leqslant \theta\mu_{\text{eff}} \sum_{E\in\mathcal{G}_h} \left(\frac{h_E}{h_E}\right)^{\frac{1}{2}} \|\nabla\boldsymbol{u}_h\cdot\boldsymbol{n}_E\|_{0,E}\left\|\boldsymbol{w}_h^{\boldsymbol{u}_h}\right\|_{0,E} \\ &\leqslant (\widetilde{c}_{\text{DTI}}\mu_{\text{eff}})^{\frac{1}{2}} \|\nabla\boldsymbol{u}_h\|_0 \underbrace{\theta^{\frac{1}{2}}}_{\leqslant 1} \left(\sum_{E\in\mathcal{G}_h} \theta\,\frac{\mu_{\text{eff}}}{h_E}\|\boldsymbol{u}_h\|_{0,E}^2\right)^{\frac{1}{2}}. \end{aligned} \tag{5.3.24}$$

In order to bound the term $\mathcal{Q}_1$, we use the integration by parts formula, the Cauchy–Schwarz inequality (Thm. 3.1.5), the Cauchy–Schwarz inequality for sums (Lem. 3.1.2), the inequalities (5.3.14), (5.3.12), $\theta\hat{\mu} = \mu_{\text{eff}}$, and $\theta \leqslant 1$ to obtain

$$\begin{aligned} \mathcal{Q}_1 &= \sum_{\boldsymbol{T}\in\mathcal{T}_h} \theta\left(\nabla p_h, \boldsymbol{w}_h^{\boldsymbol{u}_h}\right)_{\boldsymbol{T}} \leqslant \theta\left(\alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\|\nabla p_h\|_{0,\boldsymbol{T}}^2\right)^{\frac{1}{2}}\left(\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{\hat{\mu}}{\alpha h_{\boldsymbol{T}}^2}\left\|\boldsymbol{w}_h^{\boldsymbol{u}_h}\right\|_{0,\boldsymbol{T}}^2\right)^{\frac{1}{2}} \\ &\leqslant \left(\alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\|\nabla p_h\|_{0,\boldsymbol{T}}^2\right)^{\frac{1}{2}}\left(\frac{\theta^2\hat{\mu}\widetilde{c_3}}{\alpha}\left\|\nabla\boldsymbol{w}_h^{\boldsymbol{u}_h}\right\|_0^2\right)^{\frac{1}{2}} \\ &\leqslant \left(\frac{c_2\widetilde{c_3}}{\alpha}\right)^{\frac{1}{2}}\left(\alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\|\nabla p_h\|_{0,\boldsymbol{T}}^2\right)^{\frac{1}{2}}\left(\sum_{E\in\mathcal{G}_h}\theta\,\frac{\mu_{\text{eff}}}{h_E}\|\boldsymbol{u}_h\|_{0,E}^2\right)^{\frac{1}{2}}. \end{aligned} \tag{5.3.25}$$

Next, we observe that the term $\mathcal{Q}_2$, coming from the pressure stabilization, can be bounded using the Cauchy–Schwarz inequality (Thm. 3.1.5), (5.2.7), $\theta \leqslant 1$, the Cauchy–Schwarz inequality for sums (Lem. 3.1.2), Young's inequality (Thm. 3.1.1 (ii)) with $\varepsilon = 2$, and (5.3.13) as

$$\begin{aligned} \mathcal{Q}_2 &\geqslant -\theta\alpha^{\frac{1}{2}}\left(\alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\|\nabla p_h\|_{0,\boldsymbol{T}}^2\right)^{\frac{1}{2}}\left(\sum_{\boldsymbol{T}\in\mathcal{T}_h}\underbrace{\frac{\sigma h_{\boldsymbol{T}}^2}{\hat{\mu}}}_{\leqslant 1}\sigma\left\|\boldsymbol{w}_h^{\boldsymbol{u}_h}\right\|_{0,\boldsymbol{T}}^2\right)^{\frac{1}{2}} - \theta\alpha\sigma c_3\|\boldsymbol{u}_h\|_0^2 \\ &\geqslant -\left(\alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\|\nabla p_h\|_{0,\boldsymbol{T}}^2\right)^{\frac{1}{2}}(\alpha\sigma)^{\frac{1}{2}}c_3\|\boldsymbol{u}_h\|_0 - \alpha\sigma c_3\|\boldsymbol{u}_h\|_0^2 \\ &\geqslant -\frac{1}{4}\alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\|\nabla p_h\|_{0,\boldsymbol{T}}^2 - \left(c_3^2 + c_3\right)\alpha\sigma\|\boldsymbol{u}_h\|_0^2. \end{aligned} \tag{5.3.26}$$

Finally, the Cauchy–Schwarz inequality (Thm. 3.1.5), Corollary 3.1.4, (5.3.12),

and $(\hat{\mu}\theta)^{\frac{1}{2}} = \mu_{\text{eff}}^{\frac{1}{2}}$ allow also to conclude

$$\begin{aligned}\mathcal{Q}_3 &\geqslant -\delta\,(\hat{\mu}\theta N)^{\frac{1}{2}}\,\|\nabla\cdot\boldsymbol{u}_h\|_0\,(\hat{\mu}\theta)^{\frac{1}{2}}\,\left\|\nabla\boldsymbol{w}_h^{\boldsymbol{u}_h}\right\|_0 \\ &\geqslant -\left(\delta N c_2\right)^{\frac{1}{2}}\left(\delta\,\hat{\mu}\|\nabla\cdot\boldsymbol{u}_h\|_0^2\right)^{\frac{1}{2}}\left(\sum_{E\in\mathcal{G}_h}\theta\,\frac{\mu_{\text{eff}}}{h_E}\|\boldsymbol{u}_h\|_{0,E}^2\right)^{\frac{1}{2}}. \qquad (5.3.27)\end{aligned}$$

Inserting (5.3.24), (5.3.25), (5.3.26), and (5.3.27) into (5.3.23), and reordering the terms yields

$$\begin{aligned}\mathcal{A}_h\left[(\boldsymbol{u}_h,p_h)\,;(\boldsymbol{w}_h,0)\right] \geqslant\; & c_0\sum_{E\in\mathcal{G}_h}\theta\,\frac{\mu_{\text{eff}}}{h_E}\|\boldsymbol{u}_h\|_{0,E}^2 - c_1\mu_{\text{eff}}\|\nabla\boldsymbol{u}_h\|_0^2 \\ & - \left(c_3+\alpha\left(c_3^2+c_3\right)\right)\sigma\|\boldsymbol{u}_h\|_0^2 \\ & - \left(c_2^{\frac{1}{2}}+\widetilde{c}_{\text{DTI}}^{\frac{1}{2}}\right)\mu_{\text{eff}}^{\frac{1}{2}}\|\nabla\boldsymbol{u}_h\|_0\left(\sum_{E\in\mathcal{G}_h}\theta\,\frac{\mu_{\text{eff}}}{h_E}\|\boldsymbol{u}_h\|_{0,E}^2\right)^{\frac{1}{2}} \\ & - \frac{1}{4}\alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\|\nabla p_h\|_{0,\boldsymbol{T}}^2 \\ & - \left(\frac{c_2\widetilde{c_3}}{\alpha}\right)^{\frac{1}{2}}\left(\alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\|\nabla p_h\|_{0,\boldsymbol{T}}^2\right)^{\frac{1}{2}}\left(\sum_{E\in\mathcal{G}_h}\theta\,\frac{\mu_{\text{eff}}}{h_E}\|\boldsymbol{u}_h\|_{0,E}^2\right)^{\frac{1}{2}} \\ & - \left(\delta N c_2\right)^{\frac{1}{2}}\left(\delta\,\hat{\mu}\|\nabla\cdot\boldsymbol{u}_h\|_0^2\right)^{\frac{1}{2}}\left(\sum_{E\in\mathcal{G}_h}\theta\,\frac{\mu_{\text{eff}}}{h_E}\|\boldsymbol{u}_h\|_{0,E}^2\right)^{\frac{1}{2}}.\end{aligned}$$

Applying three times the Young inequality yields, for any $\varepsilon>0$,

$$\begin{aligned}\mathcal{A}_h\left[(\boldsymbol{u}_h,p_h)\,;(\boldsymbol{w}_h,0)\right] \geqslant\; & c_0\sum_{E\in\mathcal{G}_h}\theta\,\frac{\mu_{\text{eff}}}{h_E}\|\boldsymbol{u}_h\|_{0,E}^2 - c_1\mu_{\text{eff}}\|\nabla\boldsymbol{u}_h\|_0^2 \\ & - \left(c_3+\alpha\left(c_3^2+c_3\right)\right)\sigma\|\boldsymbol{u}_h\|_0^2 - \left(c_2^{\frac{1}{2}}+\widetilde{c}_{\text{DTI}}^{\frac{1}{2}}\right)^2\frac{\mu_{\text{eff}}}{2\varepsilon}\|\nabla\boldsymbol{u}_h\|_0^2 \\ & - \frac{\varepsilon}{2}\sum_{E\in\mathcal{G}_h}\theta\,\frac{\mu_{\text{eff}}}{h_E}\|\boldsymbol{u}_h\|_{0,E}^2 - \frac{1}{4}\,\alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\|\nabla p_h\|_{0,\boldsymbol{T}}^2 \\ & - \frac{c_2\widetilde{c_3}}{2\alpha\varepsilon}\,\alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\|\nabla p_h\|_{0,\boldsymbol{T}}^2 - \frac{\varepsilon}{2}\sum_{E\in\mathcal{G}_h}\theta\,\frac{\mu_{\text{eff}}}{h_E}\|\boldsymbol{u}_h\|_{0,E}^2 \\ & - \frac{\delta N c_2}{2\varepsilon}\,\delta\,\hat{\mu}\|\nabla\cdot\boldsymbol{u}_h\|_0^2 - \frac{\varepsilon}{2}\sum_{E\in\mathcal{G}_h}\theta\,\frac{\mu_{\text{eff}}}{h_E}\|\boldsymbol{u}_h\|_{0,E}^2.\end{aligned}$$

Choosing $\varepsilon = \frac{c_0}{2}$ allows to conclude

$$\begin{aligned}\mathcal{A}_h\left[(\boldsymbol{u}_h, p_h); (\boldsymbol{w}_h, 0)\right] \geqslant \frac{c_0}{4} \sum_{E\in\mathcal{G}_h} \theta\, \frac{\mu_{\text{eff}}}{h_E}\|\boldsymbol{u}_h\|_{0,E}^2 \\ - \left(c_1 + \frac{\left(c_2^{\frac{1}{2}} + \widetilde{c}_{\text{DTI}}^{\frac{1}{2}}\right)^2}{c_0}\right)\mu_{\text{eff}}\|\nabla \boldsymbol{u}_h\|_0^2 \\ - \left(c_3 + \alpha\left(c_3^2 + c_3\right)\right)\sigma\|\boldsymbol{u}_h\|_0^2 - \frac{\delta N c_2}{c_0}\left(\delta\,\hat{\mu}\|\nabla\cdot\boldsymbol{u}_h\|_0^2\right) \\ - \left(\frac{1}{4} + \frac{c_2\widetilde{c_3}}{\alpha c_0}\right)\alpha \sum_{\boldsymbol{T}\in\mathcal{T}_h} \frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\|\nabla p_h\|_{0,\boldsymbol{T}}^2.\end{aligned}$$

The proof is completed by defining

$$C_2 := \max\left\{c_1 + \frac{\left(c_2^{\frac{1}{2}} + \widetilde{c}_{\text{DTI}}^{\frac{1}{2}}\right)^2}{c_0}, \left(c_3 + \alpha\left(c_3^2 + c_3\right)\right), \frac{\delta N c_2}{c_0}, \frac{1}{4} + \frac{c_2\widetilde{c_3}}{\alpha c_0}\right\}.$$

□

**Remark 5.3.8** (On the Behavior of $C_2$)**.**
*The constant $C_2$ depends on the stabilization parameters $\alpha$ and $\delta$, and for $\alpha \ll 1$ we have $C_2 \sim \frac{1}{\alpha}$.*

**Remark 5.3.9** (On the Role of $\theta$)**.**
*We observe that the scaling of the velocity test function in Lemma 5.3.7 (and the norm of the velocity at the boundary) by $\theta$ allows to derive robust (independent of the physical parameters) stability estimates for the terms involving $\hat{\mu}\nabla\cdot\boldsymbol{w}_h$. In detail, this scaling (and the equality $\hat{\mu}\theta = \mu_{\text{eff}}$) is in particular crucial for obtaining the estimates for $Q_1$ and $Q_3$ in the proof of Theorem 5.3.7, which are valid for $\mu_{\text{eff}} \longrightarrow 0$. Notice as well that the scaling by $\theta$ implies that the test function $\theta\boldsymbol{w}_h^{\boldsymbol{u}_h}$ vanishes in the Darcy limit ($\mu_{\text{eff}} = 0$).*

The fourth test function pair, utilized to show the fulfillment of the inf-sup condition, is related to the control of the normal velocity at the boundary, which is particularly important in order to guarantee stability towards the Darcy limit, i.e., for $\mu_{\text{eff}} = 0$.

For the proof of the subsequent lemma, in particular the assumptions $\mathbf{M}_1$ and $\mathbf{M}_2$ on the mesh ((5.2.2) and (5.2.4)) will be used.

**Lemma 5.3.10.** *Let us assume that the family of triangulations $\{\mathcal{T}_h\}_{h>0}$ satisfies $\mathbf{M}_1$, $\mathbf{M}_2$ ((5.2.2), (5.2.4)), and assumption $\mathbf{M}_3$. For a given $\boldsymbol{u}_h \in \boldsymbol{V}_h^1$, let us define $q_h^{\boldsymbol{u}_h} \in Q_h^1$ as the function whose values at the boundary nodes are uniquely defined to satisfy the $L^2$-projection property*

$$\left\langle q_h^{\boldsymbol{u}_h}, \varphi_h\right\rangle = -\sum_{E\in\mathcal{G}_h}\frac{1}{h_E}\left\langle \boldsymbol{u}_h\cdot\boldsymbol{n}_E, \varphi_h\right\rangle_E, \quad \forall \varphi_h \in Q_h^1, \tag{5.3.28}$$

*and its value at the interior nodes is given by a constant $c_q$, chosen in order to satisfy $\int_{\boldsymbol{\Omega}} q_h^{\boldsymbol{u}_h} = 0$.*
*Then the function $q_h^{\boldsymbol{u}_h}$ has the following properties:*

1. *There exists a constant $c_4 > 0$, depending only on $\eta_0$, such that*

$$\sum_{E \in \mathcal{G}_h} h_E \left\| q_h^{\boldsymbol{u}_h} \right\|_{0,E}^2 \leqslant c_4 \sum_{E \in \mathcal{G}_h} \frac{1}{h_E} \| \boldsymbol{u}_h \cdot \boldsymbol{n}_E \|_{0,E}^2. \tag{5.3.29}$$

2. *There exists a constant $c_5 > 0$, depending only on $\eta_0$, such that*

$$- \left\langle q_h^{\boldsymbol{u}_h}, \boldsymbol{u}_h \cdot \boldsymbol{n} \right\rangle \geqslant \frac{1}{2} \sum_{E \in \mathcal{G}_h} \frac{1}{h_E} \| \boldsymbol{u}_h \cdot \boldsymbol{n}_E \|_{0,E}^2 - c_5 \sum_{\boldsymbol{x} \in \mathcal{C}} \left| [\![ \boldsymbol{u}_h \cdot \boldsymbol{n} ]\!] (\boldsymbol{x}) \right|^2. \tag{5.3.30}$$

3. *There exists a constant $c_6 > 0$, depending only on the properties of the mesh, such that*

$$\left\| q_h^{\boldsymbol{u}_h} \right\|_0^2 \leqslant c_6 \sum_{E \in \mathcal{G}_h} \frac{1}{h_E} \| \boldsymbol{u}_h \cdot \boldsymbol{n}_E \|_{0,E}^2 \tag{5.3.31}$$

*and*

$$\sum_{\boldsymbol{T} \in \mathcal{T}_h} h_{\boldsymbol{T}}^2 \left\| \nabla q_h^{\boldsymbol{u}_h} \right\|_{0,\boldsymbol{T}}^2 \leqslant c_6 c_I^2 \sum_{E \in \mathcal{G}_h} \frac{1}{h_E} \| \boldsymbol{u}_h \cdot \boldsymbol{n}_E \|_{0,E}^2. \tag{5.3.32}$$

*Proof:*
(1): In order to prove (5.3.29), for simplicity and without loss of generality, consider the case of a boundary with a single connected component. In this case, let us number the boundary nodes as $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N_\Gamma}$ and the boundary edges as $E_1, \ldots, E_{N_\Gamma}$ such that the edge $E_i$ connects the nodes $\boldsymbol{x}_i$ and $\boldsymbol{x}_{i+1}$, for all $i \in \{1, \ldots, N_\Gamma\}$. Moreover, we identify $\boldsymbol{x}_{N_\Gamma+1}$ with $\boldsymbol{x}_1$, so that the above defined convention is well-defined also for $i = N_\Gamma$. To simplify the notation, let us abbreviate $h_i = h_{E_i}$ and $q_i = q_h^{\boldsymbol{u}_h}(\boldsymbol{x}_i)$. We now consider a function $\varphi_h \in Q_h^1$ defined at each node $\boldsymbol{x}_i$ of the mesh by

$$\varphi_h(\boldsymbol{x}_i) := \begin{cases} h_i q_i, & \boldsymbol{x}_i \in \Gamma, \\ c_\varphi, & \text{otherwise}, \end{cases}$$

where $c_\varphi$ is a constant defined in order to have $\int_\Omega \varphi_h = 0$.

We will first prove that there exists a constant $\hat{c} > 0$, independent of $q_h^{\boldsymbol{u}_h}$ and $h_i$, such that for any $E_i \in \mathcal{G}_h$ it holds

$$h_i \int_{E_i} |q_h^{\boldsymbol{u}_h}|^2 \leqslant \hat{c} \int_{E_i} q_h^{\boldsymbol{u}_h} \varphi_h. \tag{5.3.33}$$

This will justify the particular bound in assumption (5.2.2). On any boundary edge $E_i$, by the linearity of $q_h^{\boldsymbol{u}_h}$ and $\varphi_h$, the application of Simpson's rule (Thm. 5.1.1) yields

$$\int_{E_i} q_h^{\boldsymbol{u}_h} \varphi_h = \frac{h_i}{6} \left( 2 h_i q_i^2 + (h_i + h_{i+1}) q_i q_{i+1} + 2 h_{i+1} q_{i+1}^2 \right)$$

and

$$\int_{E_i} |q_h^{\boldsymbol{u}_h}|^2 = \frac{h_i}{3}\left(q_i^2 + q_i q_{i+1} + q_{i+1}^2\right).$$

If $q_i = 0$, (5.3.33) holds if $\frac{h_i}{h_{i+1}} \leqslant \hat{c}$, i.e., with $\hat{c} \geqslant \eta_0$ in (5.2.2). Assume now that $q_i \neq 0$ and set $\eta := \frac{h_{i+1}}{h_i}$ and $t := \frac{q_{i+1}}{q_i}$. The inequality (5.3.33) is then equivalent to the inequality

$$1 + t + t^2 \leqslant \frac{\hat{c}}{2}\left(2 + (1+\eta)\,t + 2\eta t^2\right).$$

Since $\eta > 0$, $1+t+t^2 > 0$, for all $t \in \mathbb{R}$, and (by assumption (5.2.2)) $\eta \in [1/\eta_0, \eta_0]$, the above condition is equivalent to

$$\inf_{\substack{\frac{1}{\eta_0} \leqslant \eta \leqslant \eta_0 \\ t \in \mathbb{R}}} \frac{2 + (1+\eta)\,t + 2\eta t^2}{1 + t + t^2} > 0.$$

Since the polynomial in the denominator is always strictly positive, the whole infimum is strictly positive if the numerator as a polynomial in $t$ is strictly positive for all $t \in \mathbb{R}$ and for all $\eta \in [1/\eta_0, \eta_0]$. We observe that it has a strictly positive value for $t = 0$. Hence, this polynomial is strictly positive for all $t \in \mathbb{R}$ if and only if it has no zeros, which is equivalent to its discriminant $D(\eta) := (1+\eta)^2 - 16\eta$ having no real square roots for the selected range of $\eta$. The discriminant $D(\eta)$ vanishes if $\eta = 7 \pm 4\sqrt{3}$ and is strictly negative if $\eta \in \left(7 - 4\sqrt{3},\, 7 + 4\sqrt{3}\right) = \left(\frac{1}{7+4\sqrt{3}},\, 7 + 4\sqrt{3}\right)$. Hence, $D(\eta)$ is negative if $\eta_0 < 7 + 4\sqrt{3}$, which is exactly the bound given in assumption (5.2.2). Estimate (5.3.33) is therefore proven by setting

$$\hat{c} = \hat{c}(\eta_0) := \max\left\{\eta_0,\, 2\left(\inf_{\substack{\frac{1}{\eta_0} \leqslant \eta \leqslant \eta_0 \\ t \in \mathbb{R}}} \frac{2 + (1+\eta)\,t + 2\eta t^2}{1 + t + t^2}\right)^{-1}\right\}.$$

Using Simpson's rule (Thm. 5.1.1) for $\|\varphi_h\|_{0,E_i}^2$ and $\left\|q_h^{\boldsymbol{u}_h}\right\|_{0,E_i}^2$, and the assumption $\frac{h_{i+1}}{h_i} \leqslant \eta_0$, yields on any boundary edge $E_i$

$$\begin{aligned}
\frac{1}{h_i}\|\varphi_h\|_{0,E_i}^2 &= \frac{1}{h_i}\frac{h_i}{6}\left(h_i^2|q_i|^2 + |h_i q_i + h_{i+1} q_{i+1}|^2 + h_{i+1}^2|q_{i+1}|^2\right) \\
&\leqslant \frac{1}{h_i}\left(h_i^2\left\|q_h^{\boldsymbol{u}_h}\right\|_{0,E_i}^2 + h_i h_{i+1}\left\|q_h^{\boldsymbol{u}_h}\right\|_{0,E_i}^2 + h_{i+1}^2\left\|q_h^{\boldsymbol{u}_h}\right\|_{0,E_i}^2\right) \\
&= h_i\left(1 + \frac{h_{i+1}}{h_i} + \left(\frac{h_{i+1}}{h_i}\right)^2\right)\left\|q_h^{\boldsymbol{u}_h}\right\|_{0,E_i}^2 \\
&\leqslant \left(1 + \eta_0 + \eta_0^2\right) h_i\left\|q_h^{\boldsymbol{u}_h}\right\|_{0,E_i}^2. \qquad (5.3.34)
\end{aligned}$$

Summing (5.3.33) over all boundary edges and using (5.3.28), the Cauchy–Schwarz inequality (Thm. 3.1.5), the Cauchy–Schwarz inequality for sums

(Lem. 3.1.2), (5.3.34), and the Young inequality (Thm. 3.1.1) yields

$$\begin{aligned}\sum_{E\in\mathcal{G}_h} h_E\big\|q_h^{\boldsymbol{u}_h}\big\|_{0,E}^2 &\leqslant \hat{c}\,\big\langle q_h^{\boldsymbol{u}_h},\varphi_h\big\rangle = -\hat{c}\sum_{E\in\mathcal{G}_h}\frac{1}{h_E}\langle \boldsymbol{u}_h\cdot\boldsymbol{n}_E,\varphi_h\rangle_E\\ &\leqslant \hat{c}\left(\sum_{E\in\mathcal{G}_h}\frac{1}{h_E}\|\boldsymbol{u}_h\cdot\boldsymbol{n}_E\|_{0,E}^2\right)^{\frac{1}{2}}\left(\sum_{E\in\mathcal{G}_h}\frac{1}{h_E}\|\varphi_h\|_{0,E}^2\right)^{\frac{1}{2}}\\ &\leqslant \hat{c}\left(\frac{\varepsilon}{2}\sum_{E\in\mathcal{G}_h}\frac{1}{h_E}\|\boldsymbol{u}_h\cdot\boldsymbol{n}_E\|_{0,E}^2+\frac{\left(1+\eta_0+\eta_0^2\right)}{2\varepsilon}\sum_{E\in\mathcal{G}_h}h_E\big\|q_h^{\boldsymbol{u}_h}\big\|_{0,E}^2\right).\end{aligned}$$

Choosing $\varepsilon := \hat{c}\left(1+\eta_0+\eta_0^2\right)$ leads to

$$\sum_{E\in\mathcal{G}_h} h_E\big\|q_h^{\boldsymbol{u}_h}\big\|_{0,E}^2 \leqslant \hat{c}^2\left(1+\eta_0+\eta_0^2\right)\sum_{E\in\mathcal{G}_h}\frac{1}{h_E}\|\boldsymbol{u}_h\cdot\boldsymbol{n}_E\|_{0,E}^2.$$

Estimate (5.3.29) is obtained defining $c_4 := \hat{c}^2\left(1+\eta_0+\eta_0^2\right)$, which only depends on $\eta_0$.

(2): To prove the inequality (5.3.30), let us consider the function $\varphi_h \in Q_h^1$ such that, at the mesh nodes $\boldsymbol{x}$, it holds

$$\varphi_h(\boldsymbol{x}) := \begin{cases} \boldsymbol{u}_h\cdot\boldsymbol{n}(\boldsymbol{x}), & \text{for } \boldsymbol{x}\in\Gamma\setminus\mathcal{C},\\ \frac{1}{2}\left(\boldsymbol{u}_h\cdot\boldsymbol{n}_E(\boldsymbol{x})+\boldsymbol{u}_h\cdot\boldsymbol{n}_{E'}(\boldsymbol{x})\right), & \text{for } \boldsymbol{x}\in\mathcal{C}, \text{ with}\\ & E\cap E'=\boldsymbol{x},\ E,E'\in\mathcal{G}_h,\\ c_\varphi, & \text{otherwise,}\end{cases}$$

with a constant $c_\varphi$ defined in order to have $\int_\Omega \varphi_h = 0$. Remember that $\mathcal{C}$ is the set of corner nodes at the boundary, see (5.2.1).

Using (5.3.28), the Cauchy–Schwarz inequality (Thm. 3.1.5), and the Young inequality (Thm. 3.1.1) we obtain, for any $\varepsilon > 0$,

$$\begin{aligned}-\big\langle q_h^{\boldsymbol{u}_h},\boldsymbol{u}_h\cdot\boldsymbol{n}\big\rangle &= -\big\langle q_h^{\boldsymbol{u}_h},\varphi_h\big\rangle - \big\langle q_h^{\boldsymbol{u}_h},\boldsymbol{u}_h\cdot\boldsymbol{n}-\varphi_h\big\rangle\\ &= \sum_{E\in\mathcal{G}_h}\frac{1}{h_E}\left\langle \boldsymbol{u}_h\cdot\boldsymbol{n}_E,\varphi_h\right\rangle_E - \sum_{E\in\mathcal{G}_h}\left\langle q_h^{\boldsymbol{u}_h},\boldsymbol{u}_h\cdot\boldsymbol{n}_E-\varphi_h\right\rangle_E\\ &= \sum_{E\in\mathcal{G}_h}\frac{1}{h_E}\left\langle \boldsymbol{u}_h\cdot\boldsymbol{n}_E,\boldsymbol{u}_h\cdot\boldsymbol{n}_E-(\boldsymbol{u}_h\cdot\boldsymbol{n}_E-\varphi_h)\right\rangle_E\\ &\quad - \sum_{E\in\mathcal{G}_h}\left\langle q_h^{\boldsymbol{u}_h},\boldsymbol{u}_h\cdot\boldsymbol{n}_E-\varphi_h\right\rangle_E\\ &= \sum_{E\in\mathcal{G}_h}\frac{1}{h_E}\|\boldsymbol{u}_h\cdot\boldsymbol{n}_E\|_{0,E}^2 - \sum_{E\in\mathcal{G}_h}\frac{1}{h_E}\langle \boldsymbol{u}_h\cdot\boldsymbol{n}_E,\boldsymbol{u}_h\cdot\boldsymbol{n}_E-\varphi_h\rangle_E\\ &\quad - \sum_{E\in\mathcal{G}_h}\left\langle q_h^{\boldsymbol{u}_h},\boldsymbol{u}_h\cdot\boldsymbol{n}_E-\varphi_h\right\rangle_E\\ &\geqslant \sum_{E\in\mathcal{G}_h}\frac{1}{h_E}\left(\left(1-\frac{\varepsilon}{2}\right)\|\boldsymbol{u}_h\cdot\boldsymbol{n}_E\|_{0,E}^2-\frac{1}{\varepsilon}\|\boldsymbol{u}_h\cdot\boldsymbol{n}_E-\varphi_h\|_{0,E}^2\right)\\ &\quad -\frac{\varepsilon}{2}\sum_{E\in\mathcal{G}_h}h_E\big\|q_h^{\boldsymbol{u}_h}\big\|_{0,E}^2. \end{aligned}\tag{5.3.35}$$

By the definition of $\varphi_h$, the function $(\boldsymbol{u}_h \cdot \boldsymbol{n} - \varphi_h)\,|_{E\in\mathcal{G}_h}$ might be different from zero only on boundary edges that are adjacent to a corner. Thus, let us consider a corner node $\boldsymbol{x}_c \in \mathcal{C}$ with an adjacent edge $E = \overline{\boldsymbol{x}_i\boldsymbol{x}_c}$. It holds

$$|(\boldsymbol{u}_h \cdot \boldsymbol{n}_E - \varphi_h)(\boldsymbol{x}_c)| = \frac{1}{2}\,|[\![\boldsymbol{u}_h \cdot \boldsymbol{n}_E]\!](\boldsymbol{x}_c)| \qquad \text{and} \qquad (\boldsymbol{u}_h \cdot \boldsymbol{n}_E - \varphi_h)(\boldsymbol{x}_i) = 0,$$

which yields due to Simpson's rule (Thm. (5.1.1))

$$\|\boldsymbol{u}_h \cdot \boldsymbol{n}_E - \varphi_h\|_{0,E}^2 = \frac{1}{12} h_E\,|[\![\boldsymbol{u}_h \cdot \boldsymbol{n}_E]\!](\boldsymbol{x}_c)|^2\,. \tag{5.3.36}$$

Thus, inserting (5.3.29) and (5.3.36) into (5.3.35) and choosing $\varepsilon := \frac{1}{c_4+1}$ we obtain

$$-\left\langle q_h^{\boldsymbol{u}_h}, \boldsymbol{u}_h \cdot \boldsymbol{n}\right\rangle \geqslant \frac{1}{2}\sum_{E\in\mathcal{G}_h}\frac{1}{h_E}\|\boldsymbol{u}_h \cdot \boldsymbol{n}_E\|_{0,E}^2 - c_5\sum_{\boldsymbol{x}\in\mathcal{C}}|[\![\boldsymbol{u}_h \cdot \boldsymbol{n}]\!](\boldsymbol{x})|^2\,,$$

with $c_5 := 2\frac{c_4+1}{12}$ (since every corner is contained in two boundary edges), which only depends on $\eta_0$.
(3): To prove (5.3.31), let us first introduce a continuous, element-wise linear function $q_0^{\boldsymbol{u}_h}$, which coincides with $q_h^{\boldsymbol{u}_h}$ on the boundary $\Gamma$ and vanishes at all the interior nodes, i.e.,

$$q_0^{\boldsymbol{u}_h}(\boldsymbol{x}_i) := \begin{cases} q_h^{\boldsymbol{u}_h}(\boldsymbol{x}_i)\,, & \boldsymbol{x}_i \in \Gamma, \\ 0, & \text{otherwise,} \end{cases}$$

and a continuous, element-wise linear function $\psi_h$ vanishing on the boundary and being equal to 1 at all the interior nodes, i.e.,

$$\psi_h(\boldsymbol{x}_i) := \begin{cases} 0, & \boldsymbol{x}_i \in \Gamma, \\ 1, & \text{otherwise.} \end{cases}$$

By linearity we have the representation

$$q_h^{\boldsymbol{u}_h} = q_0^{\boldsymbol{u}_h} + c_q\psi_h, \tag{5.3.37}$$

where $c_q$ is described in Lemma 5.3.10.

Integrating (5.3.37) and using the property $\int_{\boldsymbol{\Omega}} q_h^{\boldsymbol{u}_h} = 0$ yields

$$c_q = -\frac{\int_{\boldsymbol{\Omega}} q_0^{\boldsymbol{u}_h}}{\int_{\boldsymbol{\Omega}}\psi_h} \qquad \text{and} \qquad |c_q| \leqslant \frac{\|q_0^{\boldsymbol{u}_h}\|_0|\boldsymbol{B}_h|^{\frac{1}{2}}}{\int_{\boldsymbol{\Omega}}\psi_h},$$

where for the second estimate the Cauchy–Schwarz inequality (Thm. 3.1.5), the fact that $q_0^{\boldsymbol{u}_h}$ is different from 0 only on $\boldsymbol{B}_h$ (defined in (5.2.3)), and $\int_{\boldsymbol{\Omega}}\psi_h > 0$ were used. Hence, with the triangle inequality, the definition of $\psi_h$, and assumption (5.2.4) (which guarantees the existence of $\omega < 1$ such that for all $h$ it holds $|\boldsymbol{B}_h| \leqslant \omega|\boldsymbol{\Omega}|$), it follows

$$\begin{aligned}\left\|q_h^{\boldsymbol{u}_h}\right\|_0 &\leqslant \|q_0^{\boldsymbol{u}_h}\|_0 + |c_q|\|\psi_h\|_0 \leqslant \left(1 + \frac{\|\psi_h\|_0|\boldsymbol{B}_h|^{\frac{1}{2}}}{\int_{\boldsymbol{\Omega}}\psi_h}\right)\|q_0^{\boldsymbol{u}_h}\|_0 \\ &\leqslant \left(1 + \frac{|\boldsymbol{\Omega}|^{\frac{1}{2}}|\boldsymbol{B}_h|^{\frac{1}{2}}}{|\boldsymbol{\Omega}| - |\boldsymbol{B}_h|}\right)\|q_0^{\boldsymbol{u}_h}\|_0 \leqslant \left(1 + \frac{\omega^{\frac{1}{2}}}{1-\omega}\right)\|q_0^{\boldsymbol{u}_h}\|_0. \end{aligned} \tag{5.3.38}$$

Since $0 < \omega < 1$, the coefficient inside the parentheses is always strictly larger than one (and it approaches one on fine meshes).

By the definition of $q_0^{\boldsymbol{u}_h}$ we have

$$\|q_0^{\boldsymbol{u}_h}\|_0^2 = \sum_{\boldsymbol{T}\in\boldsymbol{B}_h} \|q_0^{\boldsymbol{u}_h}\|_{0,\boldsymbol{T}}^2.$$

Let us first consider a triangle $\boldsymbol{T}_E$ with $\boldsymbol{T}_E \cap \Gamma = E$ and without loss of generality $E = \overline{\boldsymbol{x}_1\boldsymbol{x}_2}$, i.e., in particular $\boldsymbol{x}_0$ is an interior node and thus $q_0^{\boldsymbol{u}_h}(\boldsymbol{x}_0) = 0$. Then the scaling arguments (Lem. 5.1.2) yield

$$\|q_0^{\boldsymbol{u}_h}\|_{0,\boldsymbol{T}_E}^2 = \frac{|\boldsymbol{T}_E|}{6}\left(q_h^{\boldsymbol{u}_h}(\boldsymbol{x}_1)^2 + q_h^{\boldsymbol{u}_h}(\boldsymbol{x}_2)^2 + q_h^{\boldsymbol{u}_h}(\boldsymbol{x}_1)\, q_h^{\boldsymbol{u}_h}(\boldsymbol{x}_2)\right)$$

and

$$h_E\big\|q_h^{\boldsymbol{u}_h}\big\|_{0,E}^2 = \frac{h_E^2}{3}\left(q_h^{\boldsymbol{u}_h}(\boldsymbol{x}_1)^2 + q_h^{\boldsymbol{u}_h}(\boldsymbol{x}_2)^2 + q_h^{\boldsymbol{u}_h}(\boldsymbol{x}_1)\, q_h^{\boldsymbol{u}_h}(\boldsymbol{x}_2)\right).$$

The area of the triangle $\boldsymbol{T}_E$ is given by $|\boldsymbol{T}_E| = \frac{1}{2}h_E h_{E,\perp}$, with $h_{E,\perp}$ denoting the height of $\boldsymbol{T}_E$ with respect to the edge $E$. Including also triangles with only one vertex at the boundary (recall assumption $\mathbf{M}_3$) we obtain

$$\|q_0^{\boldsymbol{u}_h}\|_0^2 \leqslant c_{\mathrm{NB}}\frac{1}{4}\max_{E\in\mathcal{G}_h}\left\{\frac{h_{E,\perp}}{h_E}\right\}\sum_{E\in\mathcal{G}_h} h_E\big\|q_h^{\boldsymbol{u}_h}\big\|_{0,E}^2, \tag{5.3.39}$$

where $c_{\mathrm{NB}}$ is the maximum number of triangles adjacent to a boundary node. Inserting (5.3.39) into (5.3.38) and using (5.3.29) proves (5.3.31) with a constant

$$c_6 := c_{NB}\,\frac{c_4}{4}\max_{E\in\mathcal{G}_h}\left\{\frac{h_{E,\perp}}{h_E}\right\}\left(1 + \frac{\omega^{\frac{1}{2}}}{1-\omega}\right)^2,$$

thus depending on $c_{\mathrm{NB}}$, $\omega$, the shape regularity of the mesh, and $\eta_0$ through $c_4$. Finally, (5.3.32) can be obtained combining the inverse inequality (4.1.3a) on each triangle and (5.3.31):

$$\sum_{\boldsymbol{T}\in\mathcal{T}_h} h_{\boldsymbol{T}}^2\big\|\nabla q_h^{\boldsymbol{u}_h}\big\|_{0,\boldsymbol{T}}^2 \leqslant c_I^2\sum_{\boldsymbol{T}\in\mathcal{T}_h}\big\|q_h^{\boldsymbol{u}_h}\big\|_{0,\boldsymbol{T}}^2 \leqslant c_6 c_I^2\sum_{E\in\mathcal{G}_h}\frac{1}{h_E}\|\boldsymbol{u}_h\cdot\boldsymbol{n}_E\|_{0,E}^2.$$

□

Finally, we are able to show control of the normal velocity for arbitrary values of physical parameters.

**Lemma 5.3.11** (Part IV – Boundary Control II)**.**
*Let $\delta, \rho > 0$, $\alpha \geqslant 0$, and $\mu_{\mathrm{eff}}, \sigma \geqslant 0$ with $\mu_{\mathrm{eff}} + \sigma > 0$ and let us assume that the family of triangulations $\{\mathcal{T}_h\}_{h>0}$ satisfies* (5.2.2), (5.2.4), *and assumption* $\mathbf{M}_3$. *Then, for any $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h^1 \times Q_h^1$, there exists a function $q_h \in Q_h^1$ and*

*a constant $C_3 = C_3(\alpha, \delta, \rho) > 0$ independent of the physical parameters, of $\boldsymbol{u}_h$, and of $h$, such that*

$$\mathcal{A}_h\left[(\boldsymbol{u}_h, p_h); (\boldsymbol{0}, q_h)\right] \geqslant \frac{1}{4}\sum_{E\in\mathcal{G}_h}\frac{\hat{\mu}}{h_E}\|\boldsymbol{u}_h\cdot\boldsymbol{n}_E\|_{0,E}^2 - C_3\left(|||\boldsymbol{u}_h|||^2 + \rho\hat{\mu}\sum_{\boldsymbol{x}\in\mathcal{C}}|[\![\boldsymbol{u}_h\cdot\boldsymbol{n}]\!](\boldsymbol{x})|^2 + \alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\|\nabla p_h\|_{0,\boldsymbol{T}}^2\right).$$

*Proof:* Let $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h^1 \times Q_h^1$ and let $q_h := \hat{\mu}\, q_h^{\boldsymbol{u}_h}$, where $q_h^{\boldsymbol{u}_h}$ is the function defined in Lemma 5.3.10. Using the Cauchy–Schwarz inequality (Thm. 3.1.5), Young's inequality (Thm. 3.1.1), (5.3.30), (5.3.31), (5.3.32), and (5.2.7) we obtain

$$\begin{aligned}
\mathcal{A}_h\left[(\boldsymbol{u}_h, p_h); (\boldsymbol{0}, q_h)\right] &= \left(\nabla\cdot\boldsymbol{u}_h, \hat{\mu}\, q_h^{\boldsymbol{u}_h}\right) - \left\langle\hat{\mu}\, q_h^{\boldsymbol{u}_h}, \boldsymbol{u}_h\cdot\boldsymbol{n}\right\rangle \\
&\quad + \alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\left(\sigma\boldsymbol{u}_h + \nabla p_h, \hat{\mu}\nabla q_h^{\boldsymbol{u}_h}\right)_{\boldsymbol{T}} \\
&\geqslant -\frac{1}{2\varepsilon}\hat{\mu}\delta\|\nabla\cdot\boldsymbol{u}_h\|_0^2 - \frac{\varepsilon\hat{\mu}}{2\delta}\left\|q_h^{\boldsymbol{u}_h}\right\|_0^2 - \hat{\mu}\left\langle q_h^{\boldsymbol{u}_h}, \boldsymbol{u}_h\cdot\boldsymbol{n}\right\rangle \\
&\quad - \alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\left(\sigma^2\frac{1}{2\varepsilon'}\|\boldsymbol{u}_h\|_{0,\boldsymbol{T}}^2 + \hat{\mu}^2\frac{\varepsilon'}{2}\left\|\nabla q_h^{\boldsymbol{u}_h}\right\|_{0,\boldsymbol{T}}^2\right) \\
&\quad - \alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\left(\frac{1}{2\varepsilon'}\|\nabla p_h\|_{0,\boldsymbol{T}}^2 + \hat{\mu}^2\frac{\varepsilon'}{2}\left\|\nabla q_h^{\boldsymbol{u}_h}\right\|_{0,\boldsymbol{T}}^2\right) \\
&\geqslant \left(\frac{1}{2} - \frac{\varepsilon}{2\delta}c_6 - \alpha c_I^2 c_6\varepsilon'\right)\sum_{E\in\mathcal{G}_h}\frac{\hat{\mu}}{h_E}\|\boldsymbol{u}_h\cdot\boldsymbol{n}_E\|_{0,E}^2 \\
&\quad - \frac{c_5}{\rho}\rho\,\hat{\mu}\sum_{\boldsymbol{x}\in\mathcal{C}}\left|[\![\boldsymbol{u}_h\cdot\boldsymbol{n}]\!](\boldsymbol{x})\right|^2 - \frac{1}{2\varepsilon}\hat{\mu}\delta\|\nabla\cdot\boldsymbol{u}_h\|_0^2 \\
&\quad - \alpha\frac{1}{2\varepsilon'}\sum_{\boldsymbol{T}\in\mathcal{T}_h}\underbrace{\frac{\sigma h_{\boldsymbol{T}}^2}{\hat{\mu}}}_{\leqslant 1}\sigma\|\boldsymbol{u}_h\|_{0,\boldsymbol{T}}^2 - \alpha\frac{1}{2\varepsilon'}\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\|\nabla p_h\|_{0,\boldsymbol{T}}^2,
\end{aligned}$$

for any $\varepsilon, \varepsilon' > 0$. The proof is completed by choosing $\varepsilon = \frac{\delta}{4c_6}$, $\varepsilon' = \frac{1}{8\alpha c_6 c_I^2}$, and defining

$$C_3 := \max\left\{\frac{c_5}{\rho}, \frac{2c_6}{\delta}, 4\alpha^2 c_6 c_I^2, 4\alpha c_6 c_I^2\right\},$$

which depends on the shape-regularity of the mesh and on the three stabilization parameters. □

**Remark 5.3.12** (On the Behavior of $C_3$)**.**
*In the proof of Lemma 5.3.11, $\delta^{-1}$ appears in the stability constant $C_3$. This is due to the fact that the integral $\left(\nabla\cdot\boldsymbol{u}_h, \hat{\mu}\, q_h^{\boldsymbol{u}_h}\right)$ has to be estimated in particular in the Darcy limit ($\mu_{\text{eff}} = 0$) where the gradient of the velocity has disappeared. Notice that, in order to assure the validity of Lemma 5.3.11, both, grad-div and corner stabilization, are required (i.e., $\delta, \rho > 0$). In particular, it holds $C_3 = O\left(\delta^{-1} + \rho^{-1}\right)$ for small values of $\delta$ and $\rho$. Moreover, $\alpha > 0$ is not required.*

**Remark 5.3.13** (On the Role of the Grad-Div Stabilization)**.**
*The proof of Lemma 5.3.11 clarifies the importance of the grad-div stabilization. Its presence provides not only control of the divergence norm of the velocity but it also allows to control the terms containing the divergence of the velocity in the formulation in a robust way for* $\mu_{\text{eff}}, \sigma \geqslant 0$*. Alternative estimates (without using a grad-div stabilization) could introduce non-robust dependencies, e.g., on* $\mu_{\text{eff}}^{-1}$ *(appearing when trying to obtain control via Corollary 3.1.4 through* $\mu_{\text{eff}}\|\nabla \boldsymbol{u}_h\|_0^2$*), on* $\sigma^{-1}$ *(when trying to control the terms via inverse estimates and* $\sigma\|\boldsymbol{u}_h\|_0^2$*), or on* $h^{-1}$ *(via integration by parts and incorporation into the pressure gradient term in the energy norm* (5.2.11a)*).*

**Remark 5.3.14** (On the Role of $\hat{\mu}$)**.**
*The scaling of the stabilization terms by* $\hat{\mu}$*, depending on the characteristic length* $\ell_{\boldsymbol{\Omega}}$*, allows to obtain stability estimates independent of the physical parameters. An element-dependent scaling* $\hat{\mu}_{\boldsymbol{T}} := \mu_{\text{eff}} + \sigma h_{\boldsymbol{T}}^2$ *(a suitable alternative for the case of essential boundary conditions, see, e.g., [BC09]) instead of* $\hat{\mu}$ *for the grad-div stabilization (and thus an element-wise divergence norm in* (5.2.11a)*) would not allow to uniformly bound the term* $\left(\nabla \cdot \boldsymbol{u}_h, \hat{\mu}\, q_h^{\boldsymbol{u}_h}\right)$ *in the proof of Lemma 5.3.11, since* $\hat{\mu} \geqslant \hat{\mu}_{\boldsymbol{T}}$ *and not the opposite. Note that a function* $q_h$ *of the form* $q_h|_{\boldsymbol{T}} := \hat{\mu}_{\boldsymbol{T}}\, q_h^{\boldsymbol{u}_h}|_{\boldsymbol{T}}$ *is in general not continuous and thus cannot be used as a pressure test function.*

**Remark 5.3.15** (On the Role of the Corner Stabilization)**.**
*The Lemmata 5.3.10 and 5.3.11 highlight the role of the corner stabilization. This stabilization is required here in order to allow control on the normal velocity at the boundary through the integral* $\langle q_h \boldsymbol{n}, \boldsymbol{u}_h \rangle$*.*

*Since* $\boldsymbol{n}$ *(and thus also the normal velocity* $\boldsymbol{u}_h \cdot \boldsymbol{n}$*) is discontinuous at corners,* $q_h^{\boldsymbol{u}_h}$ *of the form* $q_h^{\boldsymbol{u}_h}|_{\boldsymbol{T}} := (\nabla \cdot \boldsymbol{u}_h)\,|_{\boldsymbol{T}}$ *such that* $\left(q_h^{\boldsymbol{u}_h} \cdot \boldsymbol{n}\right)|_{E \in \mathcal{G}_h} = (\boldsymbol{u}_h \cdot \boldsymbol{n})\,|_{E \in \mathcal{G}_h}$ *- which would yield the desired control in Lemma 5.3.10 in a straightforward manner - is discontinuous and thus not allowed as test function in* $Q_h^1$*.*

*We observe that the corner stabilization is only required to prove stability in the Darcy limit (*$\mu_{\text{eff}} = 0$*). In fact, restricting to the case of a positive effective viscosity, Lemma 5.3.7 is sufficient to assure the required control on the velocity at the boundary. In other words, the Lemmata 5.3.10 and 5.3.11 are important for the stability of the penalty-free non-symmetric Nitsche method for the dual Darcy problem with continuous, element-wise linear finite element spaces.*

### Inf-Sup Stability

The previously proven lemmata allow to prove inf-sup stability of the bilinear form (5.2.10b), which is stated in the following theorem.

**Theorem 5.3.16** (Inf-Sup Stability)**.**
*Let* $\alpha, \delta, \rho > 0$ *and* $\mu_{\text{eff}}, \sigma \geqslant 0$ *with* $\mu_{\text{eff}} + \sigma > 0$*, and let us assume that the family of triangulations* $\{\mathcal{T}_h\}_{h>0}$ *fulfills the assumptions* $\mathbf{M}_1$*,* $\mathbf{M}_2$ *(*(5.2.2)*,* (5.2.4)*), and* $\mathbf{M}_3$*. Then there exists a constant* $\beta_h > 0$*, independent of the physical*

*parameters and of $h$, such that*

$$\inf_{(\boldsymbol{u}_h,p_h)\in \boldsymbol{V}_h^1\times Q_h^1\setminus\{(\boldsymbol{0},0)\}} \left( \sup_{(\boldsymbol{v}_h,q_h)\in \boldsymbol{V}_h^1\times Q_h^1\setminus\{(\boldsymbol{0},0)\}} \left( \frac{\mathcal{A}_h\left[(\boldsymbol{u}_h,p_h)\,;(\boldsymbol{v}_h,q_h)\right]}{|||(\boldsymbol{u}_h,p_h)|||_h |||(\boldsymbol{v}_h,q_h)|||_h} \right) \right) \geqslant \beta_h.$$

*Moreover,* $\beta_h^{-1} = O\left(\alpha^{-1}\left(\alpha^{-1}+(\delta+\rho)^{-1}\right)\right)$ *for* $\alpha,\delta,\rho \ll 1$.

*Proof:* Let $(\boldsymbol{u}_h,p_h)\in \boldsymbol{V}_h^1\times Q_h^1$. For the sake of simplicity, let us introduce the following notation:

$$\begin{aligned}
\xi_0 &:= |||(\boldsymbol{u}_h,p_h)|||_h^2 - \frac{\|p_h\|_0^2}{\hat{\mu}} - \sum_{E\in\mathcal{G}_h} \theta\,\frac{\mu_{\text{eff}}}{h_E}\|\boldsymbol{u}_h\|_{0,E}^2 - \sum_{E\in\mathcal{G}_h}\frac{\hat{\mu}}{h_E}\|\boldsymbol{u}_h\cdot\boldsymbol{n}_E\|_{0,E}^2,\\
\xi_1 &:= \frac{\|p_h\|_0^2}{\hat{\mu}},\\
\xi_2 &:= \sum_{E\in\mathcal{G}_h} \theta\,\frac{\mu_{\text{eff}}}{h_E}\|\boldsymbol{u}_h\|_{0,E}^2,\\
\xi_3 &:= \sum_{E\in\mathcal{G}_h}\frac{\hat{\mu}}{h_E}\|\boldsymbol{u}_h\cdot\boldsymbol{n}_E\|_{0,E}^2,\\
\xi_4 &:= \rho\,\hat{\mu}\sum_{\boldsymbol{x}\in\mathcal{C}} \left|[\![\boldsymbol{u}_h\cdot\boldsymbol{n}]\!](\boldsymbol{x})\right|^2,
\end{aligned}$$

such that $\xi_0,\xi_1,\xi_2,\xi_3,\xi_4 \geqslant 0$ and $|||(\boldsymbol{u}_h,p_h)|||_h^2 = \xi_0+\xi_1+\xi_2+\xi_3$ (see (5.2.11a)). Now we can rewrite and summarize the estimates proven in Lemmata 5.3.2, 5.3.4 (with $\boldsymbol{v}_h$ denoting the test function (5.3.3)), 5.3.7, and 5.3.11 as

$$\begin{aligned}
&\mathcal{A}_h\left[(\boldsymbol{u}_h,p_h)\,;C_0^{-1}(\boldsymbol{u}_h,p_h)\right] \geqslant \xi_0,\\
&\mathcal{A}_h\left[(\boldsymbol{u}_h,p_h)\,;2(\boldsymbol{v}_h,0)\right] \geqslant \xi_1 - 2C_1(\xi_0+\xi_2-\xi_4)\\
&\qquad\qquad \geqslant \xi_1 - 2C_1(\xi_0+\xi_2) \geqslant \xi_1 - 2C_1(\xi_0+\xi_2+\xi_3),\\
&\mathcal{A}_h\left[(\boldsymbol{u}_h,p_h)\,;4c_0^{-1}\left(\theta\boldsymbol{w}_h^{\boldsymbol{u}_h},0\right)\right] \geqslant \xi_2 - \frac{4C_2}{c_0}(\xi_0-\xi_4) \geqslant \xi_2 - \frac{4C_2}{c_0}\xi_0,\\
&\mathcal{A}_h\left[(\boldsymbol{u}_h,p_h)\,;4\left(\boldsymbol{0},\hat{\mu}\,q_h^{\boldsymbol{u}_h}\right)\right] \geqslant \xi_3 - 4C_3\xi_0.
\end{aligned}$$

Summing up the last two inequalities leads to

$$\mathcal{A}_h\left[(\boldsymbol{u}_h,p_h)\,;\left(4c_0^{-1}\theta\boldsymbol{w}_h^{\boldsymbol{u}_h},4\hat{\mu}\,q_h^{\boldsymbol{u}_h}\right)\right] \geqslant \xi_2+\xi_3-\widehat{C}_2\xi_0,$$

where $\widehat{C}_2 := \frac{4C_2}{c_0}+4C_3$. Consider first a test function $\left(\boldsymbol{z}_h^1,r_h^1\right)\in \boldsymbol{V}_h^1\times Q_h^1$ of the form

$$\left(\boldsymbol{z}_h^1,r_h^1\right) := (1-\eta_1)\,C_0^{-1}(\boldsymbol{u}_h,p_h) + \eta_1\left(4c_0^{-1}\theta\boldsymbol{w}_h^{\boldsymbol{u}_h},4\hat{\mu}\,q_h^{\boldsymbol{u}_h}\right),$$

depending on a parameter $\eta_1\in(0,1)$ (such that both coefficients are strictly positive) which will be determined later. Then it holds

$$\mathcal{A}_h\left[(\boldsymbol{u}_h,p_h)\,;\left(\boldsymbol{z}_h^1,r_h^1\right)\right] \geqslant \left(1-\eta_1-\widehat{C}_2\eta_1\right)\xi_0 + \eta_1(\xi_2+\xi_3).$$

Hence, defining

$$\eta_1 := \frac{1}{\widehat{C}_2 + 2} \in \left(0, \frac{1}{2}\right)$$

gives $1 - \eta_1 - \widehat{C}_2\eta_1 = \eta_1$ and thus

$$\mathcal{A}_h\left[(\boldsymbol{u}_h, p_h); (\boldsymbol{z}_h^1, r_h^1)\right] \geqslant \frac{1}{\widehat{C}_2 + 2}\left(\xi_0 + \xi_2 + \xi_3\right).$$

Next, consider a test function $(\boldsymbol{z}_h^2, r_h^2) \in \boldsymbol{V}_h^1 \times Q_h^1$ of the form

$$(\boldsymbol{z}_h^2, r_h^2) := (1 - \eta_2)\left(\widehat{C}_2 + 2\right)(\boldsymbol{z}_h^1, r_h^1) + \eta_2\,(2\boldsymbol{v}_h, 0),$$

depending on a parameter $\eta_2 \in (0,1)$ to be determined later. This yields

$$\mathcal{A}_h\left[(\boldsymbol{u}_h, p_h); (\boldsymbol{z}_h^2, r_h^2)\right] \geqslant (1 - \eta_2 - 2C_1\eta_2)\left(\xi_0 + \xi_2 + \xi_3\right) + \eta_2\xi_1.$$

Therefore, the choice

$$\eta_2 := \frac{1}{2C_1 + 2} \in \left(0, \frac{1}{2}\right)$$

gives $1 - \eta_2 - 2C_1\eta_2 = \eta_2$ and leads to

$$\mathcal{A}_h\left[(\boldsymbol{u}_h, p_h); (\boldsymbol{z}_h^2, r_h^2)\right] \geqslant \frac{1}{2C_1 + 2}\left(\xi_0 + \xi_1 + \xi_2 + \xi_3\right) = \frac{1}{2C_1 + 2}|||(\boldsymbol{u}_h, p_h)|||_h^2.$$

It remains to control the norm of the above constructed test function $(\boldsymbol{z}_h^2, r_h^2)$. From Corollary 3.1.4, the definition of $\boldsymbol{v}_h$ in (5.3.3), its properties stated in (5.3.4), and (5.2.7) we have

$$|||(\boldsymbol{v}_h, 0)|||_h^2 = \mu_{\text{eff}}\|\nabla\boldsymbol{v}_h\|_0^2 + \sigma\|\boldsymbol{v}_h\|_0^2 + \delta\,\hat{\mu}\|\nabla\cdot\boldsymbol{v}_h\|_0^2 \leqslant c_{\boldsymbol{\Omega}}^2\left(1 + N\delta\right)\frac{\|p_h\|_0^2}{\hat{\mu}}.$$

Moreover, from Lemma 5.3.6 and using $\theta \leqslant 1$ and $\theta\hat{\mu} = \mu_{\text{eff}}$, we infer

$$\left|\left|\left|\left(\theta\boldsymbol{w}_h^{\boldsymbol{u}_h}, 0\right)\right|\right|\right|_h^2 \leqslant \left(c_3^2 + c_2 + \delta N c_2 + 1\right)|||(\boldsymbol{u}_h, p_h)|||_h^2,$$

and (5.3.31) and (5.3.32) result in

$$\begin{aligned}\left|\left|\left|\left(\boldsymbol{0}, \hat{\mu}\, q_h^{\boldsymbol{u}_h}\right)\right|\right|\right|_h^2 &= \hat{\mu}^2\frac{\left\|q_h^{\boldsymbol{u}_h}\right\|_0^2}{\hat{\mu}} + \hat{\mu}^2\alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\left\|\nabla q_h^{\boldsymbol{u}_h}\right\|_{0,\boldsymbol{T}}^2 \\ &\leqslant c_6\left(1 + \alpha c_I^2\right)|||(\boldsymbol{u}_h, p_h)|||_h^2.\end{aligned}$$

Hence, we can estimate with the triangle inequality, Corollary 3.1.3, and since $0 < \eta_1 < \frac{1}{2}$ and $0 < \eta_2 < \frac{1}{2}$,

$$\begin{aligned}\left|\left|\left|\left(\boldsymbol{z}_h^1, r_h^1\right)\right|\right|\right|_h^2 &\leqslant 2\left(1 - \eta_1\right)^2 C_0^{-2}|||(\boldsymbol{u}_h, p_h)|||_h^2 \\ &\quad + 2\eta_1^2\left(32c_0^{-2}\left|\left|\left|\left(\theta\boldsymbol{w}_h^{\boldsymbol{u}_h}, 0\right)\right|\right|\right|_h^2 + 32\left|\left|\left|\left(\boldsymbol{0}, \hat{\mu}\, q_h^{\boldsymbol{u}_h}\right)\right|\right|\right|_h^2\right) \\ &\leqslant \left(2C_0^{-2} + 16\left(\frac{\left(c_3^2 + c_2 + \delta N c_2 + 1\right)}{c_0^2} + c_6\left(1 + \alpha c_I^2\right)\right)\right)|||(\boldsymbol{u}_h, p_h)|||_h^2,\end{aligned}$$

$$\begin{aligned}\left|\!\left|\!\left|\left(\boldsymbol{z}_h^2, r_h^2\right)\right|\!\right|\!\right|_h^2 &\leqslant 2\eta_2^2 |\!|\!|(2\boldsymbol{v}_h, 0)|\!|\!|_h^2 + 2\left(1-\eta_2\right)^2 \left(\widehat{C}_2+2\right)^2 \left|\!\left|\!\left|\left(\boldsymbol{z}_h^1, r_h^1\right)\right|\!\right|\!\right|_h^2 \\ &\leqslant \left(2c_{\boldsymbol{\Omega}}^2\left(1+N\delta\right) + 4\left(\widehat{C}_2+2\right)^2 \left(C_0^{-2}\right.\right. \\ &\quad \left.\left. + \frac{8\left(c_3^2+c_2+\delta N c_2+1\right)}{c_0^2} + 8c_6\left(1+\alpha c_I^2\right)\right)\right) |\!|\!|(\boldsymbol{u}_h, p_h)|\!|\!|_h^2.\end{aligned}$$

which allows to conclude

$$\mathcal{A}_h\left[(\boldsymbol{u}_h, p_h)\,;\left(\boldsymbol{z}_h^2, r_h^2\right)\right] \geqslant \beta_h\, |\!|\!|(\boldsymbol{u}_h, p_h)|\!|\!|_h \left|\!\left|\!\left|\left(\boldsymbol{z}_h^2, r_h^2\right)\right|\!\right|\!\right|_h$$

with

$$\begin{aligned}\beta_h = (2C_1+2)^{-1} \Bigg[ & 2c_{\boldsymbol{\Omega}}^2\left(1+N\delta\right) \\ & + 4\left(\widehat{C}_2+2\right)^2 \left(C_0^{-2} + \frac{8\left(c_3^2+c_2+\delta N c_2+1\right)}{c_0^2} + 8c_6\left(1+\alpha c_I^2\right)\right)\Bigg]^{-\frac{1}{2}}.\end{aligned}$$

The behavior for small values of stabilization parameters follows from $C_0 = O(1)$, $C_1 = O\left(\alpha^{-1}\right)$, $C_2 = O(\alpha^{-1})$, and $C_3 = O\left(\delta^{-1}+\rho^{-1}\right)$.

□

**Corollary 5.3.17** (Existence of a Unique Solution)**.**
*The discrete problem* (5.2.10) *has a unique solution (is well-posed).*

*Proof:* This is a direct consequence of the inf-sup stability (Thm. 5.3.16), as described in Theorem 4.1.1. □

### 5.3.2 A Priori Error Analysis

Based on the well-posedness result (Thm. 5.3.16 and Cor. 5.3.17), the approximation error of the finite element method 5.2.10 can be analyzed. Therefore, we will discuss the consistency, the Galerkin orthogonality, approximability properties, and finally derive an error estimate predicting the convergence order.

#### Consistency and Galerkin Orthogonality

Firstly, let us observe that the proposed stabilized penalty-free non-symmetric Nitsche method is consistent and satisfies the Galerkin orthogonality property.

**Lemma 5.3.18** (Consistency)**.**
*The stabilized, penalty-free non-symmetric Nitsche formulation* (5.2.10) *is consistent with the original problem* (3.2.8)*, i.e., for the solution* $(\boldsymbol{u}, p) \in \boldsymbol{H}^2(\boldsymbol{\Omega}) \times H^1(\boldsymbol{\Omega})$ *of* (3.2.8) *it holds*

$$\mathcal{A}_h\left[(\boldsymbol{u}, p)\,;(\boldsymbol{v}_h, q_h)\right] = \mathcal{L}_h\left[(\boldsymbol{v}_h, q_h)\right], \quad \forall\,(\boldsymbol{v}_h, q_h) \in \boldsymbol{V}_h^1 \times Q_h^1.$$

*Proof:* Let $(\boldsymbol{u}, p) \in \boldsymbol{H}^2(\boldsymbol{\Omega}) \times H^1(\boldsymbol{\Omega})$ be the solution of (3.2.8). Integration by parts can be used to obtain

$$(\mu_{\text{eff}} \nabla \boldsymbol{u}, \nabla \boldsymbol{v}_h) - (p, \nabla \cdot \boldsymbol{v}_h) - \langle (\mu_{\text{eff}} \nabla \boldsymbol{u} - p\mathbb{I}) \cdot \boldsymbol{n}, \boldsymbol{v}_h \rangle = (-\mu_{\text{eff}} \Delta \boldsymbol{u} + \nabla p, \boldsymbol{v}_h).$$

Further using the momentum balance and the divergence constraint of the Brinkman problem (3.2.8), combined with the consistency of the Nitsche method and the stabilizations (see Subsection 4.3.2 and Remarks 5.2.3 and 5.2.4) we can show

$$\begin{aligned}
\mathcal{A}_h\left[(\boldsymbol{u}, p); (\boldsymbol{v}_h, q_h)\right] &= (-\mu_{\text{eff}} \Delta \boldsymbol{u} + \nabla p + \sigma \boldsymbol{u}, \boldsymbol{v}_h) + (\nabla \cdot \boldsymbol{u}, q_h) \\
&\quad + \langle (\mu_{\text{eff}} \nabla \boldsymbol{v}_h - q_h \mathbb{I}) \cdot \boldsymbol{n}, \boldsymbol{u} \rangle + S_{h,\alpha}^{GLS,lhs}\left[(\boldsymbol{u}, p); (\boldsymbol{v}_h, q_h)\right] \\
&\quad + S_{h,\delta}^{GD,lhs}\left[(\boldsymbol{u}, \boldsymbol{v}_h)\right] + S_{h,\rho}^{C,lhs}\left[(\boldsymbol{u}, \boldsymbol{v}_h)\right] \\
&= (\boldsymbol{f}, \boldsymbol{v}_h) + (g, q_h) + \langle (\mu_{\text{eff}} \nabla \boldsymbol{v}_h - q_h \mathbb{I}) \cdot \boldsymbol{n}, \boldsymbol{0} \rangle \\
&\quad + S_{h,\alpha}^{GLS,rhs}\left[(\boldsymbol{v}_h, q_h)\right] + S_{h,\delta}^{GD,rhs}\left[\boldsymbol{v}_h\right] + S_{h,\rho}^{C,rhs}\left[\boldsymbol{v}_h\right] \\
&= \mathcal{L}_h\left[(\boldsymbol{v}_h, q_h)\right], \qquad \forall\, (\boldsymbol{v}_h, q_h) \in \boldsymbol{V}_h^1 \times Q_h^1.
\end{aligned}$$

In the last equality, the boundary conditions of problem (3.2.8) were used. □

**Lemma 5.3.19** (Galerkin Orthogonality)**.**
*Assume that* $(\boldsymbol{u}, p) \in \boldsymbol{H}^2(\boldsymbol{\Omega}) \times H^1(\Omega)$ *solves* (3.2.8) *and let* $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h^1 \times Q_h^1$ *be the solution of* (5.2.10a)*. Then,*

$$\mathcal{A}_h\left[(\boldsymbol{u} - \boldsymbol{u}_h, p - p_h); (\boldsymbol{v}_h, q_h)\right] = 0, \quad \forall\, (\boldsymbol{v}_h, q_h) \in \boldsymbol{V}_h^1 \times Q_h^1.$$

*Proof:* The result follows from the consistency of the discrete formulation (Lem. 5.3.18) and the conformity of the triangulation. □

### Approximability

The next lemma is related to the quality of approximations in $\boldsymbol{V}_h^1/Q_h^1$ with respect to the mesh-dependent norm.

**Lemma 5.3.20** (Approximability)**.**
*Let* $I_{h,1}^{\text{La}}(\cdot)$ *be the Lagrange interpolation operator onto* $\boldsymbol{V}_h^1$ *and* $I_{h,1}^{\text{SZ}^{f^0}}(\cdot)$ *the modified Scott–Zhang interpolation operator onto* $Q_h^1$ *(see Section 5.1). Then for* $(\boldsymbol{u}, p) \in \boldsymbol{H}^2(\boldsymbol{\Omega}) \times H^1(\boldsymbol{\Omega})$ *and* $\sigma, \mu_{\text{eff}} \geqslant 0$ *with* $\mu_{\text{eff}} + \sigma > 0$*, it holds*

$$\begin{aligned}
\left|\!\left|\!\left| \left( \boldsymbol{u} - I_{h,1}^{\text{La}}(\boldsymbol{u}), p - I_{h,1}^{\text{SZ}^{f^0}}(p) \right) \right|\!\right|\!\right|_h^2 \leqslant h^2 \Big[ & c_{\text{La}}^2 \hat{\mu} \left(1 + N\delta + 2c_{\text{DT}}\right) |\boldsymbol{u}|_2^2 \\
& + \tilde{c}_{\text{SZ}}^2 \frac{1}{\hat{\mu}} \left(1 + \alpha\right) |p|_1^2 \Big],
\end{aligned} \tag{5.3.40}$$

*where* $c_{\text{La}}$ *and* $\bar{c}_{\text{SZ}}$ *in* $\tilde{c}_{\text{SZ}}^2 = 2\bar{c}_{\text{SZ}}^2 \max_{\boldsymbol{T} \in \mathcal{T}_h} \{\#S(\boldsymbol{T})\}$ *are the constants defined in* (5.1.1) *and* (5.1.2b)*.*

*Proof:* We start by estimating the bulk terms of the norm $|||(\cdot,\cdot)|||_h^2$ using Corollary 3.1.4 and the properties of the interpolation operators (5.1.1) and (5.1.2b) (with $l=1$ and $m \in \{0,1\}$):

$$\begin{aligned}
&\mu_{\text{eff}}\left\|\nabla\left(\boldsymbol{u}-I_{h,1}^{\text{La}}(\boldsymbol{u})\right)\right\|_0^2+\sigma\left\|\boldsymbol{u}-I_{h,1}^{\text{La}}(\boldsymbol{u})\right\|_0^2+\delta\,\hat{\mu}\left\|\nabla\cdot\left(\boldsymbol{u}-I_{h,1}^{\text{La}}(\boldsymbol{u})\right)\right\|_0^2\\
&\qquad+\frac{1}{\hat{\mu}}\left\|p-I_{h,1}^{\text{SZ}^{\int 0}}(p)\right\|_0^2+\alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\left\|\nabla\left(p-I_{h,1}^{\text{SZ}^{\int 0}}(p)\right)\right\|_{0,\boldsymbol{T}}^2\\
&\qquad\leqslant \mu_{\text{eff}}c_{\text{La}}^2h^2\,|\boldsymbol{u}|_2^2+\sigma c_{\text{La}}^2h^4\,|\boldsymbol{u}|_2^2+N\delta\,\hat{\mu}c_{\text{La}}^2h^2\,|\boldsymbol{u}|_2^2\\
&\qquad\quad+\frac{1}{\hat{\mu}}\tilde{c}_{\text{SZ}}^2h^2\,|p|_1^2+\frac{\alpha}{\hat{\mu}}\tilde{c}_{\text{SZ}}^2h^2\,|p|_1^2\\
&\qquad\leqslant h^2\left[c_{\text{La}}^2\left(\underbrace{\mu_{\text{eff}}+\sigma h^2}_{\leqslant\hat{\mu}}+N\delta\,\hat{\mu}\right)|\boldsymbol{u}|_2^2+\tilde{c}_{\text{SZ}}^2\left(\frac{1}{\hat{\mu}}(1+\alpha)\right)|p|_1^2\right],
\end{aligned}$$

where $\tilde{c}_{\text{SZ}}^2 := 2\bar{c}_{\text{SZ}}^2\max_{\boldsymbol{T}\in\mathcal{T}_h}\{\#S(\boldsymbol{T})\}$. For the additional boundary terms related to the penalty-free non-symmetric Nitsche method we get with (4.1.5) and $\theta \leqslant 1$:

$$\begin{aligned}
&\sum_{E\in\mathcal{G}_h}\theta\,\frac{\mu_{\text{eff}}}{h_E}\left\|\boldsymbol{u}-I_{h,1}^{\text{La}}(\boldsymbol{u})\right\|_{0,E}^2+\sum_{E\in\mathcal{G}_h}\frac{\hat{\mu}}{h_E}\left\|\left(\boldsymbol{u}-I_{h,1}^{\text{La}}(\boldsymbol{u})\right)\cdot\boldsymbol{n}_E\right\|_{0,E}^2\\
&\qquad\leqslant c_{\text{DT}}\left(\mu_{\text{eff}}+\hat{\mu}\right)\sum_{\substack{\boldsymbol{T}_E\in\mathcal{T}_h\\E\in\mathcal{G}_h}}\left(h_{\boldsymbol{T}_E}^{-2}\left\|\boldsymbol{u}-I_{h,1}^{\text{La}}(\boldsymbol{u})\right\|_{0,\boldsymbol{T}_E}^2+\left\|\nabla\left(\boldsymbol{u}-I_{h,1}^{\text{La}}(\boldsymbol{u})\right)\right\|_{0,\boldsymbol{T}_E}^2\right)\\
&\qquad\leqslant c_{\text{DT}}\left(\mu_{\text{eff}}+\hat{\mu}\right)c_{\text{La}}^2h^2\;|\boldsymbol{u}|_2^2\,.
\end{aligned}$$

Finally, we observe that the interpolation error due to the corner stabilization term vanishes since the Lagrange interpolator is exact on mesh nodes. The inequality (5.3.40) is obtained summing up the above estimates and observing that $\mu_{\text{eff}}, \mu_{\text{eff}}+\sigma h^2 \leqslant \hat{\mu}$. □

### Error Estimate

Combining the previous results, we can derive an error estimate for the proposed method.

**Theorem 5.3.21** (A Priori Error Estimate)**.**
*Let $(\boldsymbol{u},p)$ be the solution of (3.2.8) and $(\boldsymbol{u}_h,p_h)$ be the solution of problem (5.2.10), where $\alpha,\delta,\rho>0$ and $\mu_{\text{eff}},\sigma \geqslant 0$ with $\mu_{\text{eff}}+\sigma>0$. Then for $(\boldsymbol{u},p)\in \boldsymbol{H}^2(\boldsymbol{\Omega})\times H^1(\boldsymbol{\Omega})$, it holds*

$$|||(\boldsymbol{u}-\boldsymbol{u}_h,p-p_h)|||_h \leqslant h\left(c_{\text{La}}C_{\boldsymbol{u}}\;|\boldsymbol{u}|_2+\tilde{c}_{\text{SZ}}C_p\;|p|_1\right)$$

*with*

$$\begin{aligned}
C_{\boldsymbol{u}} := \hat{\mu}^{\frac{1}{2}}\Big[&(1+N\delta+2c_{\text{DTI}})^{\frac{1}{2}}\\
&+\frac{1}{\beta_h}\left((2c_{\text{DT}}\left(c_{\text{DTI}}+\widehat{c}_{\text{DTI}}\right))^{\frac{1}{2}}+N^{\frac{1}{2}}\left(\delta^{\frac{1}{2}}+1\right)+1+(2c_{\text{DT}})^{\frac{1}{2}}+2\alpha^{\frac{1}{2}}+2\alpha\right)\Big]
\end{aligned}$$

*and*

$$C_p := \hat{\mu}^{-\frac{1}{2}} \left[ (1+\alpha)^{\frac{1}{2}} + \frac{1}{\beta_h} \left( (2c_{\mathrm{DT}})^{\frac{1}{2}} + \alpha^{\frac{1}{2}} + \alpha + \delta^{-\frac{1}{2}} \right) \right].$$

*Proof:* The proof is based on the combination of the inf-sup condition, the Galerkin orthogonality, Lemma 5.3.20, and the approximation properties of the interpolation operators. Let us consider the Lagrange respectively Scott–Zhang interpolants of the solution, $\boldsymbol{v}_h := I_{h,1}^{\mathrm{La}}(\boldsymbol{u})$ and $q_h := I_{h,1}^{\mathrm{SZ}^{\int^0}}(p)$, and decompose the error as

$$|||(\boldsymbol{u}-\boldsymbol{u}_h, p-p_h)|||_h \leqslant |||(\boldsymbol{u}-\boldsymbol{v}_h, p-q_h)|||_h + |||(\boldsymbol{v}_h-\boldsymbol{u}_h, q_h-p_h)|||_h. \quad (5.3.41)$$

Exploiting the inf-sup stability (Thm. 5.3.16) and the Galerkin orthogonality (Lem. 5.3.19), it holds:

$$|||(\boldsymbol{u}-\boldsymbol{u}_h, p-p_h)|||_h \leqslant |||(\boldsymbol{u}-\boldsymbol{v}_h, p-q_h)|||_h + \frac{1}{\beta_h} \sup_{(\boldsymbol{w}_h, r_h) \in \boldsymbol{V}_h^1 \times Q_h^1 \setminus \{(\boldsymbol{0},0)\}} \frac{|\mathcal{A}_h[(\boldsymbol{u}-\boldsymbol{v}_h, p-q_h); (\boldsymbol{w}_h, r_h)]|}{|||(\boldsymbol{w}_h, r_h)|||_h}.$$

Next, we bound $|\mathcal{A}_h[(\boldsymbol{u}-\boldsymbol{v}_h, p-q_h); (\boldsymbol{w}_h, r_h)]|$. For the bulk terms we obtain, using the Cauchy–Schwarz inequality (Thm. 3.1.5), Corollary 3.1.4, and the properties (5.1.1) and (5.1.2b) (with $l=1$ and $m=0$) of the interpolation operators:

$$\begin{aligned}
|\mu_{\mathrm{eff}}(\nabla(\boldsymbol{u}-\boldsymbol{v}_h), \nabla \boldsymbol{w}_h)| &\leqslant \mu_{\mathrm{eff}}^{\frac{1}{2}} \|\nabla(\boldsymbol{u}-\boldsymbol{v}_h)\|_0 \, \mu_{\mathrm{eff}}^{\frac{1}{2}} \|\nabla \boldsymbol{w}_h\|_0 \\
&\leqslant c_{\mathrm{La}} \mu_{\mathrm{eff}}^{\frac{1}{2}} h \, |\boldsymbol{u}|_2 \, |||(\boldsymbol{w}_h, r_h)|||_h, \\
|\sigma(\boldsymbol{u}-\boldsymbol{v}_h, \boldsymbol{w}_h)| &\leqslant \sigma^{\frac{1}{2}} \|\boldsymbol{u}-\boldsymbol{v}_h\|_0 \, \sigma^{\frac{1}{2}} \|\boldsymbol{w}_h\|_0 \\
&\leqslant c_{\mathrm{La}} \sigma^{\frac{1}{2}} h^2 \, |\boldsymbol{u}|_2 \, |||(\boldsymbol{w}_h, r_h)|||_h, \\
|-(p-q_h, \nabla \cdot \boldsymbol{w}_h)| &\leqslant (\delta\,\hat{\mu})^{-\frac{1}{2}} \|p-q_h\|_0 \, (\delta\,\hat{\mu})^{\frac{1}{2}} \|\nabla \cdot \boldsymbol{w}_h\|_0 \\
&\leqslant \tilde{c}_{\mathrm{SZ}} (\delta\,\hat{\mu})^{-\frac{1}{2}} h \, |p|_1 \, |||(\boldsymbol{w}_h, r_h)|||_h, \\
|(\nabla \cdot (\boldsymbol{u}-\boldsymbol{v}_h), r_h)| &\leqslant \hat{\mu}^{\frac{1}{2}} \|\nabla \cdot (\boldsymbol{u}-\boldsymbol{v}_h)\|_0 \hat{\mu}^{-\frac{1}{2}} \|r_h\|_0 \\
&\leqslant c_{\mathrm{La}} (\hat{\mu} N)^{\frac{1}{2}} h \, |\boldsymbol{u}|_2 \, |||(\boldsymbol{w}_h, r_h)|||_h,
\end{aligned}$$

with $\tilde{c}_{\mathrm{SZ}}^2 := 2\bar{c}_{\mathrm{SZ}}^2 \max_{\boldsymbol{T} \in \mathcal{T}_h} \{\#S(\boldsymbol{T})\}$. The GLS stabilization terms can be estimated using additionally the Cauchy–Schwarz inequality for sums (Lem. 3.1.2) and

the estimates (5.2.7), yielding

$$\left|\alpha \sum_{\boldsymbol{T}\in\mathcal{T}_h} \frac{h_{\boldsymbol{T}}^2}{\hat{\mu}} \left(\sigma\left(\boldsymbol{u}-\boldsymbol{v}_h\right), \sigma \boldsymbol{w}_h\right)_{\boldsymbol{T}}\right| \leqslant c_{\mathrm{La}} \alpha \sigma^{\frac{1}{2}} h^2 |\boldsymbol{u}|_2 \; |\!|\!| (\boldsymbol{w}_h, r_h) |\!|\!|_h,$$

$$\left|\alpha \sum_{\boldsymbol{T}\in\mathcal{T}_h} \frac{h_{\boldsymbol{T}}^2}{\hat{\mu}} \left(\sigma\left(\boldsymbol{u}-\boldsymbol{v}_h\right), \nabla r_h\right)_{\boldsymbol{T}}\right| \leqslant c_{\mathrm{La}} \left(\alpha\sigma\right)^{\frac{1}{2}} h^2 |\boldsymbol{u}|_2 \; |\!|\!| (\boldsymbol{w}_h, r_h) |\!|\!|_h,$$

$$\left|\alpha \sum_{\boldsymbol{T}\in\mathcal{T}_h} \frac{h_{\boldsymbol{T}}^2}{\hat{\mu}} \left(\nabla\left(p-q_h\right), \sigma \boldsymbol{w}_h\right)_{\boldsymbol{T}}\right| \leqslant \tilde{c}_{\mathrm{SZ}} \alpha \hat{\mu}^{-\frac{1}{2}} h |p|_1 \; |\!|\!| (\boldsymbol{w}_h, r_h) |\!|\!|_h,$$

$$\left|\alpha \sum_{\boldsymbol{T}\in\mathcal{T}_h} \frac{h_{\boldsymbol{T}}^2}{\hat{\mu}} \left(\nabla\left(p-q_h\right), \nabla r_h\right)_{\boldsymbol{T}}\right| \leqslant \tilde{c}_{\mathrm{SZ}} \alpha^{\frac{1}{2}} \hat{\mu}^{-\frac{1}{2}} h |p|_1 \; |\!|\!| (\boldsymbol{w}_h, r_h) |\!|\!|_h,$$

$$\left|\alpha \sum_{\boldsymbol{T}\in\mathcal{T}_h} \frac{h_{\boldsymbol{T}}^2}{\hat{\mu}} \left(-\mu_{\mathrm{eff}} \Delta\left(\boldsymbol{u}-\boldsymbol{v}_h\right), \sigma \boldsymbol{w}_h\right)_{\boldsymbol{T}}\right| \leqslant c_{\mathrm{La}} \alpha \sigma^{\frac{1}{2}} h^2 |\boldsymbol{u}|_2 \; |\!|\!| (\boldsymbol{w}_h, r_h) |\!|\!|_h,$$

$$\left|\alpha \sum_{\boldsymbol{T}\in\mathcal{T}_h} \frac{h_{\boldsymbol{T}}^2}{\hat{\mu}} \left(-\mu_{\mathrm{eff}} \Delta\left(\boldsymbol{u}-\boldsymbol{v}_h\right), \nabla r_h\right)_{\boldsymbol{T}}\right| \leqslant c_{\mathrm{La}} \left(\alpha\mu_{\mathrm{eff}}\right)^{\frac{1}{2}} h |\boldsymbol{u}|_2 \; |\!|\!| (\boldsymbol{w}_h, r_h) |\!|\!|_h.$$

For the terms related to the grad-div stabilization, we similarly obtain

$$\begin{aligned}\left|\delta \sum_{\boldsymbol{T}\in\mathcal{T}_h} \hat{\mu} \left(\nabla\cdot\left(\boldsymbol{u}-\boldsymbol{v}_h\right), \nabla\cdot \boldsymbol{w}_h\right)_{\boldsymbol{T}}\right| &\leqslant \sum_{\boldsymbol{T}\in\mathcal{T}_h} (\delta\hat{\mu})^{\frac{1}{2}} \|\nabla\cdot\left(\boldsymbol{u}-\boldsymbol{v}_h\right)\|_{0,\boldsymbol{T}} \left(\delta\hat{\mu}\right)^{\frac{1}{2}} \|\nabla\cdot \boldsymbol{w}_h\|_{0,\boldsymbol{T}} \\ &\leqslant c_{\mathrm{La}} \left(N\delta\,\hat{\mu}\right)^{\frac{1}{2}} h |\boldsymbol{u}|_2 \, |\!|\!| (\boldsymbol{w}_h, r_h) |\!|\!|_h.\end{aligned}$$

The additional terms related to the penalty-free non-symmetric Nitsche method can be controlled using further (4.1.5), (4.1.6b), (4.1.6a), and $\theta\hat{\mu} = \mu_{\mathrm{eff}}$ as follows:

$$\begin{aligned}|\langle \mu_{\mathrm{eff}} \nabla(\boldsymbol{u}-\boldsymbol{v}_h)\cdot\boldsymbol{n}, \boldsymbol{w}_h\rangle| &\leqslant \frac{\mu_{\mathrm{eff}}^{\frac{1}{2}}}{\theta^{\frac{1}{2}}} \left(\sum_{E\in\mathcal{G}_h} h_E \|\nabla(\boldsymbol{u}-\boldsymbol{v}_h)\cdot\boldsymbol{n}_E\|_{0,E}^2\right)^{\frac{1}{2}} \left(\sum_{E\in\mathcal{G}_h} \theta \frac{\mu_{\mathrm{eff}}}{h_E} \|\boldsymbol{w}_h\|_{0,E}^2\right)^{\frac{1}{2}} \\ &\leqslant c_{\mathrm{La}} \left(2\hat{\mu}\, c_{\mathrm{DT}}\right)^{\frac{1}{2}} h |\boldsymbol{u}|_2 \, |\!|\!| (\boldsymbol{w}_h, r_h) |\!|\!|_h, \\ |\langle p-q_h, \boldsymbol{w}_h\cdot\boldsymbol{n}\rangle| &\leqslant \left(\sum_{E\in\mathcal{G}_h} \frac{h_E}{\hat{\mu}} \|p-q_h\|_{0,E}^2\right)^{\frac{1}{2}} \left(\sum_{E\in\mathcal{G}_h} \frac{\hat{\mu}}{h_E} \|\boldsymbol{w}_h\cdot\boldsymbol{n}_E\|_{0,E}^2\right)^{\frac{1}{2}} \\ &\leqslant \tilde{c}_{\mathrm{SZ}} \left(2c_{\mathrm{DT}}\right)^{\frac{1}{2}} \hat{\mu}^{-\frac{1}{2}} h |p|_1 \, |\!|\!| (\boldsymbol{w}_h, r_h) |\!|\!|_h,\end{aligned}$$

$$
\begin{aligned}
|\langle \mu_{\text{eff}} \nabla \boldsymbol{w}_h \cdot \boldsymbol{n}, \boldsymbol{u} - \boldsymbol{v}_h \rangle| &\leqslant \mu_{\text{eff}}^{\frac{1}{2}} \left( \sum_{E \in \mathcal{G}_h} \frac{1}{h_E} \|\boldsymbol{u} - \boldsymbol{v}_h\|_{0,E}^2 \right)^{\frac{1}{2}} \left( \sum_{E \in \mathcal{G}_h} h_E \mu_{\text{eff}} \|\nabla \boldsymbol{w}_h \cdot \boldsymbol{n}_E\|_{0,E}^2 \right)^{\frac{1}{2}} \\
&\leqslant (\mu_{\text{eff}} c_{\text{DT}} \widehat{c}_{\text{DTI}})^{\frac{1}{2}} \left( \sum_{\substack{\boldsymbol{T}_E \in \mathcal{T}_h \\ E \in \mathcal{G}_h}} h_{\boldsymbol{T}_E}^{-2} \left( \|\boldsymbol{u} - \boldsymbol{v}_h\|_{0,\boldsymbol{T}_E}^2 \right. \right. \\
&\qquad \left. \left. + h_{\boldsymbol{T}_E}^2 \|\nabla (\boldsymbol{u} - \boldsymbol{v}_h)\|_{0,\boldsymbol{T}_E}^2 \right) \right)^{\frac{1}{2}} \times \left( \sum_{\substack{\boldsymbol{T}_E \in \mathcal{T}_h \\ E \in \mathcal{G}_h}} \mu_{\text{eff}} \|\nabla \boldsymbol{w}_h\|_{0,\boldsymbol{T}_E}^2 \right)^{\frac{1}{2}} \\
&\leqslant c_{\text{La}} \left(2 \mu_{\text{eff}} c_{\text{DT}} \widehat{c}_{\text{DTI}}\right)^{\frac{1}{2}} h \, |\boldsymbol{u}|_2 \, |||(\boldsymbol{w}_h, r_h)|||_h, \\
|\langle (\boldsymbol{u} - \boldsymbol{v}_h) \cdot \boldsymbol{n}, r_h \rangle| &\leqslant \left( \sum_{E \in \mathcal{G}_h} \frac{1}{h_E} \|\boldsymbol{u} - \boldsymbol{v}_h\|_{0,E}^2 \right)^{\frac{1}{2}} \left( \sum_{E \in \mathcal{G}_h} h_E \|r_h\|_{0,E}^2 \right)^{\frac{1}{2}} \\
&\leqslant (c_{\text{DT}} c_{\text{DTI}} \hat{\mu})^{\frac{1}{2}} \left( \sum_{\substack{\boldsymbol{T}_E \in \mathcal{T}_h \\ E \in \mathcal{G}_h}} \left( h_{\boldsymbol{T}_E}^{-2} \|\boldsymbol{u} - \boldsymbol{v}_h\|_{0,\boldsymbol{T}_E}^2 \right. \right. \\
&\qquad \left. \left. + \|\nabla (\boldsymbol{u} - \boldsymbol{v}_h)\|_{0,\boldsymbol{T}_E}^2 \right) \right)^{\frac{1}{2}} \times \left( \frac{1}{\hat{\mu}} \sum_{\substack{\boldsymbol{T}_E \in \mathcal{T}_h \\ E \in \mathcal{G}_h}} \|r_h\|_{0,\boldsymbol{T}_E}^2 \right)^{\frac{1}{2}} \\
&\leqslant c_{\text{La}} \left(2 c_{\text{DT}} c_{\text{DTI}} \hat{\mu}\right)^{\frac{1}{2}} h \, |\boldsymbol{u}|_2 \, |||(\boldsymbol{w}_h, r_h)|||_h.
\end{aligned}
$$

The terms related to the corner stabilization vanish as the Lagrange interpolator is exact on mesh nodes. The proof is concluded summing up all the contributions, using (5.2.7), and Lemma 5.3.20 for the first summand in (5.3.41). □

## 5.4 Summary

In this chapter we have proposed a stabilized, penalty-free, non-symmetric Nitsche method for the Brinkman problem and have proven stability and optimal convergence in the energy norm for any combination of physical coefficients, including the Stokes limit and the Darcy limit. The robustness of the method is a consequence of its ingredients.

Linear equal-order interpolation is chosen with the objective to obtain a low-order routine with limited computational cost that can be easily implemented in existing finite element frameworks. This finite element pair necessitates (pressure-) stabilization in order assure stability, which is the role of the non-symmetric GLS method.

Based on that choice, we recover control of the divergence of the velocity in the Darcy limit via a grad-div stabilization which is also important to prove robust stability in the Darcy limit.

A parameter-driven transition of the essential boundary conditions is enabled by using the penalty-free non-symmetric Nitsche method, which weakly imposes these boundary conditions. Control on the velocity at the boundary in the Darcy limit concerns solely the normal velocity and is guaranteed by incorporating a corner stabilization.

For the proof of inf-sup stability, we have constructed a pair of test functions which yielded the desired control with an inf-sup constant, independent from the Brinkman coefficients $\mu_{\text{eff}}$ and $\sigma$ and from the mesh-size.

The error estimation assures convergence of order one (optimal) in the energy norm and shows a typical scaling in the estimate, concerning the velocity and pressure coefficients, by the physical coefficients (in form of $\hat{\mu}^{\frac{1}{2}}$ and $\hat{\mu}^{-\frac{1}{2}}$, respectively) of the problem.

# 6 Numerical Results

The goal of this chapter is to validate with numerical experiments the results of the analysis of the stabilized penalty-free non-symmetric Nitsche method (5.2.10) proposed in the previous chapter.

To that end, we will start with two examples, whose analytical solution is available and which comprise different physical parameter settings, in order to validate the robustness of the method. A convergence study for different values of the model parameters is performed and supplemented by studying the dependence of the condition number of the finite element coefficient matrix on the stabilization parameters.

The numerical study is enriched by simulation results for three more examples, which can be interpreted as schematically describing groundwater flow scenarios. These pose different additional challenges (partly not covered by the theory in Chapter 5), including heterogeneity with straight, perpendicular, or inclined fluid-porous interfaces (with respect to the background flow), non-square domains, and non-uniform, unstructured meshes. The exact solutions to the problems are not available such that the simulation results will be evaluated qualitatively and in particular discussed in comparison with results from the literature.

The third example will be used to demonstrate the crucial impact of the GLS stabilization parameter on the finite element solution and to compare the method proposed here with publicized results.

In the last two examples, the method (5.2.10) will be numerically investigated and partly compared with results from the literature, concerning its suitability as a single domain approach for regions with fluid-porous quasi-interfaces (without additional imposition of interface conditions).

The simulations were performed using the finite element library ParMooN, see [Gan+16] and [Wil+17a]. Linear systems of equations were solved directly using the external library UMFPACK, [Dav04]. Non-uniform and unstructured meshes were constructed with the free mesh generator Gmsh, [GR09], and Paraview, [AGL05], was used for the visualizations. For dimensional quantities we consider SI units if not otherwise stated. Parts of Sections 6.1 and 6.4 are published in [Bla+18].

## 6.1 Homogeneous Domains with Analytical Solutions – Parameter and Convergence Study

The proposed method (5.2.10) will be investigated for two test problems with known analytical solution in a homogeneous porous medium.

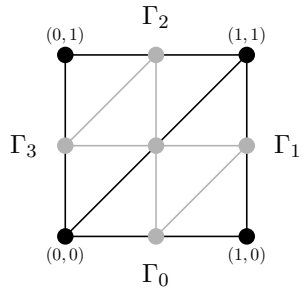
In order to validate the robustness of the formulation in practice, we consider firstly an example with $\mu_{\text{eff}} > 0$ (Stokes regime) and secondly an example with $\mu_{\text{eff}} = 0$ (Darcy limit). In both cases the asymptotic behavior ($h \to 0$) of the computational result is compared with the predicted convergence rates, for varying physical ($\mu_{\text{eff}}$ respectively $\sigma$) and numerical ($\alpha$, $\delta$, $\rho$, and $\ell_{\mathbf{\Omega}}$) parameters.

### The Family of Meshes

The domain for the convergence studies considered in this section is the unit square, i.e., $\mathbf{\Omega} := (0,1)^2$, whose boundary consists of four straight segments referred to as $\Gamma_i$, $i = 0, 1, 2, 3$, with

$$\begin{aligned} \Gamma_0 &:= \{(x,0) \ : \ x \in [0,1]\}\,, & \Gamma_2 &:= \{(x,1) \ : \ x \in [0,1]\}\,, \\ \Gamma_1 &:= \{(1,y) \ : \ y \in [0,1]\}\,, & \Gamma_3 &:= \{(0,y) \ : \ y \in [0,1]\}\,. \end{aligned} \tag{6.1.1}$$

The spatial discretization of the domain is obtained by regular refinements of the diagonally split domain and results in uniform triangular meshes, see Figure 6.1.




(a) Meshes corresponding to level 0 (black) and level 1 (black & gray).

| level | $h_{\boldsymbol{T}}$ | # Cells (Triangles) | # Dofs $(\boldsymbol{v}, p)$ |
|---|---|---|---|
| 0 | 1.41421 | 2 | $(8, 4)$ |
| 1 | 0.707107 | 8 | $(18, 9)$ |
| 2 | 0.353553 | 32 | $(50, 25)$ |
| 3 | 0.176777 | 128 | $(162, 81)$ |
| 4 | 0.0883883 | 512 | $(578, 289)$ |
| 5 | 0.0441942 | 2048 | $(2178, 1089)$ |
| 6 | 0.0220971 | 8192 | $(8450, 4225)$ |
| 7 | 0.0110485 | 32768 | $(33282, 16641)$ |
| 8 | 0.00552427 | 131072 | $(132098, 66049)$ |
| 9 | 0.00276214 | 524288 | $(526338, 263169)$ |

(b) Refinement levels with corresponding characteristic element size, number of elements, and amount of degrees of freedom.

Figure 6.1: Information about the uniform triangular meshes used in the numerical computations for Examples I and II.

**Remark 6.1.1** (On the Meshes)**.**
*Concerning the assumptions on the mesh made in the theoretical analysis (see* $\mathbf{M}_1$*,* $\mathbf{M}_2$*, and* $\mathbf{M}_3$ *in Subsection 5.2.1), the condition* (5.2.2) *is satisfied for all refinements (with* $\eta_0 = 1$ *due to uniformity). Starting from level* 2 *also the condition* (5.2.4) *is fulfilled. However, there are two triangles present in all meshes which were excluded from the theory (they violate the assumption (*$\mathbf{M}_3$*)). Since there are no related perturbations observable in the simulations these are not altered (by appropriately splitting them into two triangles each). For level* 3 *and* $\ell_{\mathbf{\Omega}} = 0.1$*, the assumption* (5.2.6) *is slightly violated, else - for higher levels - satisfied.*

### 6.1.1 Example I: A Generalized Poiseuille/Brinkman Flow

The exact solution for the first example, taken from [HJS11], is given by

$$\begin{aligned}
\boldsymbol{u}^{\mathrm{sol}}(y) &:= \left(u_1^{\mathrm{sol}}(y), 0\right)^T, \\
u_1^{\mathrm{sol}}(y) &:= \begin{cases} \sigma^{-1}\left(1 - \dfrac{\exp\left((1-y)\sqrt{\frac{\sigma}{\mu_{\mathrm{eff}}}}\right) + \exp\left(y\sqrt{\frac{\sigma}{\mu_{\mathrm{eff}}}}\right)}{1+\exp\left(\sqrt{\frac{\sigma}{\mu_{\mathrm{eff}}}}\right)}\right), & \text{if } \frac{\mu_{\mathrm{eff}}}{\sigma} > 0, \\ \sigma^{-1}, & \text{if } \frac{\mu_{\mathrm{eff}}}{\sigma} = 0, \end{cases} \\
p^{\mathrm{sol}}(x,y) &:= 0.5 - x.
\end{aligned} \tag{6.1.2}$$

Note that the velocity solution $\boldsymbol{u}^{\mathrm{sol}}$ does not depend on $x$ and the pressure function $p^{\mathrm{sol}}$ is independent from $y$.

A straightforward calculation shows that the pair $\left(\boldsymbol{u}^{\mathrm{sol}}, p^{\mathrm{sol}}\right)$ in (6.1.2) solves the Brinkman problem (3.2.8a)-(3.2.8b) (with $\sigma > 0$ and $\mu_{\mathrm{eff}} \geqslant 0$) for $\boldsymbol{f} = \boldsymbol{0}$, $g = 0$, natural (Neumann) boundary conditions on the left and right boundaries

$$(-\mu_{\mathrm{eff}}\nabla\boldsymbol{u} + p\mathbb{I}) \cdot \boldsymbol{n} = \begin{cases} -0.5\,\boldsymbol{n} & \text{on } \Gamma_1, \\ +0.5\,\boldsymbol{n} & \text{on } \Gamma_3, \end{cases}$$

and homogeneous essential (Dirichlet) boundary conditions on the top and bottom boundaries

$$\boldsymbol{u}(x,y) = \boldsymbol{0} \quad \text{on } \Gamma_0 \cup \Gamma_2. \tag{6.1.3}$$

#### Boundary Layers – Strong Versus Weak Boundary Conditions

Figure 6.2 depicts the velocity profile $(u_1^{\mathrm{sol}}(y))$ for a few values of $\mu_{\mathrm{eff}}$ and $\sigma$. Notice that, for smaller values of the ratio $\frac{\mu_{\mathrm{eff}}}{\sigma}$, the solution has boundary layers near the Dirichlet boundaries $\Gamma_0$ and $\Gamma_2$, which get steeper if $\frac{\mu_{\mathrm{eff}}}{\sigma}$ decreases.

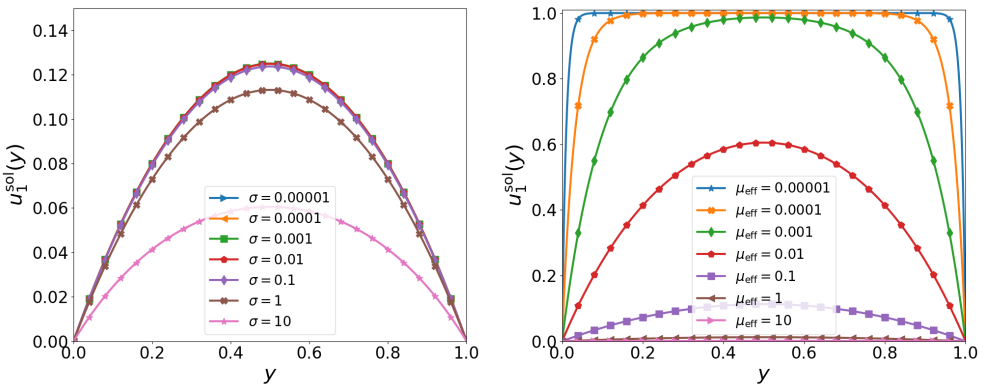
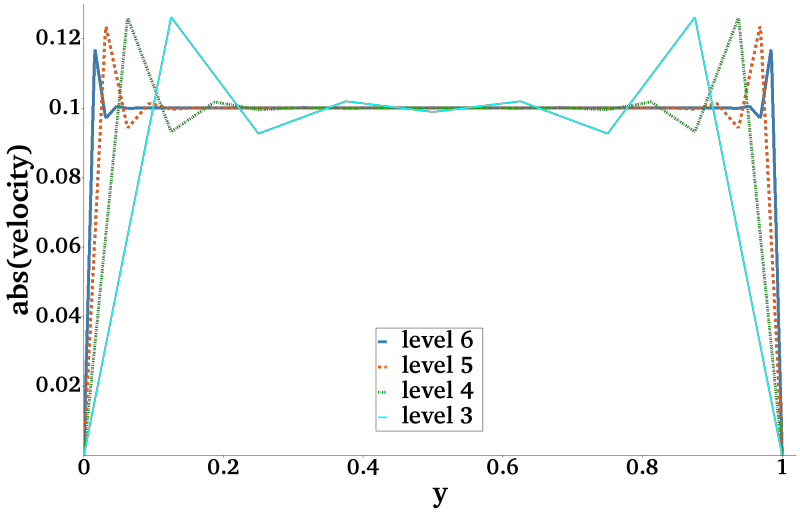


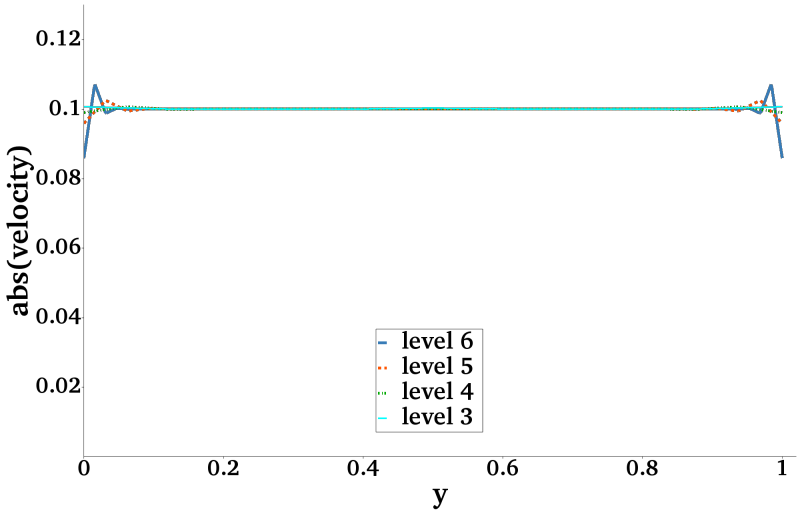

Figure 6.2: Example I: The function $u_1^{\mathrm{sol}}(y)$ defined in (6.1.2) with fixed $\mu_{\mathrm{eff}} = 1$ for different values of $\sigma$ (left) and with fixed $\sigma = 1$ for different values of $\mu_{\mathrm{eff}}$ (right).

The strong imposition of the essential boundary condition (6.1.3) might lead to strong unphysical oscillations in the solution - so-called *overshoots* and *undershoots* - near the boundary, if the mesh is not fine enough to resolve the boundary layer, see also [HJS11].

To examine this phenomenon, the finite element solutions obtained with the method (5.2.10) are compared with the results for the strong imposition of essential boundary conditions (method (5.2.10)). Therefore, the absolute velocity along the line $x = 0.5$ is computed for $(\mu_{\text{eff}}, \sigma) = (0.0001, 10)$ and visualized in Figures 6.3a and 6.3b for different levels of mesh refinement.




(a) Strong imposition of essential boundary conditions on $\Gamma_0$ and $\Gamma_2$.




(b) Weak imposition of essential boundary conditions (via the penalty-free non-symmetric Nitsche method) on $\Gamma_0$ and $\Gamma_2$.

Figure 6.3: Example I: Magnitude of the finite element velocity along the line $x = 0.5$. The parameters $(\mu_{\text{eff}}, \sigma) = (0.0001, 10)$ and levels 3, 4, 5, and 6 were considered, with $\alpha = \delta = 0.1$, $\rho = 0$, $\ell_{\mathbf{\Omega}} = 0.2$ for level 3, and $\ell_{\mathbf{\Omega}} = 0.1$ else.

The strong imposition of boundary conditions yields over- and undershoots near the boundaries, which propagate further into the domain (see Figure 6.3a). We can observe, as expected, that these oscillations reduce for decreasing mesh size (i.e., with increasing level) and become more local.

The penalty-free non-symmetric Nitsche method (see Figure 6.3b) yields a much smoother result due to the relaxation of the essential boundary conditions, although the result of the Nitsche method is not necessarily non-oscillatory. For decreasing mesh-sizes we observe that the deviations from the essential boundary conditions are reduced.

The simulation results suggest that the approximation of the exact solution on coarser meshes and away from the boundary improves when using the penalty-free Nitsche method instead of strongly imposing the respective boundary conditions - at least for problems exhibiting boundary layers.

### Convergence

Numerical simulations are performed for two different physical regimes, considering $\mu_{\text{eff}} = \sigma = 1$, i.e., $\frac{\mu_{\text{eff}}}{\sigma} = 1$ and $\mu_{\text{eff}} = 0.001$, $\sigma = 10$, i.e., $\frac{\mu_{\text{eff}}}{\sigma} = 0.0001$. Moreover, results considering different values of the stabilization parameters are compared. The legend for the forthcoming plots is shown in Figure 6.4. In particular, line colors denote different values of $\alpha$ (GLS stabilization), line styles are related to $\delta$ (grad-div stabilization) and line markers refer to the value of the characteristic length $\ell_{\boldsymbol{\Omega}}$. Due to the absence of corners between Dirichlet boundaries, the corner stabilization is omitted for this example, i.e., $\rho = 0$.

| | | | |
|---|---|---|---|
| -x- | $\alpha = 0.1$, | $\delta = 0.1$, | $\ell_{\boldsymbol{\Omega}} = 0.1$ |
| -o- | $\alpha = 0.1$, | $\delta = 0.1$, | $\ell_{\boldsymbol{\Omega}} = 1$ |
| —x— | $\alpha = 0.1$, | $\delta = 1$, | $\ell_{\boldsymbol{\Omega}} = 0.1$ |
| —o— | $\alpha = 0.1$, | $\delta = 1$, | $\ell_{\boldsymbol{\Omega}} = 1$ |
| ··x·· | $\alpha = 0.1$, | $\delta = 10$, | $\ell_{\boldsymbol{\Omega}} = 0.1$ |
| ··o·· | $\alpha = 0.1$, | $\delta = 10$, | $\ell_{\boldsymbol{\Omega}} = 1$ |
| -x- | $\alpha = 1$, | $\delta = 0.1$, | $\ell_{\boldsymbol{\Omega}} = 0.1$ |
| -o- | $\alpha = 1$, | $\delta = 0.1$, | $\ell_{\boldsymbol{\Omega}} = 1$ |
| —x— | $\alpha = 1$, | $\delta = 1$, | $\ell_{\boldsymbol{\Omega}} = 0.1$ |
| —o— | $\alpha = 1$, | $\delta = 1$, | $\ell_{\boldsymbol{\Omega}} = 1$ |
| ··x·· | $\alpha = 1$, | $\delta = 10$, | $\ell_{\boldsymbol{\Omega}} = 0.1$ |
| ··o·· | $\alpha = 1$, | $\delta = 10$, | $\ell_{\boldsymbol{\Omega}} = 1$ |
| -x- | $\alpha = 10$, | $\delta = 0.1$, | $\ell_{\boldsymbol{\Omega}} = 0.1$ |
| -o- | $\alpha = 10$, | $\delta = 0.1$, | $\ell_{\boldsymbol{\Omega}} = 1$ |
| —x— | $\alpha = 10$, | $\delta = 1$, | $\ell_{\boldsymbol{\Omega}} = 0.1$ |
| —o— | $\alpha = 10$, | $\delta = 1$, | $\ell_{\boldsymbol{\Omega}} = 1$ |
| ··x·· | $\alpha = 10$, | $\delta = 10$, | $\ell_{\boldsymbol{\Omega}} = 0.1$ |
| ··o·· | $\alpha = 10$, | $\delta = 10$, | $\ell_{\boldsymbol{\Omega}} = 1$ |

Figure 6.4: In each plot, we compare the errors varying the GLS stabilization parameter $\alpha$ (orange: 0.1, yellow: 1, purple: 10), the grad-div stabilization parameter $\delta$ (dashed line: 0.1, solid line: 1, dotted line: 10) and the characteristic length $\ell_{\boldsymbol{\Omega}}$.

The convergence of the error in the mesh-dependent norm (5.2.11a) is shown in Figure 6.5, comparing the results for different values of the parameters $\alpha$, $\delta$, and $\ell_{\mathbf{\Omega}}$, described in Figure 6.4.

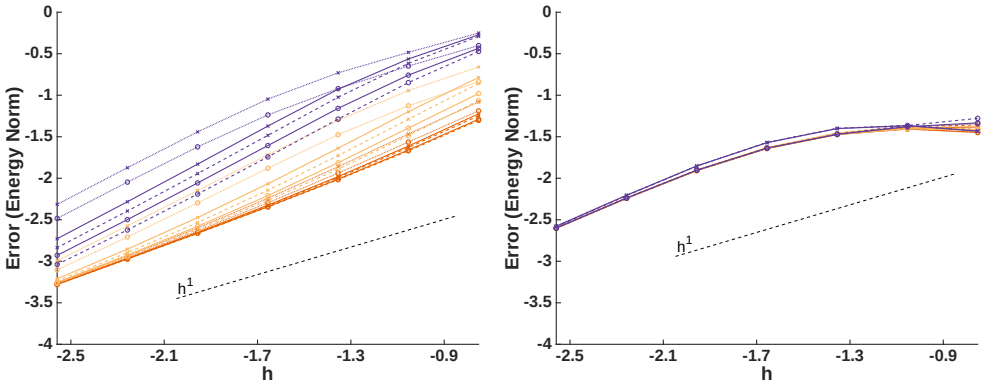



Figure 6.5: Example I: Error in the mesh-dependent norm (5.2.11a) against the mesh size (levels 3 to 9) in double logarithmic scale, for the cases $(\mu_{\text{eff}}, \sigma) = (1, 1)$ (left) and $(\mu_{\text{eff}}, \sigma) = (0.001, 10)$ (right). A dashed line with slope equal to one is also shown. For comparable values of $\mu_{\text{eff}}$ and $\sigma$, choosing larger values of $\alpha$ (purple), larger values of $\delta$ (dotted), and smaller values for $\ell_{\mathbf{\Omega}}$ (crosses) yields less accurate results.

We observe that, in both cases, the predicted convergence rate is obtained in the considered range of stabilization parameters. We also notice that the magnitude of the error for $\mu_{\text{eff}} = \sigma = 1$ slightly increases, the larger $\alpha$ is chosen. The grad-div stabilization parameter $\delta$ seems to have a similar effect on the results.

In Figures 6.6 and 6.7, the different components of the norm (5.2.11a), including the error with respect to the velocity on the boundary, are depicted for the same values of the stabilization parameters, see Figure 6.4. The theoretical first order convergence is obtained in all cases.

For $(\mu_{\text{eff}}, \sigma) = (1, 1)$, the best results (in terms of accuracy and convergence) are obtained choosing $\alpha = 0.1$, $\delta = 0.1$, and $\ell_{\mathbf{\Omega}} = 1$, except for the divergence error of the velocity which reduces with increasing grad-div parameter (hence is smallest with $\alpha = 0.1$, $\delta = 10$, and $\ell_{\mathbf{\Omega}} = 1$). In all cases shown in Figure 6.6, decreasing the GLS parameter $\alpha$ improves the absolute accuracy of the results. This behavior is particularly distinct for the errors concerning the normal velocity at the boundary, the divergence of the velocity, and the pressure.

The plots for $(\mu_{\text{eff}}, \sigma) = (0.001, 10)$ in Figure 6.7 reveal that only the divergence of the velocity, the normal velocity at the boundary, and the pressure errors do depend significantly on the considered parameter variations. Moreover, except for the $L^2$-norm of the pressure error, the results with $\ell_{\mathbf{\Omega}} = 1$ are in general better (smaller errors) than the ones with $\ell_{\mathbf{\Omega}} = 0.1$, when keeping the other parameters constant. For the pressure, we observe an opposite behavior (in

accordance with the scaling in the norm (5.2.11a)). Note that the error of the $L^2$-norm of the pressure behaves irregular for large values of $\alpha$, but linearly interpolating the error between level 3 and level 9 shows the desired convergence.

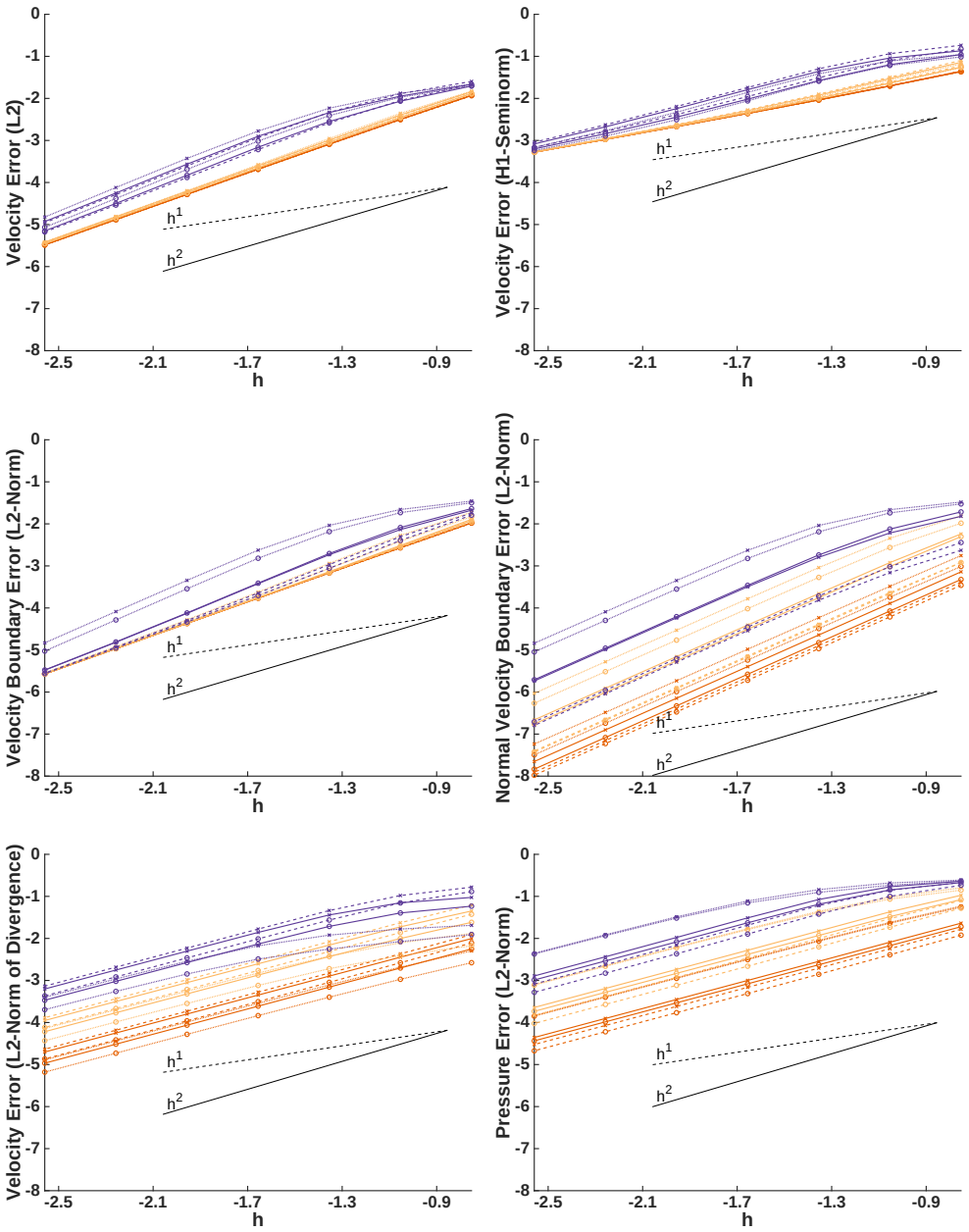



Figure 6.6: Example I: Velocity and pressure errors against the mesh size (levels 3 to 9) in double logarithmic scale, for the case $(\mu_{\text{eff}}, \sigma) = (1, 1)$. The lines with slope equal to 1 (dashed) and 2 (solid) are also shown. In all cases, decreasing the parameter $\alpha$ (orange: 0.1, yellow: 1, purple: 10) improves the absolute accuracy of the results.

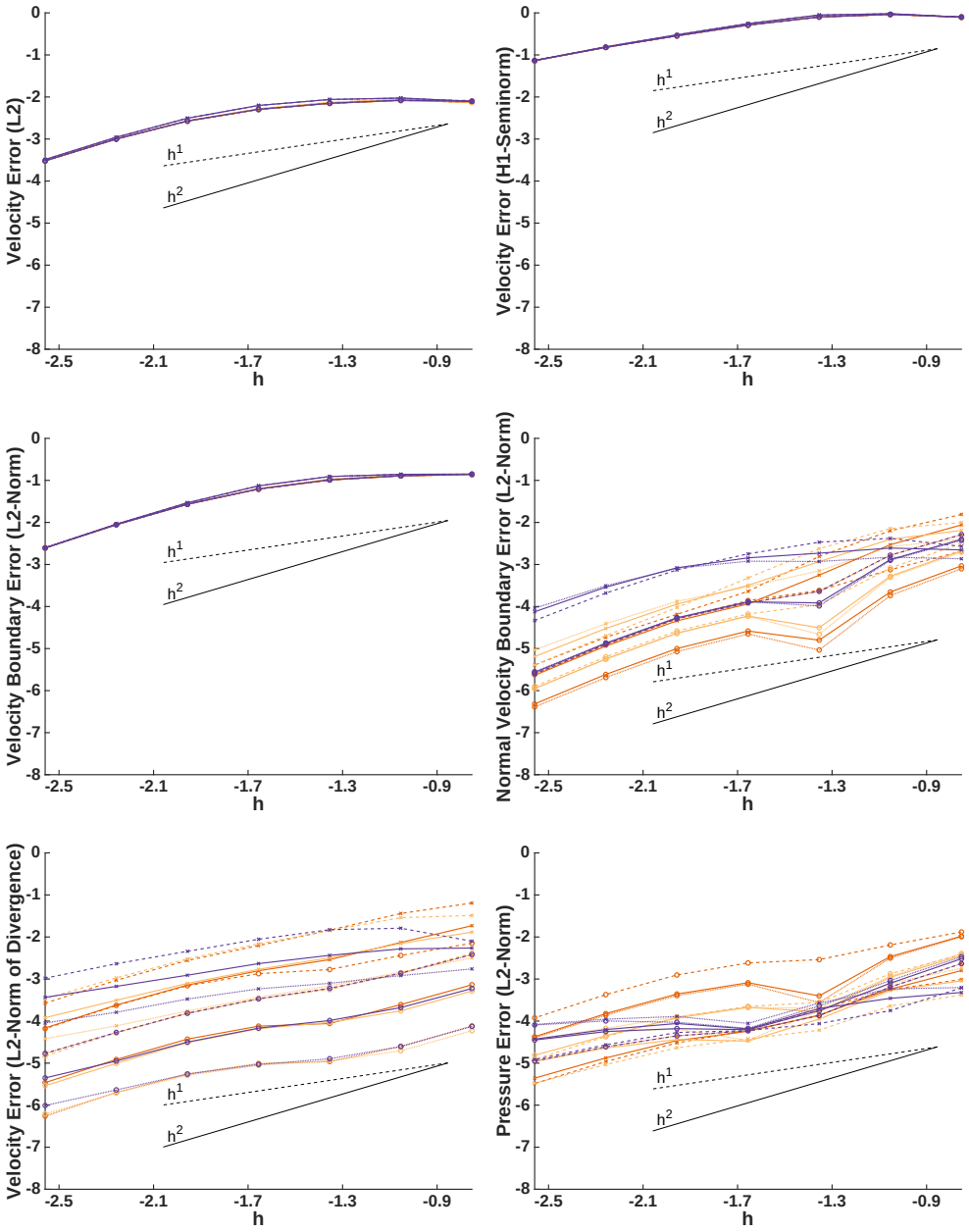



Figure 6.7: Example I: Velocity and pressure errors against the mesh size (levels 3 to 9) in double logarithmic scale, for the case $(\mu_{\text{eff}}, \sigma) = (0.001, 10)$. The lines with slope equal to 1 (dashed) and 2 (solid) are also shown. In this case, the choice $\ell_{\mathbf{\Omega}} = 1$ (circle) and $\delta = 10$ (dotted line) yields lower errors in the velocity, whereas the choice $\ell_{\mathbf{\Omega}} = 0.1$ (crosses) performs better w.r.t. the magnitude of the pressure errors.

### Conditioning

The condition number $\mathcal{K}$ of the coefficient matrix provides information about the propagation of errors from the data (matrix and right-hand side) into the solution.

In particular, for an ill-conditioned matrix ($\mathcal{K} \gg 1$), small perturbations in the data might result in large changes in the solution. When it comes to the point of solving a linear system, conditioning is important for both direct solvers, where error bounds depend on $\mathcal{K}$, and iterative solvers, whose convergence respectively iteration length is limited with respect to $\mathcal{K}$, see, e.g., [BS08, 9.6] and [ESW14]. In the case of stabilized finite element formulations, it must be expected that different choices of the stabilization parameters influence the condition number of the resulting coefficient matrix. How and to what extend this influence reaches, is therefore an important question. This issue is addressed here for Example I by performing numerical experiments.

To this aim, we consider a fixed mesh size ($h = h_{\boldsymbol{T}} \approx 0.02$ for all $\boldsymbol{T} \in \mathcal{T}_h$, corresponding to refinement level 6 in Figure 6.1b) and compute the condition number of the coefficient matrix for two different combinations of the physical parameters, namely, $(\mu_{\text{eff}}, \sigma) \in \{(1,1), (0.001, 10)\}$ and four different values $(0.01, 0.1, 1, 10)$ for $\alpha$ and $\delta$ (fixing the other parameter to one). Moreover, we compare $\ell_{\boldsymbol{\Omega}} = 0.1$ with $\ell_{\boldsymbol{\Omega}} = 1$. Since in this example Neumann boundary conditions are imposed on the left and right boundaries, the corner stabilization is omitted, i.e., $\rho = 0$. The results are shown in Figure 6.8. In all cases, the condition number (2-/Euclidean/spectral norm) has been computed using the routine `cond.m` available in MATLAB R2018a.

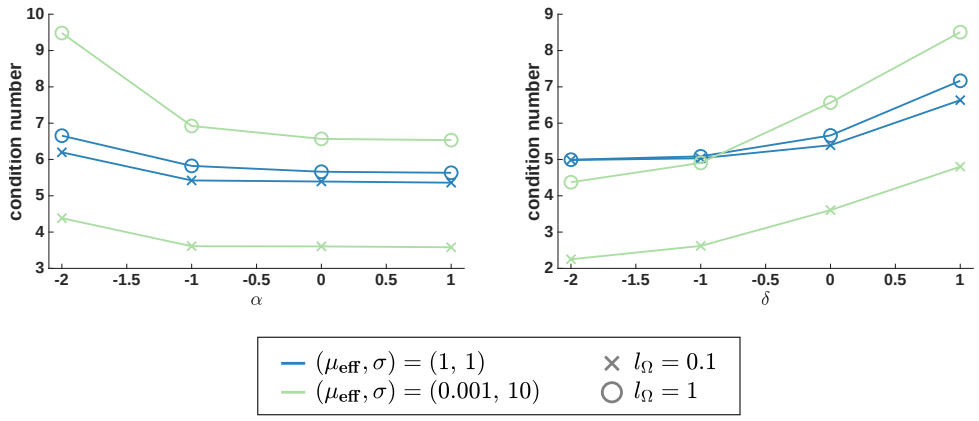



Figure 6.8: Example I: Condition numbers (double logarithmic scale) of the coefficient matrix versus $\alpha$ (left plot, fixing $\delta = 1$) and $\delta$ (right plot, fixing $\alpha = 1$) for refinement level 6. In all cases, we set $\rho = 0$ and $\ell_{\boldsymbol{\Omega}} = 0.1$ (crosses) or $\ell_{\boldsymbol{\Omega}} = 1$ (circles).

With respect to the GLS parameter, the results show that the condition number is largest when choosing very small values of $\alpha$, decreases for increasing $\alpha$, and remains almost unchanged when $\alpha \geq 0.1$. This is consistent with the fact that the GLS stabilization is required for the invertibility of the finite element matrix and it suggests, considering also the convergence results discussed in the previous paragraph, that 0.1 is a suitable order of magnitude for $\alpha$.

For the fixed values of $\alpha$, Figure 6.8 (left) also suggests that increasing $\hat{\mu} = \mu_{\text{eff}} + \sigma \ell_{\boldsymbol{\Omega}}^2$ yields an increase in the condition number. This result could be

explained by observing that the pressure-pressure block introduced by the GLS stabilization, i.e., the contribution due to the term $(\nabla p_h, \nabla q_h)_{\boldsymbol{T}}$, is multiplied by $\frac{\alpha h_{\boldsymbol{T}}^2}{\mu_{\text{eff}}+\sigma \ell_{\boldsymbol{\Omega}}^2}$. Therefore, one can expect that an increase in the denominator produces an effect similar to a decrease in $\alpha$ with respect to this term.

From Figure 6.8 (right) we deduce that increasing the grad-div stabilization parameter $\delta$ yields an increase in the condition number of the matrix.

For fixed values of $\delta \geqslant 1$, we observe a similar scaling with respect to $\hat{\mu}$ as in Figure 6.8 (left). Recalling that the matrix contribution of the grad-div stabilization has the factor $\delta\left(\mu_{\text{eff}} + \sigma \ell_{\boldsymbol{\Omega}}^2\right)$, we conclude that large values of this coefficient yield larger condition numbers. For $\delta \leqslant 0.1$, this is not the case. From a general perspective, it shall be expected that the condition number depends on a more complex interplay of physical and numerical parameters (besides the particular example), whose detailed investigation goes beyond the scopes of this work.

### 6.1.2 Example II: A Trigonometric Darcy Flow

The next example (see, e.g., [BH07b], [BC09]) focuses on a pure Darcy flow, i.e., $\mu_{\text{eff}} = 0$. Again, we consider the unit square $\boldsymbol{\Omega} := (0,1)^2$ with boundary components denoted as in (6.1.1). Then the velocity-pressure pair $\left(\boldsymbol{u}^{\text{sol}}, p^{\text{sol}}\right)$ defined by

$$\begin{aligned}
\boldsymbol{u}^{\text{sol}}(x,y) &:= \begin{pmatrix} u_1^{\text{sol}}(x,y) \\ u_2^{\text{sol}}(x,y) \end{pmatrix} = \begin{pmatrix} -2\pi \cos(2\pi x)\sin(2\pi y) \\ -2\pi \sin(2\pi x)\cos(2\pi y) \end{pmatrix}, \\
p^{\text{sol}}(x,y) &:= \sigma \sin(2\pi x)\sin(2\pi y),
\end{aligned}$$

solves the equations (3.2.8a)-(3.2.8b) with $\mu_{\text{eff}} = 0$, the right-hand sides $\boldsymbol{f} = \boldsymbol{0}$, $g(x,y) = 8\pi^2 \sin(2\pi x)\sin(2\pi y)$, and fulfills the non-homogeneous essential boundary condition

$$(\boldsymbol{u}\cdot\boldsymbol{n})(x,y) = (\boldsymbol{u}^{\text{sol}}\cdot\boldsymbol{n})(x,y) \qquad \text{on } \Gamma.$$

We observe that the pressure scales with $\sigma$ and the velocity is independent from $\sigma$.

#### Convergence

The errors in the energy norm (5.2.11a), for $\sigma \in \{0.001, 1, 1000\}$ and different values of the stabilization parameters, are shown in Figure 6.9. We recall that the legend and mesh information are given in Figures 6.4 and 6.1, respectively.

In all cases, the predicted convergence order is confirmed. As apparent, the magnitude of the errors increases with increasing $\sigma$, whereas the overall behavior shows no significant differences.

We also observe that small GLS parameters $\alpha$ and $\ell_{\boldsymbol{\Omega}} = 0.1$ yield smaller errors than using other options.

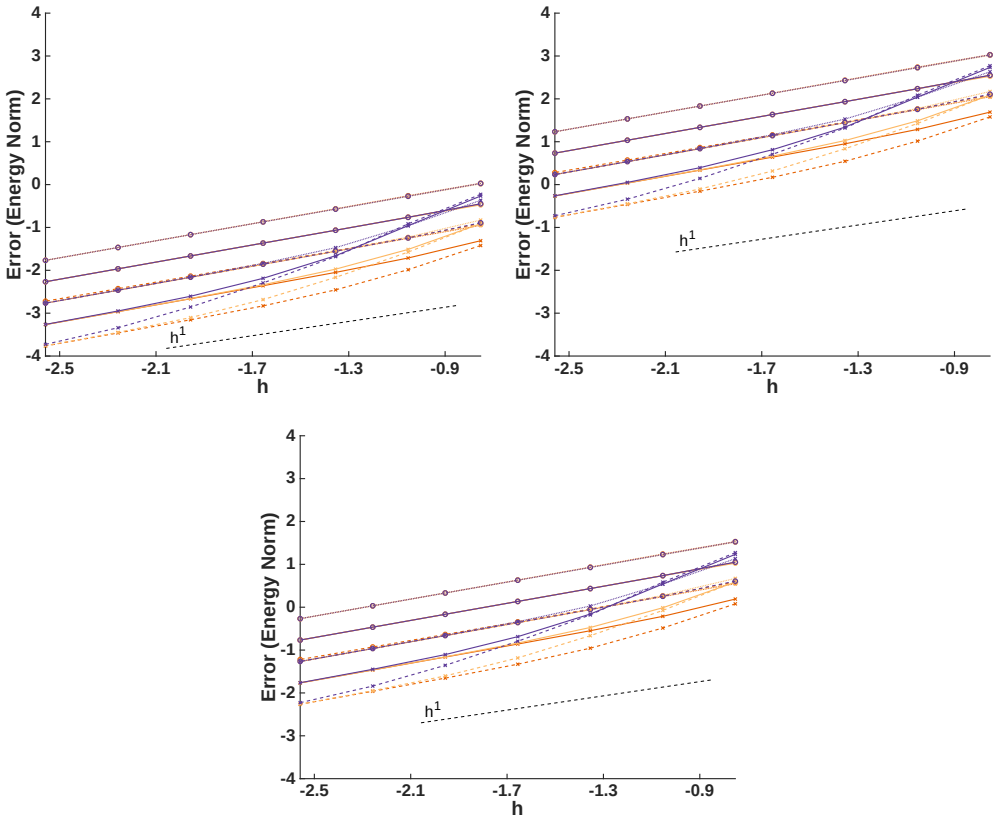



Figure 6.9: Example II: Error in the mesh-dependent norm (5.2.11a) against the mesh size (levels 3 to 9) in double logarithmic scale, for the cases $\sigma = 0.001$ (top left), $\sigma = 1$ (bottom), and $\sigma = 1000$ (top right). The dashed line visualizes a slope equal to 1. In these cases, $\ell_{\mathbf{\Omega}} = 0.1$ (crosses) and $\alpha = 0.1$ (orange) produces overall better results.

Figure 6.10 shows the error components with respect to the velocity for $\sigma = 1$, which behave equivalent for the other values of $\sigma$. The latter phenomenon could be tied to the specific example we are considering, since the exact velocity solution is actually independent from $\sigma$. While the magnitude of the $L^2$-error of the divergence is almost unaltered for fine meshes, the $L^2$-error of the velocity varies significantly, with small values of $\alpha$ and $\ell_{\mathbf{\Omega}}$ yielding the smallest errors (in accordance with the error in the energy norm in Figure 6.9).

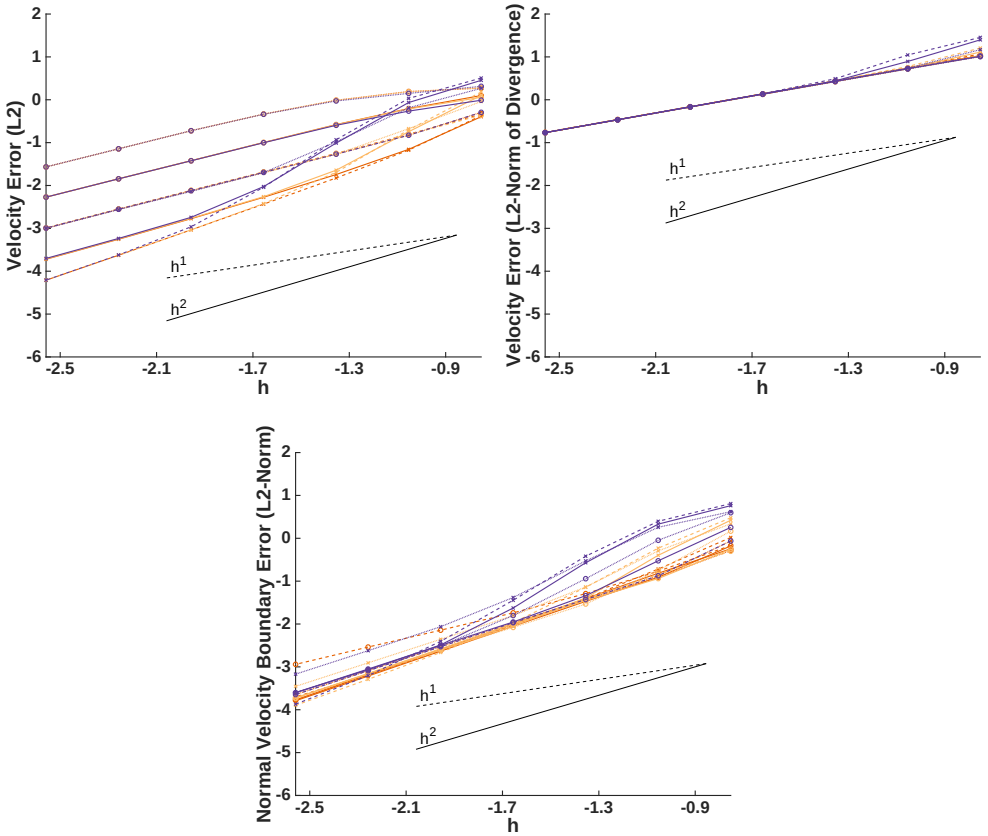



Figure 6.10: Example II: Velocity errors against the mesh size (levels 3 to 9) in double logarithmic scale, for the case $\sigma = 1$ as a representative case for any considered $\sigma$. The lines with slope equal to 1 (dashed) and 2 (solid) are also shown.

Finally, Figure 6.11 visualizes the behavior of the pressure error for different mesh resolutions and $\sigma \in \{0.001, 1, 1000\}$. As reflected in the energy norm, only the magnitude of the errors varies significantly when varying sigma, but the convergence rate is mostly preserved.

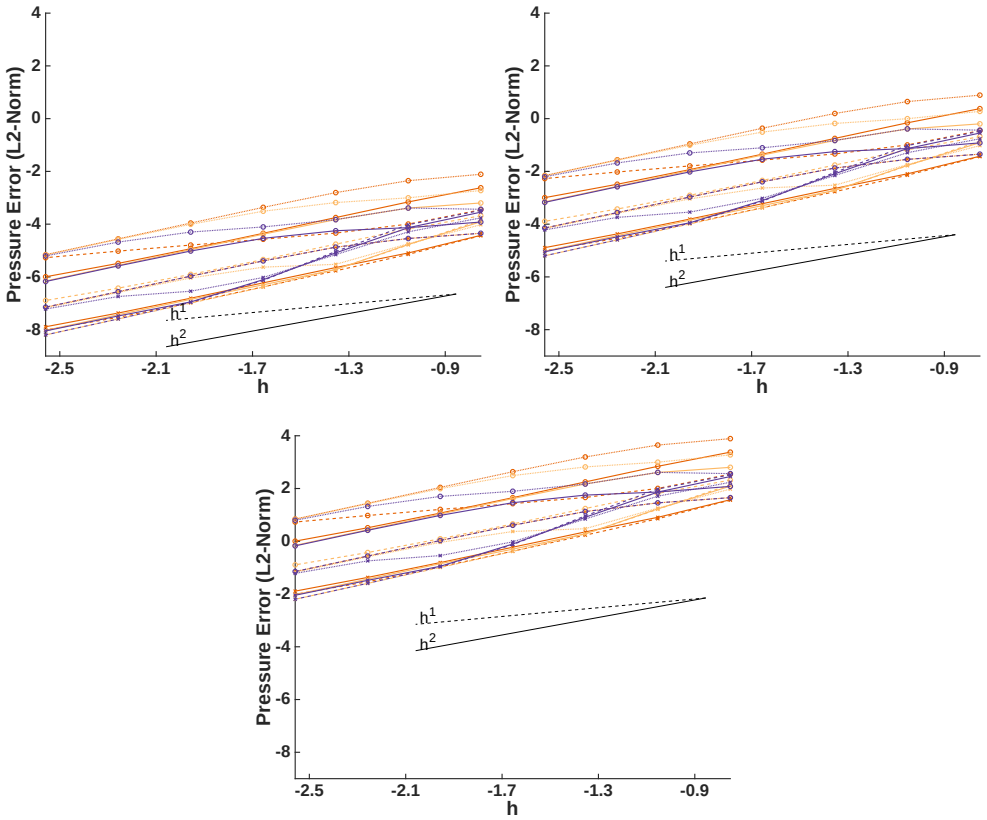



Figure 6.11: Example II: Pressure errors against the mesh size (levels 3 to 9) in double logarithmic scale, for the cases $\sigma = 0.001$ (top left), $\sigma = 1$ (top right), and $\sigma = 1000$ (bottom). The lines with slope equal to 1 (dashed) and 2 (solid) are also shown. Here, $\ell_{\mathbf{\Omega}} = 0.1$ (crosses) resulted in smaller errors.

Altogether, the predicted convergence rate was obtained for all combinations, nevertheless, the choice $\alpha = 0.1$, $\delta = 0.1$, and $\ell_{\mathbf{\Omega}} = 0.1$ seems to lead to the best result. Note, that a variation in the corner stabilization parameter did not result in any change of the convergence behavior, such that $\rho = 1$ was fixed.

### Conditioning

As next, let us numerically investigate the dependence of the conditioning of the coefficient matrix on the stabilization parameters. Therefore, we consider $\sigma \in \{0.001, 1, 1000\}$ and four different values of $\alpha$ and $\delta$ ranging from 0.01 to 10. An additional study has been performed varying the corner stabilization parameter $\rho$. However, since no significant impact on the condition number could be observed, only the outcome for the case $\rho = 1$ is visualized here. We consider the mesh refinement level 6 and compute the condition number using the routine `cond.m` available in MATLAB R2018a for $\ell_{\mathbf{\Omega}} = 0.1$ and $\ell_{\mathbf{\Omega}} = 1$. The results are displayed in Figure 6.12.

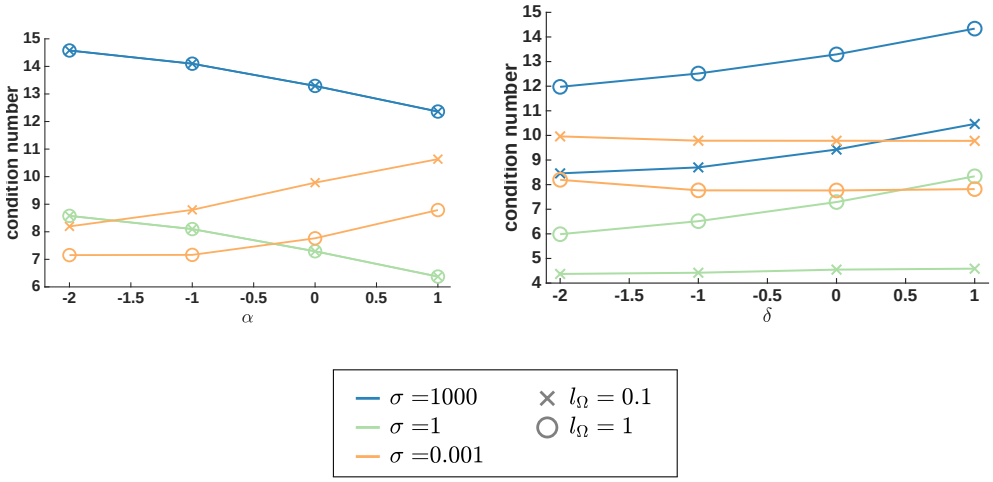



Figure 6.12: Example II: Condition numbers (double logarithmic scale) of the coefficient matrix against $\alpha$ (left plot, fixing $\delta = 1$) and $\delta$ (right plot, fixing $\alpha = 1$) for refinement level 6. In all cases, we set $\rho = 1$ and $\ell_{\mathbf{\Omega}} = 0.1$ (crosses) or $\ell_{\mathbf{\Omega}} = 1$ (circles).

For moderate and large values of $\sigma$, the results are in line with the ones obtained for the conditioning in the previous example (see Subsection 6.1.1). Namely, the condition number is large for small $\alpha$ (GLS stabilization) and tends to increase when increasing $\delta$ (grad-div stabilization). Moreover, $\sigma = 1000$, at least together with $\ell_{\mathbf{\Omega}} = 1$, appears to be the setting with the worst conditioning. As observed before, the increase in $\sigma\ell_{\mathbf{\Omega}}^2$ can be expected to produce an effect analogous to a decrease in $\alpha$ or an increase in $\delta$ (i.e., worsening the conditioning).

A different behavior is obtained for the case $\sigma = 0.001$. In this situation, the condition number increases with $\alpha$ (Figure 6.12, left), suggesting that the conditioning worsens when the pressure-pressure block (which is proportional to $\frac{\alpha h_T^2}{\sigma\ell_{\mathbf{\Omega}}^2}$) is much larger than the block diagonal of the velocity-velocity block matrix (which scales with $\sigma\left(1 + \frac{\alpha h_T^2}{\ell_{\mathbf{\Omega}}^2} + \delta\ell_{\mathbf{\Omega}}^2\right)$).

Varying the grad-div stabilization parameter we observe that, for a very small value of $\sigma\ell_{\mathbf{\Omega}}^2$, the matrix contribution due to the grad-div terms remains small, which can explain the fact that the condition number in Figure 6.12 (right: orange lines and green line with crosses) stays almost constant (in the considered range of $\delta$).

In conclusion, Figure 6.12 confirms that the condition number is strictly related to both, physical and stabilization parameters, and, in particular, that the conditioning might be qualitatively very diverse for the same stabilization parameters but very different physical settings.

## 6.2 Example III: Flow in a T-shaped (Porous) Homogeneous Cavity

This Example III is taken from [Cod01]. Besides a comparison with the results presented in [Cod01], also a sensitivity study for the GLS parameter is provided.

Fluid flow through a T-shaped (non-convex[1]) domain given by

$$\overline{\boldsymbol{\Omega}} = ([-4,4] \times [0,6]) \cup ([-7,7] \times [6,9]) \subset \mathbb{R}^2,$$

as visualized in Figure 6.13, is considered.

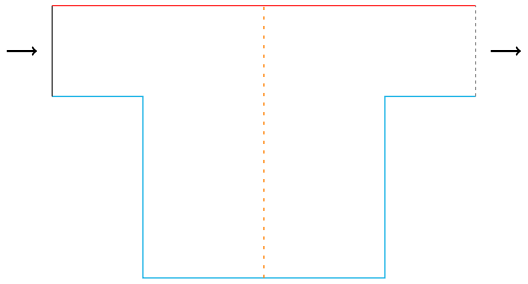


Figure 6.13: Example III: Domain for the T–shaped cavity with the line $\{0\} \times [0,9]$ in orange (for later reference).

A parabolic inflow profile is enforced at the inlet boundary $\{-7\} \times [6,9]$ (black in Figure 6.13), while free outflow is prescribed at the outlet $\{7\} \times [6,9]$ (gray, dotted in Figure 6.13). A constant velocity is imposed on the top boundary $[-7,7] \times \{9\}$ (red in Figure 6.13), while no-slip boundary conditions are prescribed on the remaining parts of the domain (blue in Figure 6.13). Formally, the problem (3.2.8a)-(3.2.8b) is equipped with the following boundary conditions:

$$\begin{aligned}
\boldsymbol{u}(x,y) &= \left(1 - \left(\frac{y}{3} - 3\right)^2, 0\right)^T && \text{on } \{-7\} \times [6,9], \\
(\mu_{\text{eff}} \nabla \boldsymbol{u} - p\mathbb{I}) \cdot \boldsymbol{n} &= \boldsymbol{0} && \text{on } \{7\} \times [6,9], \\
\boldsymbol{u}(x,y) &= (1,0)^T && \text{on } [-7,7] \times \{9\}, \\
\boldsymbol{u}(x,y) &= \boldsymbol{0}, && \text{else.}
\end{aligned}$$

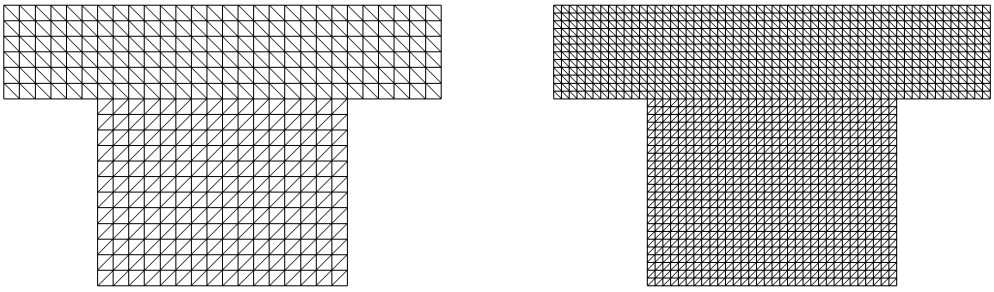
The functions $\boldsymbol{f}$ and $g$ are set equal to zero.

### 6.2.1 The Penalty-Free Non-Symmetric Nitsche Method and the Darcy Regime

For the sake of comparability, a uniform triangular mesh based on nodes with distance 0.25 in $x$- and $y$-direction, as described in [Cod01], is constructed, see Figure 6.14 (left). While this coarse mesh is used for $\mathbb{P}_2/\mathbb{P}_1$ approximations, its

[1] The non-convexity might diminish reasonable regularity assumptions on the solution and thus necessitate a reformulation of the error analysis in Subsection 5.3.2. This is not discussed here, for more information see [Gri11].

uniform refinement (Figure 6.14, right) will be adopted for $\mathbb{P}_1/\mathbb{P}_1$ computations. This has the effect that the number of dofs in total is less different and in particular the same for the velocity (for each velocity component, the Euler formula – # vertices - # edges + # triangles = 1, [Joh16, Rem. C.10] – yields 1533 dofs for level 0 and $\mathbb{P}_2$, and 1533 dofs for level 1 and $\mathbb{P}_1$).



| level | $h_{\boldsymbol{T}}$ | # Cells (Triangles) | # Vertices |
|---|---|---|---|
| 0 | 0.707107 | 720 | 407 |
| 1 | 0.353553 | 2880 | 1533 |

Figure 6.14: Example III: Level 0 (top left) and level 1 (top right) meshes with corresponding data (bottom).

To start with, let us investigate the proposed stabilized penalty-free non-symmetric Nitsche method (5.2.10) for different values of physical parameters.

Weak boundary conditions will be imposed on all boundary segments except for the inlet ($\{-7\} \times [6, 9]$, strong imposition) and the outlet ($\{7\} \times [6, 9]$, natural boundary condition).

Since we consider $\sigma \neq 0$, in general, values for the parameters $\delta$, $\rho$, and $\ell_{\boldsymbol{\Omega}}$ have to be chosen in addition to $\alpha$. According to the theoretical bound (5.2.6), the characteristic length shall be an upper bound on the mesh-size, such that, due to the information in Figure 6.14, $\ell_{\boldsymbol{\Omega}} = 0.36$ is an appropriate choice for the level 1 mesh. In the following computations we fix $\alpha = \delta = \rho = 0.1$.

Due to the homogeneity of the Brinkman momentum equation ($\boldsymbol{f} = \boldsymbol{0}$) in combination with the homogeneity of the natural boundary condition (outlet), the ratio $\frac{\mu_{\text{eff}}}{\sigma}$ characterizes the behavior of the $x$-velocity, in particular along the vertical, central line $\{0\} \times [0, 9]$. In other words, the plots in Figure 6.15 are equivalent keeping $\frac{\mu_{\text{eff}}}{\sigma}$ constant.

Firstly we observe that the curves for $\mu_{\text{eff}} = 0.001$ and $\mu_{\text{eff}} = 0$ differ only slightly. Moreover, on both boundaries $y = 0$ and $y = 9$, the boundary condition relaxes for decreasing values of $\frac{\mu_{\text{eff}}}{\sigma} < 1$. This phenomenon is much more significant close to the top boundary ($y = 9$) since the gradient here is steeper. To be precise, the boundary value decreases from approximately 0.98 (for $\frac{\mu_{\text{eff}}}{\sigma} = 1$) to 0.38 (for $\frac{\mu_{\text{eff}}}{\sigma} = 0.0001$).

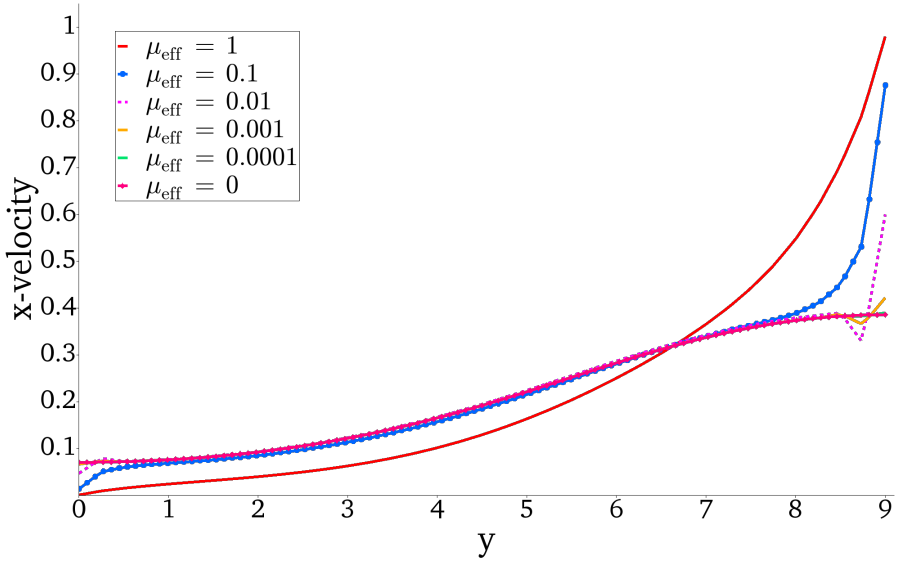



Figure 6.15: Example III: First velocity component on the line $\{0\}\times[0,9]$ (orange, dashed line in Fig. 6.13) for the level 1 mesh (see Figure 6.14) with $\ell_{\boldsymbol{\Omega}} = 0.36$, $\alpha = \delta = \rho = 0.1$, and different values of $\mu_{\text{eff}}$ (fixing $\sigma = 1$).

Let us now focus on the Brinkman problem with physical parameters

$$(\mu_{\text{eff}}, \sigma) = \left(1, 10^4\right).$$

Thus, the Darcy term $\sigma\boldsymbol{u}$ dominates the Stokes term $\mu_{\text{eff}}\Delta\boldsymbol{u}$ such that this setting corresponds to the Darcy regime. This choice of parameters has been as well considered in [Cod01], where a method for Navier–Stokes problems (including the term $\sigma\boldsymbol{u}$ accounting for medium resistance) is proposed, that in particular aims at stabilizing continuous equal-order interpolations and assuring stability for large values of $\sigma$.

Figure 6.16 shows a comparison of the x-velocity component along the vertical line $\{0\} \times [0, 9]$ for different methods. The numerical data corresponding to a pressure-stabilized method discussed in [Cod01][2] (based on $\mathbb{P}_2/\mathbb{P}_1$) have been extracted from [Cod01, Fig. 10] using the free software WebPlotDigitizer[3].

For the computations with the finite element pair $\mathbb{P}_2/\mathbb{P}_1$ (standard Galerkin approach), the coarser mesh is used (level 0) and for computations with $\mathbb{P}_1/\mathbb{P}_1$ (stabilized), the level 1 mesh is utilized, see Figure 6.14.

[2] Although the stabilization in [Cod01] is designed for any combination of continuous interpolations, the results were given for $\mathbb{P}_2/\mathbb{P}_1$, where no pressure stabilization is necessary.

[3] `http://arohatgi.info/WebPlotDigitizer/`

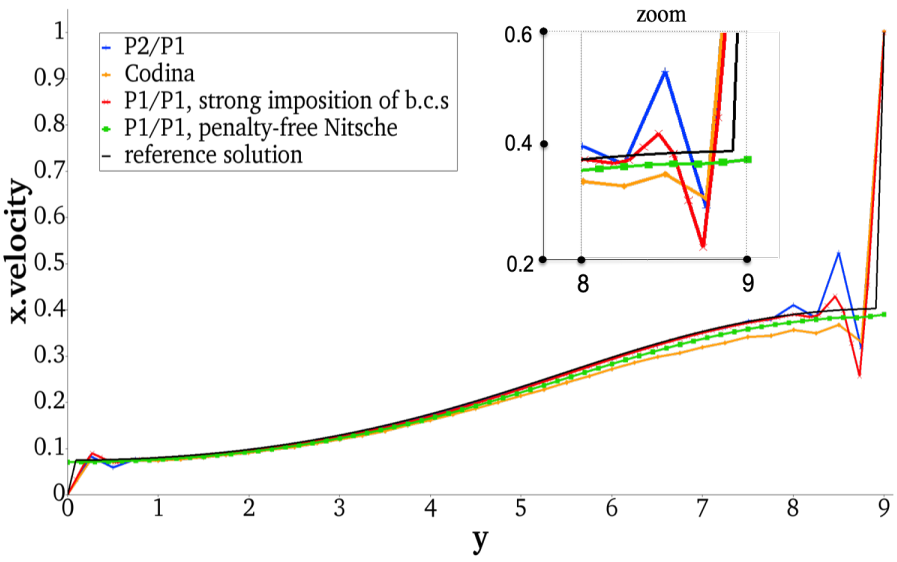



Figure 6.16: Example III: First velocity component on the line $\{0\} \times [0, 9]$ (orange, dashed line in Fig. 6.13) for $\mathbb{P}_2/\mathbb{P}_1$ with the level 0 mesh (as described in [Cod01]) and the level 1 mesh else. In any case, it is $(\mu_{\text{eff}}, \sigma) = (1, 10^4)$ and, if present, $\ell_{\mathbf{\Omega}} = 0.36$ and $\alpha = \delta = \rho = 0.1$.

For comparative purposes, the black line represents a *reference solution* obtained with a $\mathbb{P}_2/\mathbb{P}_1$ standard Galerkin finite element approach on a uniform refinement of the level 0 mesh from Figure 6.14 (with 92897 vertices, 184320 triangles), that is apparently fine enough to resolve the boundary layers.

The curves corresponding to $\mathbb{P}_2/\mathbb{P}_1$ (red) and stabilized $\mathbb{P}_1/\mathbb{P}_1$ with strongly imposed boundary conditions (blue) oscillate in particular in the vicinity of the boundary $y = 9$, and these oscillations propagate also into the domain. The reason for this is that the boundary layer is not sufficiently well resolved by the mesh.

In contrast to that, the stabilized method from Codina (yellow, which is still based on the strong imposition of b.c.s) results in small oscillations only. At the boundary, the penalty-free Nitsche method (green) relaxes the essential boundary conditions which is the expected behavior. Further, the stabilized method (5.2.10) does not yield any oscillations. In the interior of the domain, their behavior is comparable, although the result of the method proposed in this thesis is closer to the reference solution.

### 6.2.2 Pressure Stabilization and the Stokes Regime

The numerical results in Section 6.1 have confirmed that an appropriate choice of the stabilization parameters can be crucial for the quality of the approximation. This aspect is particularly delicate in applications when only a limited number of refinements is possible, and hence the magnitude of the error becomes important (not only the asymptotic behavior). Although the theoretical analysis in Chapter 5 does not predict any explicit bounds for the non-symmetric GLS method to be well-posed (except for $\alpha > 0$), the purpose of the following discussion is to

demonstrate that the choice of the GLS parameter $\alpha$ is decisive. To that end, a sensitivity study is performed, which takes into account qualitative behaviors for wider ranges of $\alpha$ than those discussed in Section 6.1.

Subsequently, we analyze the setting determined by the parameters

$$(\mu_{\text{eff}}, \sigma) = (1, 0)\,,$$

such that the Brinkman problem reduces to a Stokes problem. Since the exact solution is not known, computational results can only be compared to a *reference solution*, which will be computed using $\mathbb{P}_2/\mathbb{P}_1$ (either on a fixed mesh for comparative reasons or on a fine mesh, replacing the unknown analytical solution).

In order to qualitatively compare the finite element solutions with respect to different values of $\alpha$, we impose the boundary conditions strongly and disable redundant stabilizations. In accordance with the theory, the grad-div and corner stabilizations are not necessary for stability of the setting under consideration (Stokes limit, $\sigma = 0$) such that we set $\delta = \rho = 0$. Note that $\ell_{\mathbf{\Omega}}$ is not present in the finite element formulation for the Stokes limit. The considered unstructured, slightly non-uniform mesh is shown in Figure 6.17.

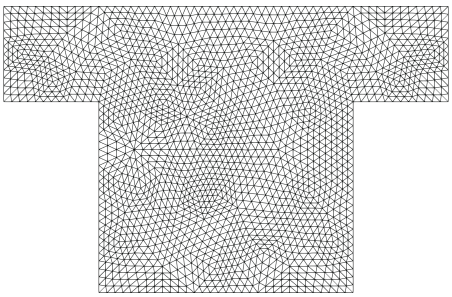


Figure 6.17: Example III: Mesh with 3840 triangles, 2013 vertices, and $(h_{\min}, h_{\max}) = (0.192, 0.339)$.

The range of values of the GLS parameter considered in the numerical study comprises the set $\left\{0, 1, 10^{\pm 1}, 0.03, 10^{\pm 2}, 10^{\pm 3}, 10^{\pm 4}\right\}$. Figures 6.18 and 6.19 compare the first component of the velocities and the pressures, respectively, both along the vertical line $\{0\} \times [0, 9]$ for $\alpha \in \left\{10^{-4}, 0.03, 10\right\}$ with the result of the standard Galerkin method for $\mathbb{P}_2/\mathbb{P}_1$ on the mesh in Figure 6.17.

We observe that the numerical solutions for the velocity obtained for all values of $\alpha$ are very similar, although, in the case $\alpha = 10$, larger discrepancies can be observed with respect to the reference solution.

In contrast to that, the quality of the corresponding pressure fields differs significantly. For $\alpha = 0.0001$ pressure oscillations can be observed, and for $\alpha = 10$ the pressure value is smaller (by a factor of $\frac{1}{2}$) than for the other variants.

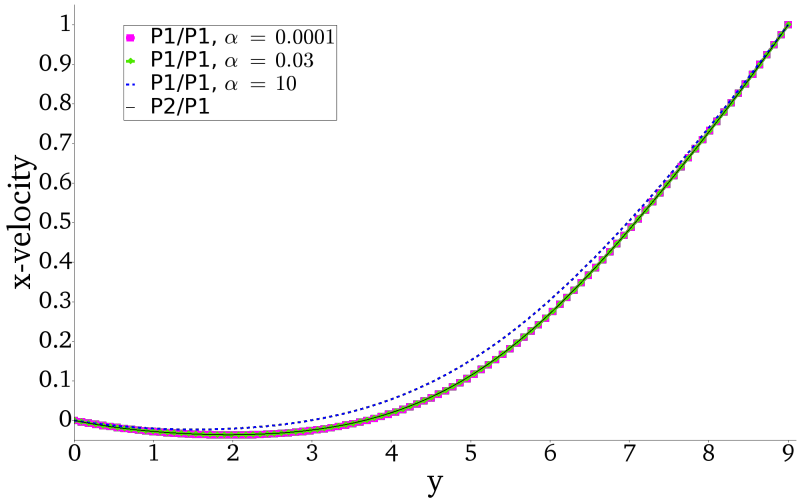



Figure 6.18: Example III: Comparison of the first velocity components along the line $\{0\} \times [0, 9]$ (orange, dashed line in Fig. 6.13) for $\mathbb{P}_2/\mathbb{P}_1$ and the stabilized $\mathbb{P}_1/\mathbb{P}_1$ finite element solutions for different values of $\alpha$ and $\delta = \rho = 0$.

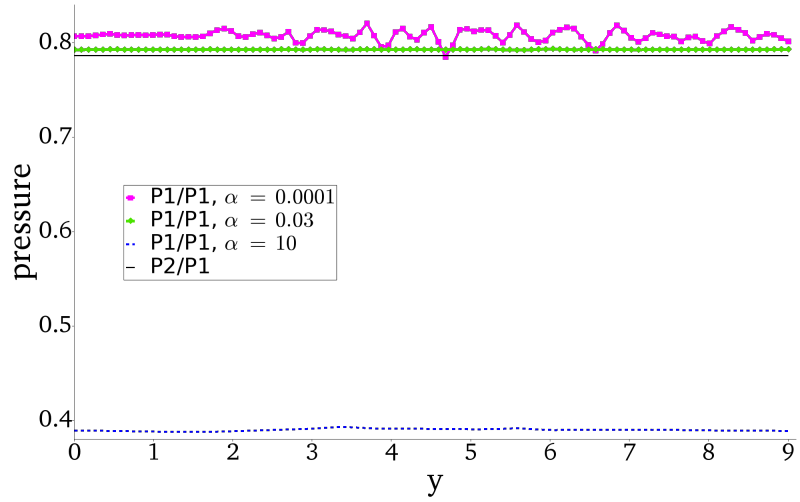



Figure 6.19: Example III: Comparison of the pressures along the line $\{0\} \times [0, 9]$ (orange, dashed line in Fig. 6.13) for $\mathbb{P}_2/\mathbb{P}_1$ and the stabilized $\mathbb{P}_1/\mathbb{P}_1$ finite element solutions for different values of $\alpha$ and $\delta = \rho = 0$.

Figure 6.20 presents a more extensive qualitative comparison of the pressure approximations over the whole domain.

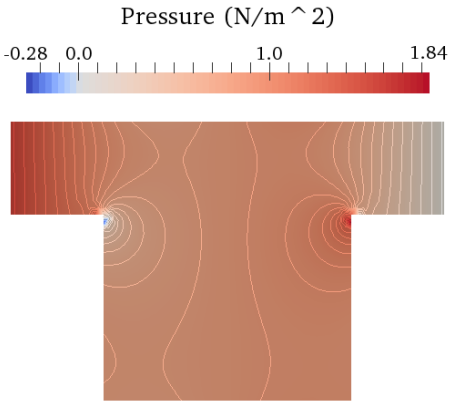
Among the tested values of $\alpha$, spurious oscillations can be observed for $0 \leqslant \alpha \leqslant 10^{-3}$ (see Figures 6.20b and 6.20c), which disappear increasing the stabilization parameter.

For large values of $\alpha$, the pressure field differs significantly from the reference solution, see Figure 6.20e. This could be explained by the fact that linear finite elements are used and hence the residual-based stabilization can only consider

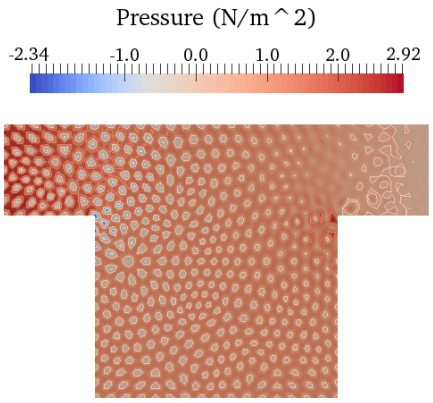
cell-wise residuals with respect to $\mathbb{P}_1$. Since $\boldsymbol{f} = \mathbf{0}$ the stabilization term in the continuity equation is solely concerning $(\nabla p_h, \nabla q_h)_{\boldsymbol{T}}$, whose strong weighting implies small pressure gradients per mesh cell.

In accordance with Figures 6.18 and 6.19, using $\alpha = 0.03$ (Figure 6.20d) yields the best coincidence with the reference solution (Figure 6.20a).

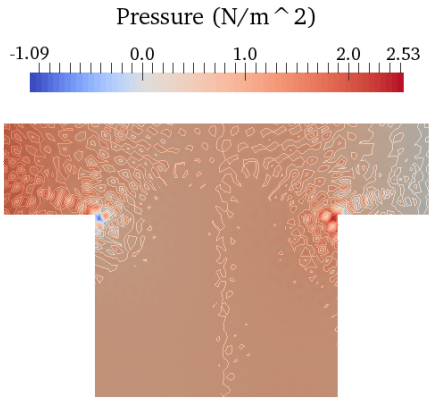
As expected from the theory, mesh refinement improves the finite element solution for large values of $\alpha$.



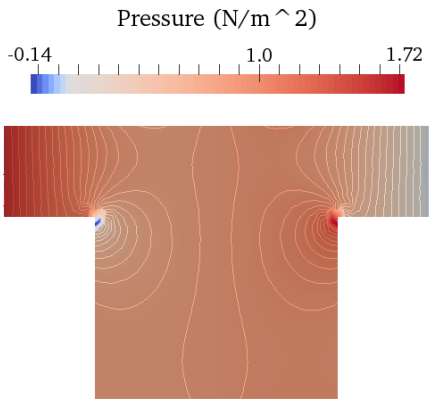

(a) $\mathbb{P}_2/\mathbb{P}_1$ – No stabilization (reference solution).



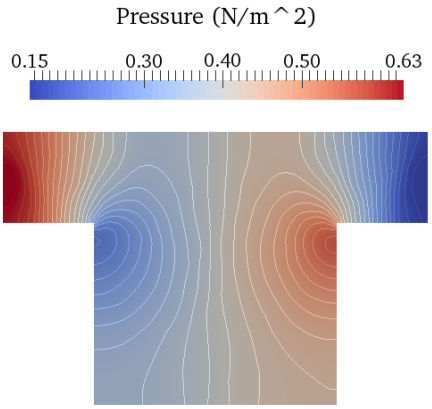

(b) $\mathbb{P}_1/\mathbb{P}_1$ – GLS parameter $\alpha = 0$.




(c) $\mathbb{P}_1/\mathbb{P}_1$ – GLS parameter $\alpha = 0.0001$.




(d) $\mathbb{P}_1/\mathbb{P}_1$ – GLS parameter $\alpha = 0.03$.




(e) $\mathbb{P}_1/\mathbb{P}_1$ – GLS parameter $\alpha = 10$.

Figure 6.20: Example III: Pressure fields with pressure isolines for $\mathbb{P}_1/\mathbb{P}_1$ with different values of the GLS parameter $\alpha$ and $\delta = \rho = 0$, compared with the $\mathbb{P}_2/\mathbb{P}_1$ solution.

## 6.3 Example IV: Flow Above and Below a Sediment-Water Interface

This example has been presented in [CW07a] and [CW07b] with the objective to investigate hydrodynamic interactions between unidirectional laminar flow coupled with porous flow, using a multi-domain approach for coupling the Stokes and Darcy problems. It was further considered in [CJW14] and [Wil19], and will be used here to compare the proposed method (5.2.10) with the results given in the aforementioned literature.

The heterogeneous domain is given by $\overline{\mathbf{\Omega}} = \overline{\mathbf{\Omega}_p} \cup \overline{\mathbf{\Omega}_f} := [0,2]^2$ and comprises a porous region $\mathbf{\Omega}_p$ (e.g., filled with sediment) and a plain region $\mathbf{\Omega}_f$, both fully saturated respectively filled with a viscous fluid (e.g., water). The interface is formally given by

$$\begin{aligned}\overline{\Gamma_I} := &\operatorname{conv}\{(0,1.5),(0.9,1.6)\} \cup \operatorname{conv}\{(0.9,1.6),(1,1.5)\} \\ &\cup \operatorname{conv}\{(1,1.5),(1.9,1.6)\} \cup \operatorname{conv}\{(1.9,1.6),(2,1.5)\},\end{aligned}$$

with $\operatorname{conv}\{\dots\}$ denoting the convex hull. It separates the porous part of the domain from the plain region and has geological large scale characteristics similar to rigid dunes (inclinations and declinations) associated with a riverbed. The inclinations are expected to cause an *interfacial exchange zone* (IEZ), i.e, a region below the interface that receives water from the plain region. The boundary of the domain is decomposed into 6 boundary segments denoted by $\Gamma_0, \dots, \Gamma_5$ as illustrated in Figure 6.21.

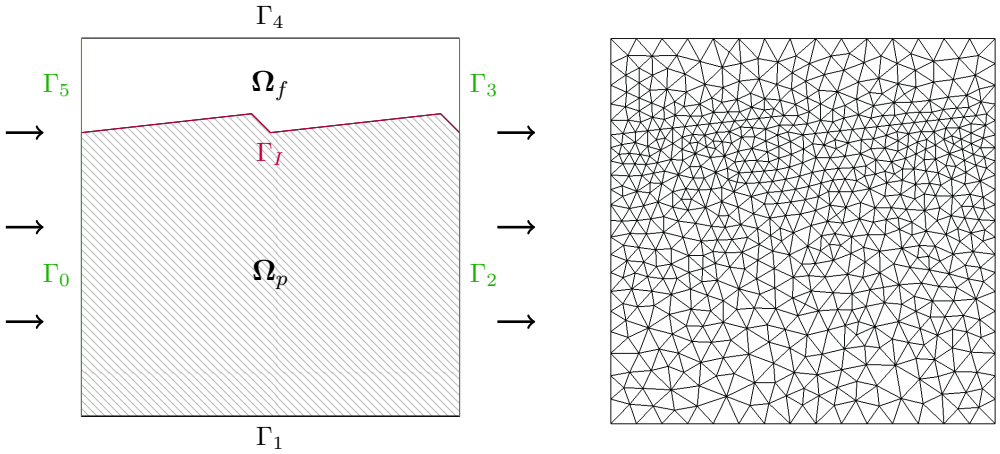



Figure 6.21: Example IV: Domain (left) and initial mesh (right), aligned with and slightly refined at the interface.

As described in [CW07b], in the subdomain below the interface $\Gamma_I$, i.e., in $\mathbf{\Omega}_p$, a homogeneous (right-hand sides vanish) Darcy problem with fluid viscosity $\mu = 0.001\,\mathrm{Pa\,s}$ and permeability $\mathbb{K} = 10^{-10}\,\mathrm{m}^2$ (e.g., coarse sand) is solved. Meanwhile, the fluid flow in the plain region, $\mathbf{\Omega}_f$, is assumed to obey a

homogeneous (Navier–)Stokes problem (with the same fluid viscosity). On top of the domain no-slip boundary conditions

$$\boldsymbol{u} = \boldsymbol{0} \quad \text{on } \Gamma_4$$

are imposed and on the bottom boundary the no flux condition (impermeability)

$$\boldsymbol{u} \cdot \boldsymbol{n} = 0 \quad \text{on } \Gamma_1$$

is prescribed. In order to model an infinite, horizontal, periodic extension of $\boldsymbol{\Omega}$, periodicity of the velocity in the $x$-direction is enforced by identifying the velocity dofs at the inlet $\Gamma_0 \cup \Gamma_5$ with those at the outlet $\Gamma_2 \cup \Gamma_3$ (green). Further, flow is driven by an induced pressure difference $p_0$ as

$$((\mu_{\text{eff}} \nabla \boldsymbol{u} - p\mathbb{I}) \cdot \boldsymbol{n})\,|_{\Gamma_0 \cup \Gamma_5} = -((\mu_{\text{eff}} \nabla \boldsymbol{u} - p\mathbb{I}) \cdot \boldsymbol{n})\,|_{\Gamma_2 \cup \Gamma_3} + p_0\,\boldsymbol{n},$$

here in form of

$$\begin{aligned} (\mu_{\text{eff}} \nabla \boldsymbol{u} - p\mathbb{I}) \cdot \boldsymbol{n} &= 0 && \text{on } \Gamma_0 \cup \Gamma_5, \\ (\mu_{\text{eff}} \nabla \boldsymbol{u} - p\mathbb{I}) \cdot \boldsymbol{n} &= p_0\,\boldsymbol{n} && \text{on } \Gamma_2 \cup \Gamma_3, \end{aligned}$$

with $p_0 = 0.001$ introducing a pressure difference that induces fluid flow from left to right. In both subregions, the external forces are assumed to be absent, i.e., $\boldsymbol{f} = \boldsymbol{0}$ and the flow is divergence-free, i.e., $g = 0$.

In [CW07a] and [CW07b], the simulations are performed in a sequential procedure, starting with the solution of a Navier–Stokes problem (including a convective term) in $\boldsymbol{\Omega}_f$ using $\mathbb{P}_2/\mathbb{P}_1$. Based on the fact that the Stokes velocity at the interface is very small compared to the bulk Stokes velocity, the no-slip condition is imposed on the interface. The resulting pressure distribution along the interface is then used as a boundary condition for the (primal) Darcy problem (via $\mathbb{P}_2$) in $\boldsymbol{\Omega}_p$ (enforcing pressure continuity across the subdomains). This approach generally yields a discontinuous velocity across the interface (Stokes velocity is zero) resulting in a slight mass imbalance.

Example IV was also investigated in [CJW14] and [Wil19] in the context of iterative subdomain methods for the solution of coupled Stokes–Darcy problems, where also here the Darcy problem has the form of a Poisson problem for the pressure. As interface conditions, continuity of the normal velocity, continuity of the normal stress and the Beavers–Joseph–Saffman condition (see Subsection 2.3.2) are imposed, which necessitates an appropriate choice of the Beavers–Joseph constant.

In the following discussion, we will numerically assess the use of the Brinkman equations for a single domain approach for fluid-porous multi-domains, without imposing explicit interface conditions. Therefore we fix the physical parameters as follows:

$$\begin{aligned} \boldsymbol{\Omega}_f &: \quad (\mu_{\text{eff}}, \sigma) = \left(10^{-3}, 0\right), \\ \boldsymbol{\Omega}_p &: \quad (\mu_{\text{eff}}, \sigma) = \left(0, 10^{7}\right). \end{aligned}$$

We observe that both coefficients jump at $\Gamma_I$ and that the jump is of the order $10^{-3}$ in $\mu_{\text{eff}}$ respectively $10^7$ in $\sigma$.

The numerical simulations are performed on a mesh obtained by three uniform refinements (yielding 41441 nodes) of the initial mesh shown in Figure 6.21 (right). Sufficiently fine meshes are necessary to resolve large fine scale variations in particular when using low polynomial order finite element spaces.

The parameters $\alpha = \delta = 0.1$, $\rho = 0$ (no Nitsche corners), and $\ell_{\mathbf{\Omega}} = 0.021 \approx h_{\max}$ (approximately satisfying condition (5.2.6)) were used and the penalty-free non-symmetric Nitsche method was applied for $\Gamma_1$ and $\Gamma_4$.

**Remark 6.3.1** (On Conditions Inherently Imposed at the Quasi-Interface)**.** *The proposed stabilized penalty-free Nitsche method* (5.2.10) *enforces continuity of the velocity (components) and continuity of the pressure throughout the whole domain. Hence, continuity is also enforced across quasi-interfaces, where coefficients jump. Interpreted as a domain decomposition approach, the conservation of mass is assured by the continuity of the normal velocity which is also an interface condition of the Stokes–Darcy coupling. The remaining transition conditions at the interface in general differ.*

The computational results in Figure 6.22 show several characteristic properties of the flow.

Firstly, the method predicts an *interfacial exchange zone* localized below the dunes in the porous region. Such IEZs are particularly important for the prediction of, e.g., thermal, chemical, and biological phenomena in the porous medium. Further, the IEZ in Figure 6.22 (left) has a similar shape and depth as for the Stokes–Darcy coupling with explicit use of interface conditions in [CJW14] or [Wil19].

Below this IEZ, horizontal flow in the porous medium which is not affected by the interface and induced by the pressure difference at the inlet and outlet boundaries, the so-called *underflow*, is observed. The streamlines show that particles located in a small layer at the interface in the porous region are transported along the interface before propagating deeper into the porous region.

The 3D visualization, Figure 6.22 (right), highlights the velocity peaks above the dune pikes.

It is important to observe that the velocity in the porous region is very small ($5{\cdot}10^{-11}$ m/s in the bulk) in comparison with the velocity in the non-porous region.

Secondly, Figure 6.23a (left) shows the corresponding pressure that drops behind the dunes and exceeds the pressure difference induced by the boundary conditions (which is $10^{-3}$). This is in agreement with the results presented in the literature.

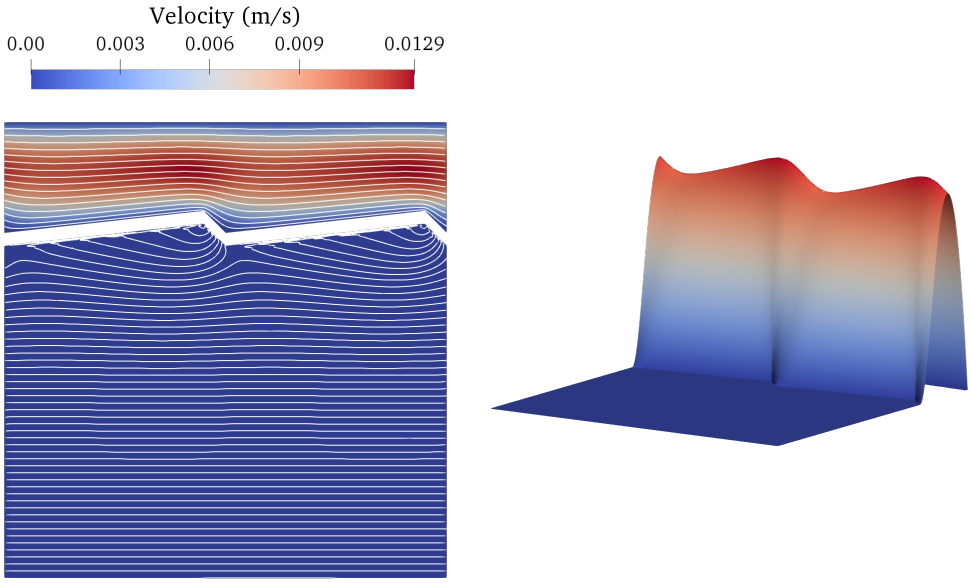
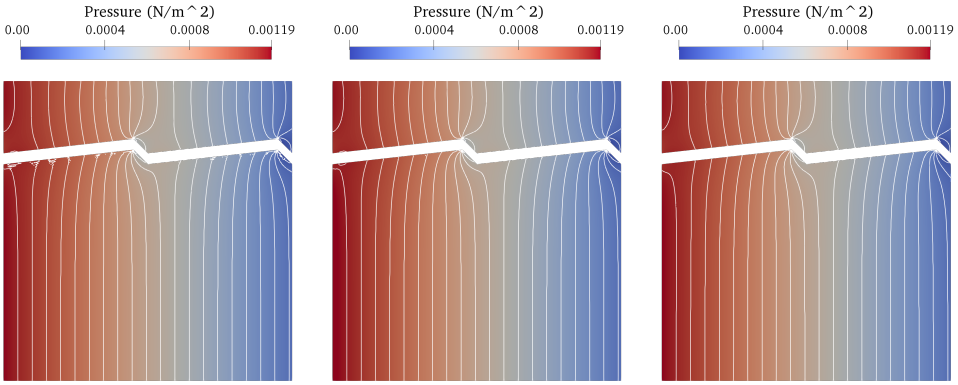


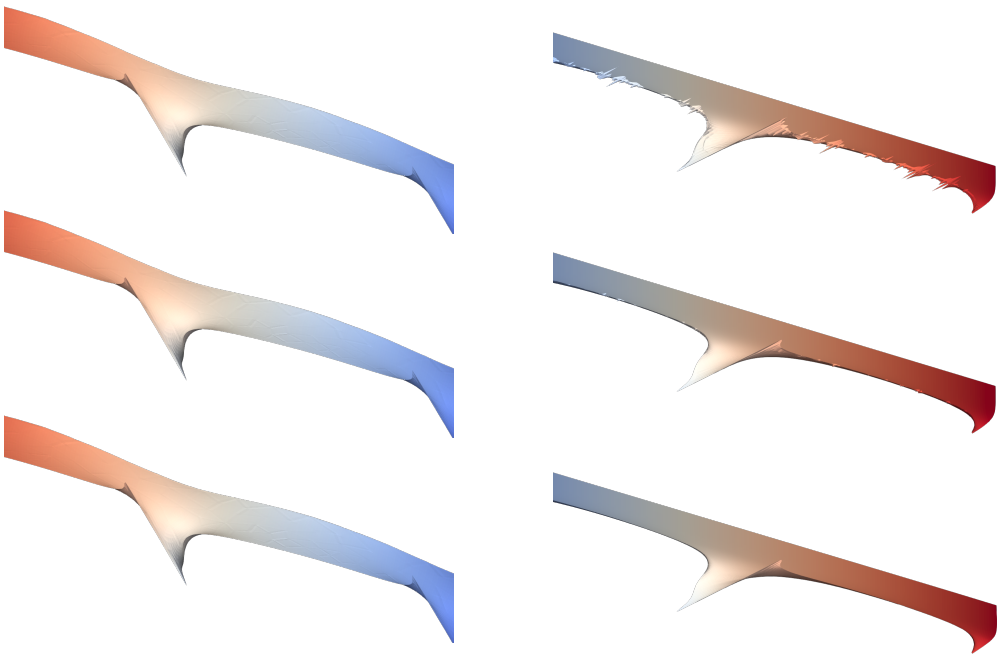

Figure 6.22: Example IV: Magnitude of the velocity with streamlines (left) and in 3D (right) using $\alpha = \delta = 0.1$, $\rho = 0$, and $\ell_{\mathbf{\Omega}} = 0.021$. For the purpose of clear visualization, the Stokes region has been spatially separated (after the simulations) from the Darcy region (left).

However, the contours are non-smooth near the interface. Numerical tests have shown that the pressure smoothens in the vicinity of the interface with decreasing values of $\ell_{\mathbf{\Omega}}$, as shown in Figure 6.23a (center and right).

A direct comparison of the pressure at the interface is shown in Figure 6.23. We observe that the behavior in the non-porous region is smooth and without significant differences for the varying values of $\ell_{\mathbf{\Omega}}$ (Figure 6.23b). In fact, this is what we expect, since whenever it is $\sigma = 0$, $\ell_{\mathbf{\Omega}}$ disappears from the formulation. In contrast to that, the pressure in the Darcy region is oscillating at the interface for $\ell_{\mathbf{\Omega}} \approx h_{\max}$, see Figure 6.23c, and becomes smoother for smaller $\ell_{\mathbf{\Omega}}$. The parameter $\ell_{\mathbf{\Omega}}$ appears in the coefficients of the stabilizations as presented in Table 6.1.




(a) Pressure with contours:
$\ell_{\mathbf{\Omega}} = 0.021 \approx h_{\max}$ (left), $\ell_{\mathbf{\Omega}} = 0.0062 \approx h_{\min}$ (center), and $\ell_{\mathbf{\Omega}} = 0.0001$ (right).



(b) Pressure – Stokes-Interface:
$\ell_{\mathbf{\Omega}} = 0.021 \approx h_{\max}$ (top),
$\ell_{\mathbf{\Omega}} = 0.0062 \approx h_{\min}$ (center),
and $\ell_{\mathbf{\Omega}} = 0.0001$ (bottom).

(c) Pressure – Darcy-Interface:
$\ell_{\mathbf{\Omega}} = 0.021 \approx h_{\max}$ (top),
$\ell_{\mathbf{\Omega}} = 0.0062 \approx h_{\min}$ (center),
and $\ell_{\mathbf{\Omega}} = 0.0001$ (bottom).

Figure 6.23: Example IV: Simulation results using $\alpha = \delta = 0.1$, $\rho = 0$, and different values of $\ell_{\mathbf{\Omega}}$.

Table 6.1: Stabilization coefficients in the subregions.

| Subregion | $(\mu_{\text{eff}}, \sigma)$ | GLS Coefficient | Grad-Div Coefficient |
|---|---|---|---|
| $\boldsymbol{\Omega}_f$ | $\left(10^{-3}, 0\right)$ | $\alpha\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}} = \alpha\frac{h_{\boldsymbol{T}}^2}{\mu_{\text{eff}}} = 10^3\alpha h_{\boldsymbol{T}}^2$ | $\delta\hat{\mu} = \delta\mu_{\text{eff}} = 10^{-3}\delta$ |
| $\boldsymbol{\Omega}_p$ | $\left(0, 10^7\right)$ | $\alpha\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}} = \alpha\frac{h_{\boldsymbol{T}}^2}{\sigma\ell_{\boldsymbol{\Omega}}^2} = \frac{10^{-7}}{\ell_{\boldsymbol{\Omega}}^2}\alpha h_{\boldsymbol{T}}^2$ | $\delta\hat{\mu} = \delta\sigma\ell_{\boldsymbol{\Omega}}^2 = 10^7\ell_{\boldsymbol{\Omega}}^2\delta$ |

In the porous region $\boldsymbol{\Omega}_p$ ($\mu_{\text{eff}} = 0$), the GLS stabilization term scales with $\left(\ell_{\boldsymbol{\Omega}}^2\right)^{-1}$ and the grad-div stabilization term is multiplied with $\ell_{\boldsymbol{\Omega}}^2$. This, together with the observed behavior with respect to $\ell_{\boldsymbol{\Omega}}$ suggests that an increase of $\alpha$ and a decrease of $\delta$ could yield a similar result as decreasing the value of $\ell_{\boldsymbol{\Omega}}$. The condition number tests in Figures 6.8 ($\mu_{\text{eff}} = \sigma = 1$) and 6.12 ($\sigma = 1000$) also suggest that this could have a positive impact on the conditioning.

Although a smaller value of $\delta$ resulted in a smoother solution in the vicinity of the interface, the oscillations could not be eliminated by solely adjusting $\delta$. Increasing the value of $\alpha$ has a smoothening effect, but spoils the pressure solution, in particular in the Stokes region, due to the strong imbalance of stabilization weights as described in Table 6.1.

Let us remark that numerical tests have shown that these oscillations are still present and even stronger if the mesh in Figure 6.21 is refined in the vicinity of the interface and $\ell_{\boldsymbol{\Omega}} \approx h_{\max}$. Also in this case the pressure is less oscillating for smaller values of $\ell_{\boldsymbol{\Omega}}$ which could not be managed by adjusting solely the stabilization parameters unless details of the solution are lost due to over stabilization.

Since $\ell_{\boldsymbol{\Omega}}$ is only present in the stabilization terms in the porous domain, we conclude that appropriate choices of the stabilization parameters do not coincide for $\boldsymbol{\Omega}_f$ and $\boldsymbol{\Omega}_p$, at least if the condition (5.2.6) shall be satisfied.

Due to the jump of the model parameters, also the artificial viscosity $\hat{\mu} := \mu_{\text{eff}} + \sigma\ell_{\boldsymbol{\Omega}}^2$, which is used in the method formulation (5.2.10) and weights the stabilization terms, jumps from $\hat{\mu}|_{\boldsymbol{\Omega}_p} = 10^7\ell_{\boldsymbol{\Omega}}^2$ in the porous region, to $\hat{\mu}|_{\boldsymbol{\Omega}_f} = \mu_{\text{eff}} = 10^{-3}$ in the pure fluid region. Smaller values of $\ell_{\boldsymbol{\Omega}}$ effectively damp the jump of $\hat{\mu}$ across the quasi-interface, see Table 6.1.

In summary, except for a small region enclosing the interface, the results of the Brinkman single-domain approach compare qualitatively well with the ones obtained with a multi-domain (Stokes–Darcy) method, which uses explicit interface conditions and does not necessarily consider continuity at the interface. In particular flow characteristics induced by the interface are captured well. However, for special interest in the interface region either a fine mesh combined with appropriate possibly domain-dependent stabilization parameters (in particular $\alpha$) have to be chosen, or the respective additional effort (explicit incorporation of interface conditions) shall be invested.

Let us point out that this example is beyond the theory provided in Chapter 5, where we have considered constant coefficients $\mu_{\text{eff}}$ and $\sigma$. Hence, this example suggests that the theory has to be extended for discontinuous coefficients representing Stokes–Darcy jumps.

## 6.4 Example V: Flow with Porous Obstacles

The following example is taken from [Dis04] and addresses a heterogeneous porous medium including three physically different settings.

A rectangular domain of length $L = 12\,\mathrm{m}$ and depth $H = 8\,\mathrm{m}$, i.e., $\mathbf{\Omega} := (0, 12) \times (0, 8)$, is considered. Within $\mathbf{\Omega}$, we distinguish between a free flow region $\mathbf{\Omega}_\mathrm{f}$, a moderately permeable region $\mathbf{\Omega}_{\mathrm{p},1}$, and a low permeability region $\mathbf{\Omega}_{\mathrm{p},2}$ with

$$\begin{aligned} \mathbf{\Omega}_f: &\qquad (\mu_\mathrm{eff}, \sigma) = \left(10^{-2}, 0\right), \\ \mathbf{\Omega}_{\mathrm{p},1}: &\qquad (\mu_\mathrm{eff}, \sigma) = \left(0, 10^{2}\right), \\ \mathbf{\Omega}_{\mathrm{p},2}: &\qquad (\mu_\mathrm{eff}, \sigma) = \left(0, 10^{6}\right), \end{aligned}$$

as visualized in Figure 6.24.

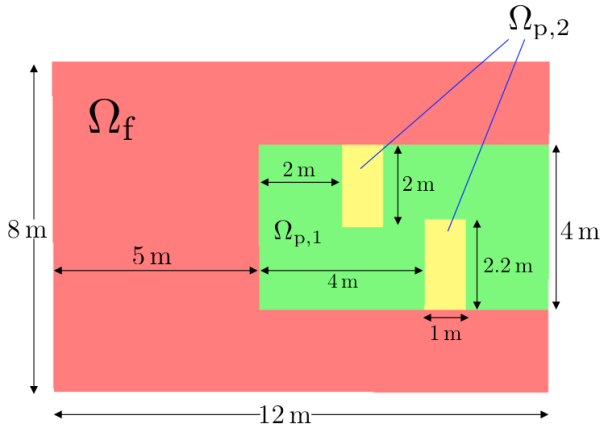



Figure 6.24: Example V: Sketch of the domain, showing the different physical regions $\Omega_\mathrm{f}$ (red), $\Omega_{\mathrm{p},1}$ (green) and $\Omega_{\mathrm{p},2}$ (yellow).

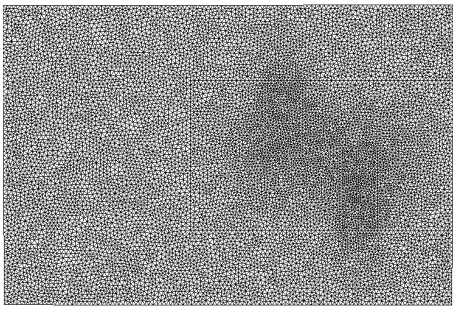
An outflow condition (natural boundary condition) is imposed on the right boundary, by setting

$$(\mu_\mathrm{eff} \nabla \boldsymbol{u} - p\mathbb{I}) \cdot \boldsymbol{n} = \mathbf{0}, \qquad \text{for } x = 12,$$

while essential boundary conditions are imposed via the penalty-free Nitsche method on the left (parabolic with maximum velocity $0.1\,\mathrm{m/s}$), top (homogeneous), and bottom (homogeneous) boundaries:

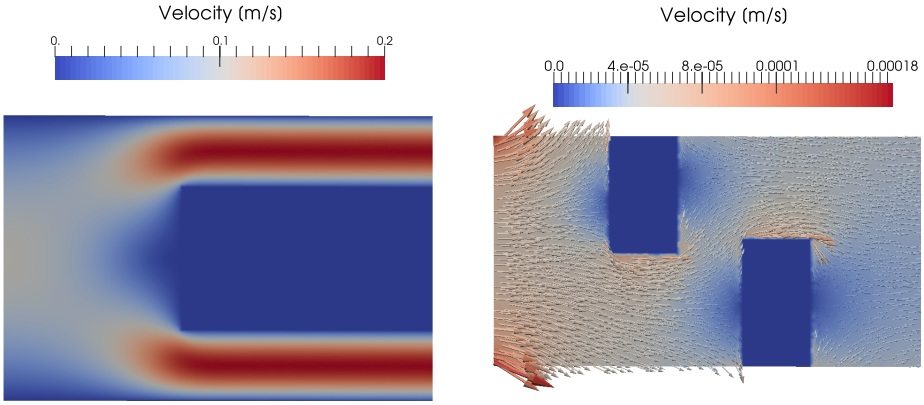
$$\begin{aligned} &\boldsymbol{u} = (u_1, u_2)^T = \left(\frac{1}{16}\left(-0.1y^2 + 0.8y\right), 0\right)^T, \qquad \text{for } x = 0, \\ &\boldsymbol{u} = \mathbf{0}, \qquad \text{for } y = 0, \text{ resp. } y = 8. \end{aligned}$$

The simulation has been performed on an unstructured, non-uniform mesh containing around 11K vertices, which has a slightly higher resolution at the implicit interfaces defined by the boundary of $\mathbf{\Omega}_{\mathrm{p},2}$ and is adapted to the interfaces, see Figure 6.25.

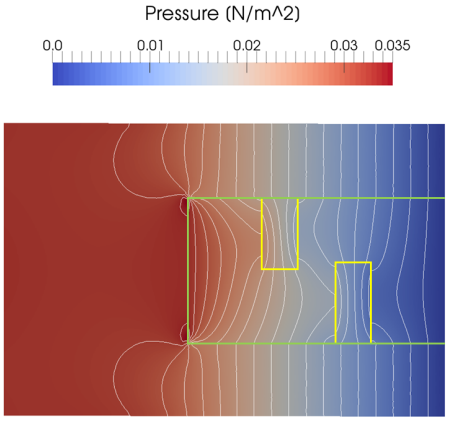


| $(h_{\min}, h_{\max})$ | # Cells (Triangles) | # Vertices |
|---|---|---|
| $(0.0598, 0.1699)$ | 22157 | 11255 |

Figure 6.25: Example V: The mesh, showing slightly refined regions.




(a) The magnitude of the velocity over the whole domain (left) and the magnitude of the velocity with velocity vectors within the porous region $\Omega_{\mathrm{p},1} \cup \Omega_{\mathrm{p},2}$ (right).




(b) Pressure isolines.

Figure 6.26: Example V: Simulation results for $\alpha = \delta = 0.1$, $\rho = 0$, and $\ell_{\mathbf{\Omega}} = 0.01$.

The parameters have been chosen as $\alpha = \delta = 0.1$, $\rho = 0$ (no Nitsche corners with $\mu_{\text{eff}} = 0$), and $\ell_{\mathbf{\Omega}} = 0.01$. The resulting magnitude of the velocity, shown in Figure 6.26a (left), visualizes that the velocity in the Stokes region $\mathbf{\Omega}_f$ is much higher than that in the porous regions, which therefore appears to be zero in the considered color scale. However, there is fluid flow present in the porous regions, which is demonstrated for $\mathbf{\Omega}_{p,1}$ in Figure 6.26a (right) with additional arrows (whose length is scaled by the absolute velocity) indicating the flow direction. The corresponding pressure approximation with isolines is depicted in Figure 6.26b. Altogether, the finite element solution is in line with the physically expected behavior.

Note that the value chosen for $\ell_{\mathbf{\Omega}}$ does not fulfill the condition (5.2.6). In Section 6.3 we have observed that smaller values of $\ell_{\mathbf{\Omega}}$ have a smoothening effect on the pressure in the vicinity of coefficient jumps of Stokes–Darcy type. To confirm this hypothesis here as well, we run an additional simulation for $\ell_{\mathbf{\Omega}} = 0.17 \approx h_{\max}$. As shown in Figure 6.27, the pressure is not well-behaved near the parameter jumps between Stokes regions and porous regions, although the corresponding velocity magnitude behaves qualitatively as shown in Figure 6.26a.

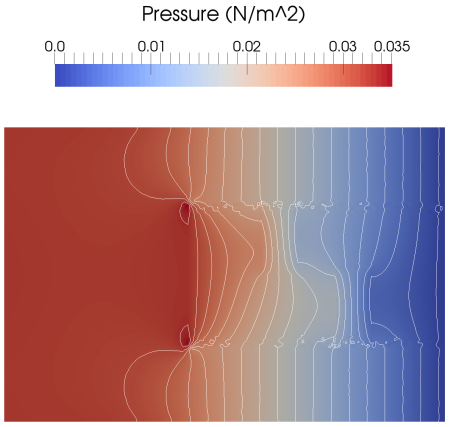



Figure 6.27: Example V: Pressure with isolines for $\alpha = \delta = 0.1$, $\rho = 0$, and $\ell_{\mathbf{\Omega}} = 0.17$.

Let us finally note that - similar as for Example IV - with the premise to obtain qualitatively similar solutions as shown in Figure 6.26, a smooth pressure approximation could not be produced with a variation in the stabilization parameters ($\alpha$ and $\delta$) solely, i.e., keeping $\ell_{\mathbf{\Omega}} = 0.17$. As already indicated, a possible explanation is that $\ell_{\mathbf{\Omega}}$ has an influence only in porous regions ($\sigma \neq 0$), while the stabilizations (controlled via $\alpha$ and $\delta$) affect the formulation in the whole domain. Hence, $\ell_{\mathbf{\Omega}}$ influences the relation of stabilization weights for Stokes regions ($\sigma = 0$) neighboring porous regions ($\sigma \neq 0$).

## 6.5 Summary

In order to complement the theory developed in Chapter 5 with computational results, several numerical studies have been presented. To that end, five examples were conducted for assessing the performance of the proposed method (5.2.10).

Examples I and II have confirmed the predicted convergence order (Thm. 5.3.21) in the Stokes regime as well as in the Darcy limit. We have also seen how the penalty-free Nitsche method acts in case of boundary layers compared to strongly imposed essential boundary conditions. Moreover, a parameter study and tests with respect to the condition number have addressed the impact of the stabilization parameters on the results. Concerning the influence of the stabilization parameters on the magnitude of the errors, an extracted general tendency is to use small values of $\alpha$, although the conditioning suggested often rather the opposite. Concerning $\delta$ and $\ell_{\mathbf{\Omega}}$, optimal choices seem to be more interwoven with each other, with the physical setting, and with parts of the error in the energy norm.

Example III was considered in order to compare the results of the method (5.2.10) with those for stabilization methods that are available in the literature. This has supported the competitiveness of the proposed formulation. Furthermore, the sensitivity with respect to the GLS parameter was tested, which has demonstrated a negative impact of too small and too large values of $\alpha$ on the quality of the (pressure) solution.

Examples IV and V considered multi-physics problems with discontinuous coefficients of Stokes–Darcy type. Although beyond the theory developed in Chapter 5, these were used in order to investigate the applicability of the method (5.2.10) in mixed flow domains. While the pressure slightly away from the interface and the velocity yielded results, comparable with those in the literature, the pressure at the interface was identified to be non-smooth. In a series of tests, we have observed that decreasing the value of $\ell_{\mathbf{\Omega}}$ (violating the condition (5.2.6)) allows to reduce the oscillations (a subdomain-dependent GLS parameter would have a similar effect).

The corner stabilization parameter was not observed to have a significant impact on the results of the considered examples, i.e., reasonable variations of $\rho$ did not affect the results.

# 7 Geothermal Energy Production – Modeling, Simulation, and Optimization for Hot Sedimentary Aquifers

Stored heat in the subsurface is recognized as *geothermal energy* and can be used for electricity production and district heating. As geothermal energy is a renewable resource, its usage can contribute to reduce the carbon footprint. This chapter discusses numerical modeling, simulation, and optimization of geothermal energy extraction as an application of the proposed method (5.2.10).

The basic functionality of an *open loop geothermal plant* can be explained by a minimal example, a so-called *doublet*, which refers to a pairing of an *injection well* and a *production well*. While water with a certain temperature is pumped into a *geothermal reservoir* (hot aquifer) at the injection well, water with a higher temperature is extracted via the production well. This procedure allows to gain energy in form of heat.

One of the main goals in geothermal development is the maximization of the economic outcome, which in turn involves maximizing the net extraction of geothermal energy.

Given a site with sound geothermal and hydrogeological conditions, the possible placement of wells is generally geographically limited to a so-called concession field which describes the area of planning permission. In particular, thermo-hydraulic encroachment into neighboring concession fields has to be prevented during the whole production time.

Furthermore, with an increasing density of geothermal installations, effects due to thermal and hydraulic interference gain importance - negative interference shall be avoided while positive interference shall be promoted.

Therefore, predictive simulations are decisive for optimal or at least reasonable placing of new wells. Among the key questions, besides economic factors like operating costs, is the optimization of (multi-)well configurations for a considered time period of economic utilization. This concerns the number, types, locations, and spacing of wells. Regarding the economic lifetime of a geothermal plant based on open loops, special attention has to be paid to the *thermal breakthrough* (the time when the temperature at a production well starts to decrease) and possible *thermal short circuits*. Hence, the spatio-temporal evolution of the cooling front emanating from the injection wells needs to be predicted. It is controlled by the geothermal and hydrogeological conditions (including porous, fractured, and sometimes even karstified geothermal reservoirs), the *exploitation*

*scheme* (induced flow rates), and possible thermo-hydraulic interaction with wells in the vicinity.

In summary, optimal field development and production management is a multi-variable optimization task with numerous parameters influencing the optimization process.

The developed numerical model, containing the method (5.2.10) as a constituent, is intended to describe the thermo-hydraulics in an essentially horizontal geothermal reservoir (hot aquifer) with multi-well configurations. It shall be further utilized in an optimization procedure in order to detect advantageous placements of geothermal wells concerning the net energy gain. Considering a two-dimensional setting is fit for purpose and advantageous (less expensive) from the computational perspective. It enables the study of diverse scenarios of multiple geothermal well arrangements for varying geothermal and hydrogeological conditions in a reasonable time frame.

Section 7.1 starts with a presentation of the main ingredients of the computational framework. Based on a description of the computational domain, we will use the Brinkman model as groundwater flow model, an immersed boundary method for the wells, and the heat transport equation.

In Section 7.2 we formulate the spatially and temporally discretized problem and perform a preliminary code and model validation focusing on a single well, respectively a doublet, investigating, in particular the sensitivity with respect to the spatial and temporal discretization. Finally we specify the objective function for the optimization procedure.

Section 7.3, is dedicated to the numerical study of the optimal placement of two specific, patterned multi-well arrangements, namely lattice-type and hexagonal configurations.

For an introduction into the topic and more information see, e.g., [SB13], [DiP08], [Gla10], [PVC18], [Ree16], and [Ern10].

The software tools used in this chapter are the finite element library ParMooN [Gan+16], [Wil+17a] and the library NLopt [Nlo] (integrated into ParMooN), which are both free and open-source. The meshes are constructed with the free open-source mesh generator Gmsh, [GR09], and the results are visualized with ParaView, [AGL05].

The content of this chapter was partly developed in collaboration with Dr. Ernesto Meneses Rioseco (Leibniz Institute for Applied Geophysics) and financially supported by a Seed Grant of the Leibniz Mathematical Modeling and Simulation (MMS) Network.

## 7.1 The Modeling Approach

This section describes the details of the numerical model. Its two fundamental components are:

(1) The Brinkman equations (3.2.8). They describe one-phase, incompressible, steady, laminar groundwater flow through a confined and fully saturated aquifer. The aquifer is predominantly composed of consolidated sedimentary material (as, e.g., sandstone and/or carbonates) and considered as an isotropic, static (nondeformable) porous medium (see Subsection 2.2.1). Compared to the frequently used Darcy model, the additional term accounting for viscous stresses in the momentum balance allows to go beyond the range of validity of Darcy's law towards regimes of higher permeability. This is particularly interesting for highly damaged zones in faulted domains and vuggy porous media manifested by karstification (see, e.g., [PEQ09] and [Joo$^+$09]).

(2) The heat transport equation (energy balance). It models the evolution of cold water fronts in hot aquifers and is of advection-diffusion(-reaction) type. The advection is determined by the geothermal reservoir exploitation scheme.

The two model problems are sequentially coupled (no feedback) for each optimization step, i.e., the Brinkman solution for the velocity determines the advective field in the heat equation.

A further important ingredient is a *non-matching (immersed) boundary approach* used to incorporate the boundary conditions at the wells represented as *singular* forces.

Regarding the *optimization* of geothermal energy extraction, our primary goal is to describe and quantify the thermo-hydraulic effects on the placement and spacing of different geothermal multi-well patterns with varying geothermal and hydrogeological conditions. Therefore, we concentrate on control variables that specify the arrangement of the wells.

One of the important capabilities of the proposed methodology is the possibility to principally consider arbitrary well locations in the geothermal reservoir during the optimization process. This potentially increases the flexibility of the optimization by reducing spatial restrictions on the feasible region (caused by mesh-adapted well models). In fact, the immersed approach used to incorporate the boundary conditions at the wells does not require the regeneration of the computational mesh during the optimization, since the imposition of boundary conditions is decoupled from the spatial discretization. From the practical viewpoint, the latter is a major advantage especially when the simulation of numerous scenarios of (multi-)well arrangements is intended.

**Remark 7.1.1** (Literature Overview)**.**
*Let us review some publications that dealt with similar questions:*

- *In [JS89, Sect. 3.5], multi-well arrangements of lattice-type were compared with single doublets and reasoned to be advantageous in terms of the occurrence time of the thermal breakthrough.*

- *A similar methodology as in this thesis has been used in, e.g., [Die14, Sect. 9.2.3] and [Wil+17b], but only concerning the primal Darcy problem (pressure Poisson problem, see Rem. 3.3.4). In [Die14] also an immersed approach is described, but is there restricted to the case of point-associated wells located at nodes of the mesh.*
- *The modeling of hot sedimentary aquifers was also investigated in [Wil+17c] and [Cro+16], taking into account the impact of sandstone reservoir heterogeneities on geothermal doublets production performance. In [Wil+17b], doublets and sizes of geothermal concession fields are discussed.*
- *Reservoir lifetimes for doublets in hot aquifers were studied in [Sae+14] and [Sae+15] as well as in [Wil+17b], including two lattice-type configurations in homogeneous domains.*
- *Very recently, automatic optimization methods based on genetic algorithms have been proposed in [RJZ19] and [Zha+19], considering heterogeneous reservoirs but limited to Cartesian meshes. Further, in [Kah+19], a grid-block-based optimization strategy with a stochastic (*10*-ensemble gradient-based) optimization is tested using (partly) the free open-source software* Open Porous Media *(OPM).*

*Let us note that most publications refer to commercial software packages as* FEFLOW *or* COMSOL Multiphysics*.*

### 7.1.1 The Geothermal and Hydrogeological Setting

Geothermal district heating development has been gaining momentum in Europe with a significant installed capacity and numerous projects currently under development, see, e.g., [Ber+17], [ABS16], and [LB16]. Especially, deep geothermal industry in the Greater Munich region in (the porous, fractured, and karstified Upper Jurassic carbonates of) the South German Molasse Basin, has been rapidly growing in the last decades, [FSS16], [Dus+16], [WBM19]. Due to this remarkable development and the related increase in the density of planed geothermal wells, the focus of this chapter will be on similar geothermal and hydrogeological conditions as encountered in the Upper Jurassic (Malm) aquifer in the Greater Munich region.

Let us mention that, although the physical setups considered in the following are concerned with similar reservoir conditions as found in the Upper Jurassic aquifer in the Munich region, the methodology that will be developed in the subsequent sections can be applied (and extended) to any hot sedimentary aquifer.

Let us begin with a brief characterization of the geothermal and hydrogeological conditions, normally encountered in hot sedimentary aquifers and in particular in the South German Molasse Basin. The crucial parameters for the optimal placement and spacing of geothermal wells are the temperature distribution, the natural hydraulic activity, and the permeability structure in a geothermal reservoir. Further details can be found in, e.g., [Dus+16] and [Dus+18].

### The Temperature Distribution

The considered Malm in the Greater Munich region reaches depths of approximately 5 km and an average geothermal gradient of around 30 °C per kilometer depth is established, such that temperatures up to 150 °C are reached. While the aquifer is in a quite constant depth in the direction east-west, an (approximately linear) inclination is found from south to north, see [Dus+16, Fig. 3]. This translates into a non-zero, constant temperature gradient in the simulations (in Section 7.3), where we will consider either a constant ambient temperature of 100 °C or a linear temperature distribution that corresponds to a slightly declined aquifer with a height difference of approximately 71m per km length.

### Natural Hydraulic Activity

The regional and local natural groundwater flow in hot sedimentary aquifers is in general overprinted by the relatively high flow rates, imposed by multi-well arrangements, [JS89]. Due to the still large uncertainties with regard to the current groundwater flow regime in the Malm aquifer and the very low hydraulic gradient, a constant underlying pressure field can alternatively be considered for the thermo-hydraulic modeling, [Dus+16].

### Hydraulic Conductivity Ranges and Permeability Structures

The aquifer in the Munich region is a geothermal reservoir with a highly variable permeability structure (typical for carbonate reservoirs) which consists of porous, fractured, and karstic carbonates of the Upper Jurassic Malm. A multidisciplinary geothermal reservoir characterization, recently conducted in [Dus+16, Table 1], presents a wide range of hydraulic conductivity values in the order of $10^{-4} - 10^{-7}$ m/s.

The Malm aquifer is confined by overriding cretaceuos and tertiary layers and an underlying crystalline basement [Dus+16, Fig. 3], which are considered as hydraulically non-conductive (*aquitards*).

Primary geologic controls on the permeability structure of carbonate geothermal reservoirs constitute structural elements as *faults or damage zones* (caused by tectonic activity) and facies distribution (caused by historical geographic development), [Dus+18], [Mra+19]. Note that karstification might even lead to solution conduits in carbonate reservoirs, [Har+14]. In [Dus+16, Fig. 10], a simplified hydrostratigraphic profile of the Upper Jurassic aquifer is given, revealing a layer-like *laterally varying* (bedded) permeability.

Such structural properties will be accounted for in the numerical study in order to assess their effect on the optimized well placement. To be precise, we will consider aquifers with laterally varying hydraulic conductivity (respectively permeability) in Subsection 7.3.3 and aquifers with faults/damage zones in Subsection 7.3.4.

### 7.1.2 Model Assumptions

The model is build upon the following main assumptions and approximations:

(A1) We assume to deal with an aquifer confined by overriding and underlying aquitards. As a consequence, fluid flow through the top and bottom boundaries of the reservoir is neglected. Moreover, assuming that the vertical dimension is much smaller than the horizontal characteristic size, we neglect the effect of gravitational forces.

(A2) We consider the flow of a single phase, incompressible (constant density), Newtonian fluid in an isotropic, saturated, non-deformable aquifer (porous medium).

(A3) We neglect the dependence of the fluid viscosity and density on the temperature.

(A4) The heat transport model is confined to the aquifer under investigation, i.e., heating due to the temperature of the aquitard is neglected.

(A5) Since the longitudinal dispersion is usually several times higher than the transversal dispersion, we neglect the transversal dispersion, assuming that its effect does not play a relevant role in the considered two-dimensional model, and use a constant heat capacity.

(A6) The wells are assumed to be vertical boreholes (no materialized surface) of cylindrical shape, perforating the aquifer.

### 7.1.3 The Aquifer

The domain of interest is an aquifer of constant thickness $H$ [m] with rectangular base of diameter $L$ [m] ($H \ll L$). It is perforated by $N^{inj}$ injection wells and $N^{prod}$ production wells.

A generic well $w_{3D}^{(\star)}$ is modeled as a vertical, cylindrical borehole. Denoting by

$$w^{(\star)} := w_{2D}^{(\star)}$$

its circular cross-section, the volume of $w_{3D}^{(\star)}$ is given by $A_{w^{(\star)}} \cdot H$, where $A_{w^{(\star)}}$ [m$^2$] is its cross-sectional area. The setup for a single well is schematically visualized in Figure 7.1 (left).

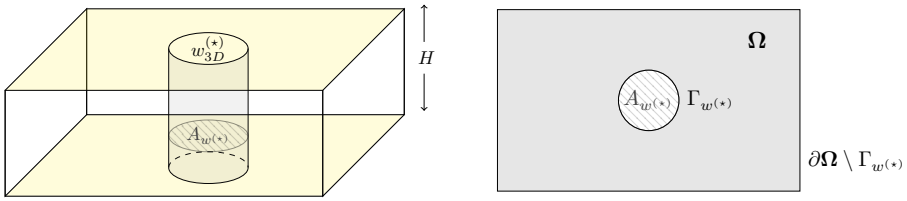



Figure 7.1: Left: Schematic 3D model of an aquifer with a single cylindrical well $w_{3D}^{(\star)}$. Top and bottom of the aquifer (yellow) are impermeable for fluid. Right: The considered horizontal cross-section of the aquifer.

Production wells and injection wells will notationally be distinguished by using the respective symbol $(\star) \in \{\text{prod}, \text{inj}\}$. The same notation will be adopted for other quantities that are associated to production and injection wells, respectively.

In the horizontal direction, the fluid is allowed to permeate through the aquifer boundaries, while eventual phenomena in the vertical direction are neglected. Hence, for the remainder of this work, we focus on an aquifer slice $\boldsymbol{\Omega} \subset \mathbb{R}^2$ parallel to the confining planes as shown in Figure 7.1 (right). Further, the boundary of a generic, circular well slice $w^{(\star)} \subset \mathbb{R}^2$ will be referred to as $\Gamma_{w^{(\star)}}$.

### 7.1.4 Groundwater Flow - The Brinkman Equations

Assuming conservation of mass and (linear) momentum, the flow is modeled by the stationary Brinkman equations (3.2.8a)-(3.2.8b) with the following boundary conditions: On the outer boundary of the computational domain $\boldsymbol{\Omega}$, natural (Neumann) boundary conditions are imposed:

$$(\mu_{\text{eff}} \nabla \boldsymbol{u} - p\mathbb{I}) \cdot \boldsymbol{n} = \boldsymbol{0}, \quad \text{on } \partial\boldsymbol{\Omega} \setminus \bigcup_{w^{(\star)}} \Gamma_{w^{(\star)}}, \tag{7.1.1a}$$

where $\boldsymbol{n}$ stands for the outer unit normal vector on $\partial\boldsymbol{\Omega} \setminus \bigcup_{w^{(\star)}} \Gamma_{w^{(\star)}}$. Along each well boundary, a constant velocity magnitude $U_{w^{(\star)}}$ [m/s] directed normally to the well boundary is assumed, yielding the Dirichlet boundary conditions

$$\boldsymbol{u} = \pm U_{w^{(\star)}} \boldsymbol{n}_{w^{(\star)}}, \quad \text{on } \Gamma_{w^{(\star)}}, \quad \forall w^{(\star)}, \tag{7.1.1b}$$

where $\boldsymbol{n}_{w^{(\star)}}$ is pointing from the well into the porous medium (inner unit normal vector). Positive and negative signs in front of $U_{w^{(\star)}}$ shall be used to represent flow into (injection) or out of (production) the aquifer with respect to the well, respectively.

The magnitude of the in-/outflow velocity, $U_{w^{(\star)}}$ in (7.1.1b), depends on the prescribed injection respectively production rate and the thickness $H$ [m] of the three-dimensional aquifer (see Subsection 7.1.3). Let us assume that the fluid is injected respectively extracted uniformly along the vertical direction of the generic cylindrical well $w_{3D}^{(\star)}$ with flow rate $Q_{w_{3D}^{(\star)}}[\text{m}^3/\text{s}]$. Then the magnitude of the velocity along the well boundary for the two-dimensional problem is given by

$$U_{w^{(\star)}} = \frac{Q_{w_{3D}^{(\star)}}}{2\pi r_{w^{(\star)}} H}, \quad \text{on } \Gamma_{w^{(\star)}}, \tag{7.1.2}$$

with $r_{w^{(\star)}}$ [m] denoting the radius of the well.

For the considered two-dimensional case, external volume forces, as well as mass sources within the aquifer vanish, i.e., $\boldsymbol{f} = \boldsymbol{0}$ and $g = 0$ in (3.2.8a)-(3.2.8b).

Let us note, that in Subsection 7.1.5, an immersed boundary method will be proposed to efficiently incorporate the wells in form of singular sources and sinks in equation (3.2.8b).

**Remark 7.1.2** (On the Boundary Condition (7.1.1a))**.**
*As discussed in Subsection 7.1.1, natural groundwater flow in the geothermal reservoir can be usually neglected in comparisons to the flow induced by injection*

*and production wells. As a consequence, we will consider a constant pressure* $P = 0\,\mathrm{Pa}$ *at the aquifer boundaries yielding the condition* (7.1.1a)*. Otherwise, one could set the external pressure equal to a given function (depending on the depth of the aquifer) and consider the boundary condition*

$$(\mu_{\text{eff}}\,\nabla \boldsymbol{u} - p\mathbb{I}) \cdot \boldsymbol{n} = P\boldsymbol{n}, \qquad \text{on } \partial\boldsymbol{\Omega} \setminus \bigcup_{w^{(\star)}} \Gamma_{w^{(\star)}},$$

*instead of* (7.1.1a)*.*

### 7.1.5 The Immersed Well Model

In practice, the radii of the cylindrical wells, e.g., 0.1 m in [Wil+16], are much smaller than the scale of the domain (several kilometers). Setting Dirichlet boundary conditions as (7.1.1b) on such small portions of the domain increases the (geometric) complexity of the problem due to the need of resolving a wide range of spatial scales within the computational mesh.

In order to circumvent this issue, a so-called *immersed boundary method (IBM)*, originating from [Pes02], will be employed. The basic idea of IBMs is to 'immerse' complex geometries in simpler geometries by extending the computational domain and introducing appropriate force terms in the model equations. From the computational point of view, one of its main advantages is that coarser spatial discretizations can be used, thus reducing the computational effort for the numerical simulation.

In the groundwater flow context, IBMs are considered also in [Die14]. However, in this case the singular problem is derived from the primal Darcy formulation (pressure Poisson problem, see Rem. 3.3.4), instead of the mixed, more general Brinkman problem. Further fields of application include, e.g., the simulation of perfusion within biological tissues, [D'A12].

The immersed well model is based on considering an extended domain, which includes the well regions. The wells are then described as singular forces defined in single points (the well centers), which can be regularized on the discrete level utilizing smooth approximations.

For simplicity of notation, we investigate in detail the case of a single well $w^{(\star)}$, as visualized in Figure 7.1 (right). In the limit case, when $r_{w^{(\star)}}$ tends to zero, the well $w^{(\star)}$ can be seen as a singular source (or sink) of mass in the continuity equation (3.2.8b), formulated in the extended domain $\boldsymbol{\Omega} \cup w^{(\star)} \subset \mathbb{R}^2$ (see Figure 7.1 right), i.e.,

$$\nabla \cdot \boldsymbol{u} = g_{w^{(\star)}} \delta_{w^{(\star)}} \qquad \text{in } \boldsymbol{\Omega} \cup w^{(\star)}. \tag{7.1.3}$$

Here, $\delta_{w^{(\star)}}$ denotes the Dirac delta distribution with respect to the well center and $g_{w^{(\star)}}$ is a proper constant which depends on the prescribed injection respectively production rate.

In order to determine the singular force $g_{w^{(\star)}}$, we note that the solution of the problem with the original boundary conditions (7.1.1b) satisfies

$$\int_{w^{(\star)}} \nabla \cdot \boldsymbol{u} = \int_{\Gamma_{w^{(\star)}}} \boldsymbol{u} \cdot \boldsymbol{n}_{w^{(\star)}} = \pm U_{w^{(\star)}} 2\pi r_{w^{(\star)}} = \pm \frac{Q_{w_{3D}^{(\star)}}}{H}, \tag{7.1.4}$$

where we have used the Gaussian theorem, the boundary condition (7.1.1b), the fact that $|\Gamma_{w^{(\star)}}| = 2\pi r_{w^{(\star)}}$, and the expression (7.1.2) for $U_{w^{(\star)}}$. On the other hand, integrating (7.1.3) over a circle of radius $r_{w^{(\star)}}$, centered at the center of $w^{(\star)}$, using (7.1.4), the fact that $g_{w^{(\star)}}$ is constant, and the property $\int_{w^{(\star)}} \delta_{w^{(\star)}} = 1$ of the Dirac delta distribution, we obtain

$$g_{w^{(\star)}} = \int_{w^{(\star)}} g_{w^{(\star)}} \delta_{w^{(\star)}} = \int_{w^{(\star)}} \nabla \cdot \boldsymbol{u} = \pm \frac{Q_{w_{3D}^{(\star)}}}{H}\,. \tag{7.1.5}$$

Hence, generalizing this ansatz to multiple wells, we consider the Brinkman problem (3.2.8a)-(3.2.8b) defined on

$$\widehat{\boldsymbol{\Omega}} := \boldsymbol{\Omega} \cup \left( \bigcup_{w^{(\star)}} w^{(\star)} \right) \subset \mathbb{R}^2, \tag{7.1.6}$$

meanwhile replacing the Dirichlet boundary conditions (7.1.1b) by a modified continuity equation of the form

$$\nabla \cdot \boldsymbol{u} = \sum_{l=1}^{N^{inj}+N^{prod}} g_{w_l^{(\star)}} \delta_{w_l^{(\star)}}. \tag{7.1.7}$$

**Remark 7.1.3** (On the Singular Mass Source)**.**
*The modified mass conservation equation* (7.1.7) *states that the divergence of the velocity vanishes everywhere except for injection and production wells (a set of given points), where it has a (positive resp. negative) value obtained from* (7.1.5).

### 7.1.6 Heat Transport – The Advection-Diffusion Equation

The model for heat transport can be obtained from the standard energy conservation equation [Sae+14], resulting in a time-dependent advection-diffusion equation for a temperature field $T(t, \boldsymbol{x})$ [K]:

$$\frac{\partial (\rho C T)}{\partial t} - \nabla \cdot (\lambda \nabla T) + \rho_f C_f \boldsymbol{u} \cdot \nabla T = 0 \quad \text{in } (0, t^L] \times \boldsymbol{\Omega}, \tag{7.1.8a}$$

where it is assumed that heat creation through friction in the well and heat loss/gain from the outside can be neglected (yielding a right-hand side equal to zero). We complete (7.1.8a) with the initial condition

$$T(0, \boldsymbol{x}) = T_0(\boldsymbol{x}) \quad \text{in } \boldsymbol{\Omega},$$

and the (steady, Dirichlet) boundary conditions

$$\begin{aligned} T &= T_{w_k^{inj}}, && \text{on } \Gamma_{w_k^{inj}},\ \forall k, \\ T &= T_0, && \text{on } \partial\boldsymbol{\Omega} \setminus \bigcup_k \Gamma_{w_k^{inj}}. \end{aligned}$$

Here, $k = 1, \ldots, N^{inj}$, $N^{inj}$ is the number of injection wells $w_k^{inj}$, $T_{w_k^{inj}}$ [K] are the corresponding injection temperatures, and $T_0$ [K] refers to the initial

temperature field, in this context called *formation or aquifer temperature*. The filed $\boldsymbol{u}$ is the groundwater flow velocity, $t$ [s] is the time variable, and $t^L$ [s] denotes the upper bound on the time interval (coinciding with the maximum operation time of the geothermal installation). In (7.1.8a) there are further coefficients associated to the fluid (index $f$) respectively the porous structure (index $s$), namely the densities $\rho_f, \rho_s$ [$^{\mathrm{kg}}/_{\mathrm{m}^3}$] and the material *specific heat capacities* $C_f$, $C_s$ [$^{\mathrm{J}}/_{\mathrm{kg\,K}}$], which are combined to the *volumetric (macroscopic) heat capacity* $\rho C$ given by

$$\rho\, C := (1-\phi)\, \rho_s\, C_s + \phi \rho_f\, C_f,$$

where $\phi$ denotes the porosity.

The total *thermal conductivity* tensor $\lambda$ [$^{\mathrm{W}}/_{\mathrm{m\,K}}$] can be modeled (see, e.g., [Sch61] or [Cro$^+$16]) as the sum of the *equivalent conductivity* $\lambda_{eq}\mathbb{I}$ and *thermal dispersion* $\lambda_{dis}$, i.e.,

$$\lambda = \lambda_{eq}\mathbb{I} + \lambda_{dis}$$

with

$$\begin{aligned}\lambda_{eq} &= (1-\phi)\,\lambda_s + \phi\lambda_f,\\ \lambda_{dis} &= \rho_f C_f \left(\alpha_T |\boldsymbol{u}|\mathbb{I} + (\alpha_L - \alpha_T)\frac{\boldsymbol{u}\boldsymbol{u}^T}{|\boldsymbol{u}|}\right),\end{aligned}$$

where $\lambda_s$ and $\lambda_f$ are the (scalar) thermal conductivities of the indexed species, $\alpha_L$ [m] is the *longitudinal thermal dispersion*, and $\alpha_T$ [m] denotes the *transversal thermal dispersion*.

As mentioned in (A2), (A3), and (A5), the transversal dispersion will be neglected and the constant densities and heat capacities can be excluded from the temporal and spatial derivative in (7.1.8).

### 7.1.7 A Penalty-Based Immersed Injection Well Model

Similar to Subsection 7.1.5 we also consider a modified problem for the temperature field, formulated in the extended domain $\widehat{\boldsymbol{\Omega}}$ defined in (7.1.6). In order to fulfill the desired Dirichlet boundary condition on the well boundary, we propose a penalty method inspired by [Ang99], which explicitly adds terms to the equation (7.1.8a), that penalize the temperature differences between the (injection) well regions and the corresponding injection temperatures, respectively. Here it is crucial to note, that the temperature field is solely disturbed at the injection wells, such that only these positions have to be included as heat sources.

The resulting penalty-based immersed heat transport problem, which approximates the original problem (7.1.8) in the extended domain $\widehat{\boldsymbol{\Omega}}$, reads

$$a\frac{\partial T}{\partial t} - \nabla\cdot(d\nabla T) + \boldsymbol{u}\cdot\nabla T + \sum_{k=1}^{N^{inj}} \widetilde{\gamma}_k \left(T - T_{w_k^{inj}}\right) = 0 \quad \text{in } (0,t^L]\times\widehat{\boldsymbol{\Omega}}, \tag{7.1.9a}$$

$$T(0,\boldsymbol{x}) = T_0(\boldsymbol{x}) \quad \text{in } \widehat{\boldsymbol{\Omega}}, \tag{7.1.9b}$$

$$T = T_0 \quad \text{on } \partial\widehat{\boldsymbol{\Omega}}, \tag{7.1.9c}$$

with

$$a := \frac{\rho C}{\rho_f C_f}, \quad d := \frac{\lambda}{\rho_f C_f}, \quad \text{and} \quad \widetilde{\gamma}_k := \begin{cases} 0, & \text{in } \mathbf{\Omega}, \\ \gamma_k, & \text{in } w_k^{inj}, \end{cases} \tag{7.1.9d}$$

for a well-wise penalty parameter $\gamma_k > 0$.

## 7.2 The Numerical Method and the Optimization Approach

The automatic, energetic optimization shall be realized on the basis of sequentially coupling the discrete flow problem with the discrete heat balance. Let us therefore introduce the discretized subproblems, validate the proposed method, and define the objective function for the optimization.

### 7.2.1 Discretization in Time and Space

The Brinkman equations and the heat transport equations are discretized in space using the finite element method. To this purpose, let us consider a shape-regular (see (4.1.2)) triangulation $\mathcal{T}_h$ of the domain $\widehat{\mathbf{\Omega}}$ and recall the notation introduced in Chapters 4 and 5.

#### The Approximation of Singular Sources and Sinks

At the discrete level, there are different approaches to include the singular sources and sinks in (7.1.7) in a numerical method.

We adopt a *non-matching* approach, decoupling the singular points from the computational mesh. In practice, the singular term is defined in such a way that the physical solution (defined only outside the well regions) is continuously extended inside the wells. To this purpose, the right-hand side of (7.1.7) is approximated using a discrete version of the Dirac delta distribution, with support on a small neighborhood of the well center. A classical approximation $\delta^{r_\varepsilon}$ (see also [Pes02]) is given by

$$\delta^{r_\varepsilon}_{w^{(\star)}}(x,y) := \frac{\pi}{r_\varepsilon^2\,(\pi^2-4)}\theta\left(\frac{\sqrt{(x-x_{w^{(\star)}})^2+(y-y_{w^{(\star)}})^2}}{r_\varepsilon}\right) \tag{7.2.1}$$

with

$$\theta(r) := \begin{cases} \cos(\pi r)+1, & \text{if } -1 < r < 1, \\ 0, & \text{otherwise,} \end{cases}$$

and $(x_{w^{(\star)}}, y_{w^{(\star)}})$ denoting the center of the well $w^{(\star)}$. In (7.2.1), $r_\varepsilon > 0$ is an arbitrary (small, compared to the domain size, i.e., $r_\varepsilon \ll \operatorname{diam}(\widehat{\mathbf{\Omega}})$) parameter that can be chosen depending on the well radius and on the suitable spatial discretization (near the well). The discrete delta function will then vanish for the points $(x,y)$ with distance from $(x_{w^{(\star)}}, y_{w^{(\star)}})$ greater than $r_\varepsilon$. Notice that the function defined in (7.2.1) also satisfies

$$\int_{\widehat{\mathbf{\Omega}}} \delta^{r_\varepsilon}_{w^{(\star)}} = 1, \quad \forall r_\varepsilon > 0.$$

For a visualization of $\delta^{r_\varepsilon}$ for different values of $r_\varepsilon$, see Figure 7.2.

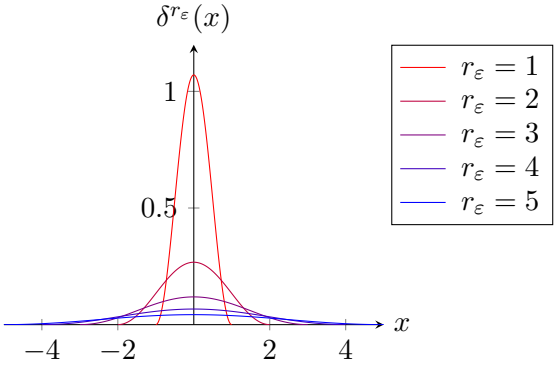



Figure 7.2: Visualization of (7.2.1) for $y = y_{w^{(\star)}}$ and $x_{w^{(\star)}} = 0$.

As a consequence of utilizing an approximation as (7.2.1), the numerical solution will have a physical meaning only at a distance from the well center greater than or equal to $r_\varepsilon$ (dictated by the support of the approximate delta function). However, this way we obtain the approximate flow field without the need of accurately resolving the fluid dynamics near the sources and sinks.

**Remark 7.2.1** (On the Non-Matching Immersed Boundary Approach)**.** *Another possibility to include the singular sources and sinks in the discrete approach is to assume that the singular points (well centers) coincide with vertices in 2D (or edges, in 3D) of the considered computational mesh (see, e.g., [D'A12] and [Die14]). This choice, however, links the (singular) wells to the mesh generation.*

*The non-matching approach instead allows to arbitrarily change the position of the well within the computational domain without the need of generating a new computational mesh. This feature will be substantial in the optimization procedure related to the (arbitrary) optimal placement of the wells.*

### The Discrete Brinkman Equations

The finite element approximations of the velocity and the pressure, $\boldsymbol{u}_h$ and $p_h$, are sought in the space of continuous, triangle-wise linear polynomials (5.2.9) with $\boldsymbol{\Omega}$ replaced by $\widehat{\boldsymbol{\Omega}}$.

Based on the discussions in Subsections 7.1.4 and 7.1.5, the stabilized finite element method (5.2.10) for the Brinkman problem (3.2.8a)-(3.2.8b) reads: Find $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h^1 \times Q_h^1$ such that

$$A_h\left[(\boldsymbol{u}_h, p_h)\,;(\boldsymbol{v}_h, q_h)\right] = L_h\left(\boldsymbol{v}_h, q_h\right), \quad \forall\ (\boldsymbol{v}_h, q_h) \in \boldsymbol{V}_h^1 \times Q_h^1, \tag{7.2.2}$$

with

$$\begin{aligned} A_h\left[(\boldsymbol{u},p)\,;(\boldsymbol{v},q)\right] &:= \mu_{\text{eff}}\left(\nabla\boldsymbol{u},\nabla\boldsymbol{v}\right)+\sigma\left(\boldsymbol{u},\boldsymbol{v}\right)-\left(p,\nabla\cdot\boldsymbol{v}\right)+\left(\nabla\cdot\boldsymbol{u},q\right)\\ &\quad+\alpha\sum_{\boldsymbol{T}\in\mathcal{T}_h}\frac{h_{\boldsymbol{T}}^2}{\hat{\mu}}\left(\sigma\boldsymbol{u}+\nabla p,\sigma\boldsymbol{v}+\nabla q\right)_{\boldsymbol{T}}+\delta\,\hat{\mu}\left(\nabla\cdot\boldsymbol{u},\nabla\cdot\boldsymbol{v}\right),\\ L_h\left(\boldsymbol{v},q\right) &:= \left(g,q\right)+\delta\,\hat{\mu}\left(g,\nabla\cdot\boldsymbol{v}\right), \end{aligned}$$

and the source term $g$ is the one defined in (7.1.7) with the approximate delta-function (7.2.1), i.e.,

$$g=\sum_{l=1}^{N^{inj}+N^{prod}} g_{w_l^{(\star)}}\delta^{r_\varepsilon}_{w^{(\star)}}.$$

We observe that only the non-symmetric GLS stabilization (with parameter $\alpha$) and the grad-div stabilization (with parameter $\delta$) are present. Since the boundary conditions (7.1.1a) with respect to $\widehat{\boldsymbol{\Omega}}$ are of natural type, the terms in (5.2.10) related to the weak imposition via the penalty-free non-symmetric Nitsche method (including the corner stabilization) are not needed.

### The Discrete Heat Equation

In order to approximately solve the time-dependent advection-diffusion problem for the temperature field (7.1.9), we utilize the approximate Delta function (7.2.1) to represent $\widetilde{\gamma}_k$ via

$$\widetilde{\gamma}_k := \gamma_k\frac{Q_{w^{inj}_{3D,k}}}{H}\delta^{r_\varepsilon}_{w^{inj}_k}.$$

In the following we consider constant values of $\gamma$ and replace the penalty term in (7.1.9a) by

$$cT-f$$

with

$$c:=\sum_{k=1}^{N^{inj}}\gamma\frac{Q_{w^{inj}_{3D,k}}}{H}\delta^{r_\varepsilon}_{w^{inj}_k}\quad\text{and}\quad f:=\sum_{k=1}^{N^{inj}}\gamma\frac{Q_{w^{inj}_{3D,k}}}{H}\delta^{r_\varepsilon}_{w^{inj}_k}T_{w^{inj}_k}. \tag{7.2.4}$$

Then we first discretize the resulting problem in time via an implicit (backward) Euler scheme, see, e.g., [JS08]. Denoting by $T^n$ the approximated temperature field at time $t^n$ and with $\Delta t_n := t^n - t^{n-1}$ the time step, the time-discretized heat equation (at time step $t^n$) reads

$$aT^n+\Delta t_n\left(-d\Delta T^n+\boldsymbol{u}\cdot\nabla T^n+cT^n\right)=aT^{n-1}+\Delta t_n f^n\quad\text{in }\widehat{\boldsymbol{\Omega}}, \tag{7.2.5}$$

with $a, d, c$, and $f$ defined as in (7.1.9) and (7.2.4). In order to further discretize (7.2.5) in space, we rewrite (7.2.5) in weak form and consider it the finite element space $\mathbb{T}_h$ of continuous and piecewise linear functions, resulting in:

$$\mathbb{T}_h := \left\{S_h\in C^0(\,\overline{\widehat{\boldsymbol{\Omega}}}\,)\;:\;S_h|_{\boldsymbol{T}}\in\mathbb{P}_1\left(\boldsymbol{T}\right),\quad\forall\,\boldsymbol{T}\in\mathcal{T}_h\right\}.$$

The finite element formulation evaluated at each time step then reads: Find $T_h^n \in \mathbb{T}_h$ with $T_h|_{\partial\widehat{\mathbf{\Omega}}} = T_0$, such that

$$\begin{aligned} a\,(T_h^n, S_h) - \Delta t_n\left((d\nabla T_h^n, \nabla S_h) + (\boldsymbol{u}\cdot\nabla T_h^n + cT_h^n)\,, S_h\right) \\ = a\,\left(T_h^{n-1}, S_h\right) + \Delta t_n\,(f^n, S_h)\,, \quad \forall S_h \in \mathbb{T}_h \text{ with } S_h|_{\partial\widehat{\mathbf{\Omega}}} = 0. \end{aligned} \tag{7.2.6}$$

In the considered computational setups, the coefficient $d$ will be of order $10^{-7}$ and $\boldsymbol{u}$ will be of order $10^{-6}$, such that a stabilization for advection-dominance is not needed.

**Remark 7.2.2** (On the Penalty Parameter $\gamma$)**.**
*The penalty parameter $\gamma$ contained in $c$ and $f$ (see* (7.2.4)*) allows to accelerate the matching of the injection temperature within the circular region of radius $r_\varepsilon$ (using* (7.2.1)*) with respect to the time discretization, i.e., for example increasing $\gamma$ could enforce the injection temperature after the first time step in a region of radius $r_\varepsilon > r_{w^{(\star)}}$.*

## 7.2.2 Preliminary Validation of the Numerical Method

The scope of this subsection is to validate the performance of the solver in a setting which is relevant for the problem of interest. Further, the dependency of the computational results on the mesh-width, the domain size (impact of boundary conditions), and the time step is investigated.

### Fluid Flow and the Immersed Boundary Method

We start with a "one-well" problem defined in the annular domain

$$\mathbf{\Omega} := \left\{(x,y) \in \mathbb{R}^2 \,:\, r_0^2 \leqslant x^2 + y^2 \leqslant r_1^2\right\} \tag{7.2.7}$$

with $r_0 := 0.2\,\mathrm{m}$ and $r_1 := 1000\,\mathrm{m}$, perforated by an injection well

$$w^{inj} := \left\{(x,y) \in \mathbb{R}^2 \,:\, x^2 + y^2 \leqslant r_0^2\right\}$$

with boundary $\Gamma_{w^{inj}} := \left\{(x,y) \in \mathbb{R}^2 \,:\, x^2 + y^2 = r_0^2\right\}$ (a circle of radius $r_0$).

For the immersed boundary approach we introduce the extended domain

$$\mathbf{\Omega} \cup w^{inj} = \left\{(x,y) \in \mathbb{R}^2 \,:\, x^2 + y^2 \leqslant r_1^2\right\}$$

and use the discrete Dirac delta function (7.2.1) with $r_\varepsilon = 50\,\mathrm{m}$. Hence, the numerical solution is expected to behave as the physical one at a distance from the origin larger than 50 m.

In order to assess the capability of the immersed finite-element method in approximating the solution near the wells, we consider porous media flow described by (3.2.8a)-(3.2.8b) in the Darcy limit, i.e., with $\mu_{\text{eff}} = 0$. Non-homogeneous boundary conditions for the velocity are imposed on $\Gamma_{w^{inj}}$, while homogeneous pressure boundary conditions are set on the outer boundary:

$$\begin{aligned} \boldsymbol{u}\cdot\boldsymbol{n}_{w^{inj}} &= U_{w^{inj}}\,, && \text{on } \Gamma_{w^{inj}}, \\ p &= 0\,, && \text{on } \partial\mathbf{\Omega} \setminus \Gamma_{w^{inj}}, \end{aligned}$$

for some constant $U_{w^{inj}} \in \mathbb{R}$. The analytical solution (in $\mathbf{\Omega}$) is then given by

$$\begin{aligned} \boldsymbol{u}^{sol}(x,y) &= \frac{U_{w^{inj}} r_0}{x^2+y^2} \begin{pmatrix} x \\ y \end{pmatrix}, \\ p^{sol}(x,y) &= \sigma U_{w^{inj}} r_0 \log\left( \frac{r_1}{\sqrt{x^2+y^2}} \right), \end{aligned} \tag{7.2.8}$$

for $\boldsymbol{f} = \mathbf{0}$ and $g = 0$.

For the numerical tests we take $\mu = 0.0003\,\mathrm{kg/m\,s}$ and $k = 3 \cdot 10^{-12}\,\mathrm{m}^2$, yielding $\sigma = 10^8\,\mathrm{kg/m^3\,s}$. Further we use $\ell_{\mathbf{\Omega}} = 50\,\mathrm{m}$, $Q_{w_{3D}^{inj}} = 100\,\mathrm{l/s}$, calculate $U_{w^{inj}} = \frac{Q_{w_{3D}^{inj}}}{2\pi r_0 H}\,[\mathrm{m/s}]$ as described in (7.1.2), and utilize the remaining data as given in Table 7.4. In Figure 7.3 we compare the exact solution $\left(\boldsymbol{u}^{sol}, p^{sol}\right)$, which is defined only outside the very small well with radius $r_0$, i.e. in $\mathbf{\Omega}$, with the numerical solution obtained with the non-matching immersed boundary method with $r_\varepsilon = 50\,\mathrm{m}$, i.e., solving the problem (7.2.2) on $\widehat{\mathbf{\Omega}} = \mathbf{\Omega} \cup w^{inj}$. Therefore, we use a triangular mesh with 17474 vertices and $(h_{\min}, h_{\max}) \approx (6.9, 27.9)$. We visualize half of a circular, centered cutout of radius 100 m on the respective domain ($\mathbf{\Omega}$ and $\widehat{\mathbf{\Omega}}$). For the velocity, the artificial well (of radius $r_\varepsilon$) has a higher transparency than the rest. The tests confirm that the resulting velocity and pressure fields approximate well the exact solution for radii larger than or equal to $r_\varepsilon$. In particular, the pressure isolines (for the immersed method) fit well with those of the exact solution.

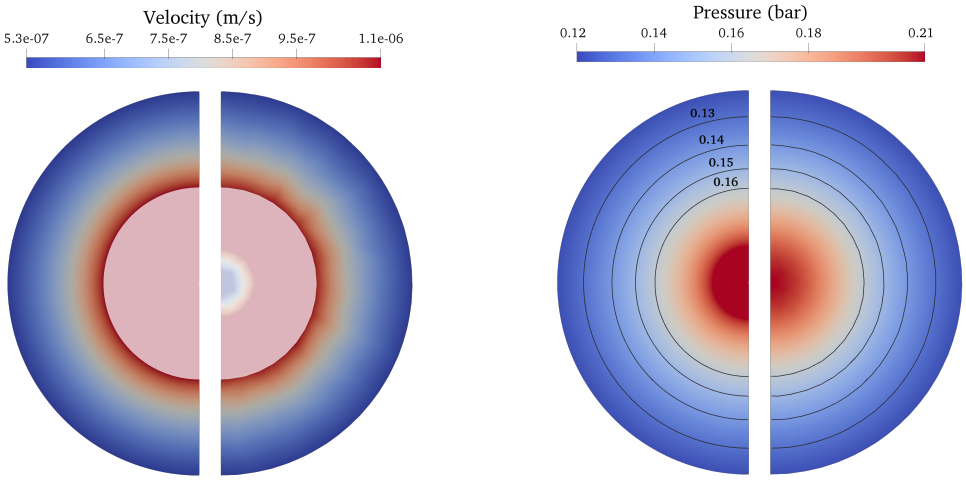



Figure 7.3: Velocity fields (left) and pressure fields (right) – centered cutout of radius 100 m. The exact solution $\left(\boldsymbol{u}^{sol}, p^{sol}\right)$ is shown in the left halves respectively, while the velocity and pressure fields obtained with the immersed method are depicted in the right halves. The area associated with a circle of radius $r_\varepsilon$ is visualized with high transparency for the velocity (left). The black lines (right) show isolines for the pressure.

### Heat Transport and the Immersed Boundary Method

Next, we assess the effect of the penalty-based immersed method, comparing the temperature field obtained with Dirichlet boundary conditions on the boundary of the well with the result from (7.2.6).

To that end we consider the annular domain $\mathbf{\Omega}$ defined in (7.2.7), and the advection field $\boldsymbol{u}^{sol}$ defined in (7.2.8).

The temperature is prescribed at the inner circle, while at the outer circle we impose a homogeneous Neumann boundary condition, and all remaining parameters in (7.1.8) and (7.2.6) are chosen as described in Table 7.4.

We compare two numerical solutions: (i) the one obtained imposing strongly the Dirichlet boundary condition at the well boundary as in (7.1.8) and (ii) the one computed via the penalty-based immersed boundary method (7.2.6).

In the following computations we use a mesh with 20 uniformly distributed nodes on $\Gamma_{w^{inj}}$ and 160 nodes on the outer circular boundary (resulting in 12100 dofs and $(h_{\min}, h_{\max}) \approx (0.064, 55.4)$ for $\mathbf{\Omega}$.

For $\widehat{\mathbf{\Omega}}$ a mesh with 20 nodes, uniformly distributed along the artificial well boundary with radius $r_\varepsilon = 50\,\mathrm{m}$ is considered instead, resulting in a coarser grid (with 4409 dofs, $(h_{\min}, h_{\max}) \approx (13.8, 55.5)$).

The results in Figure 7.4 compare the temperature fields after 12.5 years (left) and the temperature evolution at a specified point (right), and demonstrate a good agreement between the solutions of the approaches (i) and (ii).

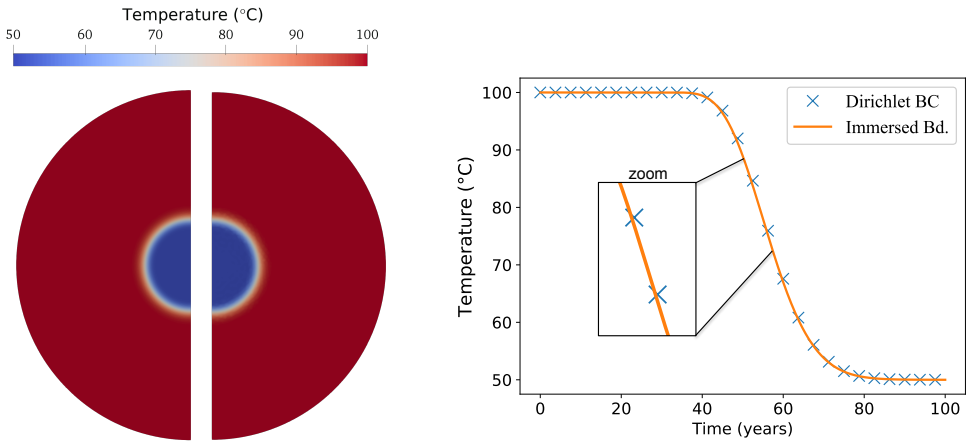



Figure 7.4: Temperature evolution in the annular domain. Left: Comparison of the numerical solutions at time $t = 12.5$ years. The left half is the solution obtained prescribing Dirichlet boundary conditions ($T|_{\Gamma_{w^{inj}}} = 50\,^\circ\mathrm{C}$), while the right half shows the solution using the immersed boundary approach (with $\gamma = 100$). Right: Evolution of the temperature at the point $(x, y) = (560, 40)\,\mathrm{m}$, comparing the solution with Dirichlet boundary conditions (blue crosses) with the result for the immersed boundary method (orange line).

In particular it is worth noticing that there is a slight delay (about 0.5 years) in the cold water front (Figure 7.4 right) computed with the penalty method. This effect is due to the penalty approach, for which the injection temperature is immediately (after the first time step) imposed at a radius $r_\varepsilon$ (larger than the actual physical well), while, in the case of the strongly imposed Dirichlet boundary condition, the cold water front reaches $r_\varepsilon$ after few time iterations.

Let us note that $\gamma$, $r_\varepsilon$, and the time step will be fixed in the numerical study in Section 7.3 such that the delay is present in all considered geometrical configurations and that solely long-term simulations will be performed. Therefore, we assume that it will only marginally affect the result of the forthcoming comparative optimization studies.

### Mesh- and Boundary-Independence Studies

The purpose of this last validation is to perform an extensive preliminary study concerning the influence of the discretization parameters on the result of the sequentially coupled problem. To that end we consider a centered doublet (pair) in a rectangular domain with constant permeability $k = 3 \cdot 10^{-12}\,\mathrm{m}^2$. The production well $w^{prod}$ and the injection well $w^{inj}$ have a fixed distance of 1 km and associated flow rates $Q_{w^{prod}} = Q_{w^{inj}}$ $[\mathrm{m}^3/\mathrm{s}]$ equivalent to $100\,\mathrm{l/s}$. All remaining physical parameters are defined as in Table 7.4 (thermal dispersion has been omitted).

The extended square domain $\widehat{\mathbf{\Omega}}$ (including $w^{prod}$ and $w^{inj}$) of varying size is discretized with an unstructured, non-uniform triangular mesh, where a centered sub-region of interest (of size $2.4 \times 1.4\,\mathrm{km}^2$) is finely resolved and embedded in a coarser mesh.

In this configuration, we monitor the *specific lifetime* of the doublet, which is set as the time until the water temperature at a measurement point of the production well drops below $T_{quit}$, and vary the size of the smallest triangles in the region of interest, the time step, and the size of the domain (i.e., the distance to the outer boundary).

Table 7.1 displays the results of the investigation concerning the time step length and the distance to the boundary. In particular, it can be seen that the boundary has a negligible effect on the *specific lifetime* of the doublet whenever it has a distance of about $\geqslant 3.5\,\mathrm{km}$ from the wells which corresponds here to a domain size $8 \times 8\,\mathrm{km}^2$ or larger. Concerning the time discretization, results become almost independent (variation of less than $\approx 1\%$) from the chosen time step for values below 3 months.

In order to assess the influence of the spatial discretization, we consider (in accordance with the previous results in Table 7.1) the case of a domain of size $10 \times 10\,\mathrm{km}^2$ and a time step length equal to 3 months. In this configuration, we refined the mesh in the region of interest (centered and of size $2.4 \times 1.4\,\mathrm{km}^2$), considering local mesh sizes between 3 m and 50 m, and monitoring in each case the *specific lifetime* of the doublet.

Table 7.1: Specific lifetimes in years varying the domain size and the time step length. Cells marked in green indicate a total variation less than 2% with respect to the subsequent temporal refinement and domain size increase. All quantities are rounded.

| $\lvert\widehat{\Omega}\rvert$ [km$^2$] (DOFS) | $3 \times 3$ (43657) | $5 \times 5$ (44203) | $8 \times 8$ (46042) | $10 \times 10$ (47549) | $15 \times 15$ (52536) |
|---|---|---|---|---|---|
| $\Delta t = 1$a | 55 $\downarrow +2.73\%$ $\xrightarrow{-5.46\%}$ | 52 $\downarrow +2.89\%$ $\xrightarrow{-1.92\%}$ | 51 $\downarrow +2.94\%$ $\xrightarrow{0\%}$ | 51 $\downarrow 2.94\%$ $\xrightarrow{0\%}$ | 51 $\downarrow +1.96\%$ |
| $\Delta t = \frac{1}{2}$a | 56.5 $\downarrow +1.77\%$ $\xrightarrow{-5.31\%}$ | 53.5 $\downarrow +1.40\%$ $\xrightarrow{-1.87\%}$ | 52.5 $\downarrow +1.91\%$ $\xrightarrow{0\%}$ | 52.5 $\downarrow +0.95\%$ $\xrightarrow{-0.95\%}$ | 52 $\downarrow +1.44\%$ |
| $\Delta t = \frac{1}{4}$a | 57.5 $\downarrow +0.57\%$ $\xrightarrow{-5.65\%}$ | 54.25 $\downarrow +0.61\%$ $\xrightarrow{-1.38\%}$ | 53.5 $\downarrow +0.62\%$ $\xrightarrow{-0.94\%}$ | 53 $\downarrow +1.09\%$ $\xrightarrow{-0.47\%}$ | 52.75 $\downarrow +0.95\%$ |
| $\Delta t = \frac{1}{12}$a | 57.83 $\xrightarrow{-5.62\%}$ | 54.58 $\xrightarrow{-1.37\%}$ | 53.83 $\xrightarrow{-0.46\%}$ | 53.58 $\xrightarrow{-0.62\%}$ | 53.25 |

The results in Table 7.2 show that the mesh width in the active subregion has an impact on the *specific lifetime*. Too coarse meshes (mesh size $51,08$ m) yield a significantly overestimated lifetime but differences for the smaller mesh sizes are relatively small. E.g., a mesh size of less than 28 m already yields a variation slightly below 1% towards a next refinement step for the considered setting. We observe that the lifetime is not uniformly varying with the mesh-size and therefore decide to use a rather fine mesh width of $\approx 6$ m in the following simulations. Note that the mesh size should be chosen in dependence of the artificial well diameter (which is $2r_\varepsilon$).

Table 7.2: Specific lifetimes in years (time step 3 months) for different mesh sizes in the active subregion $2.4 \times 1.4\,\text{km}^2$ (characterized by $h_{\min}$) embedded in the $10 \times 10\,\text{km}^2$ domain with a green color visualizing less than 1% variation to the subsequent refinement.

| Mesh Size ($h_{\min}$, $h_{\max}$) [m] | Specific Lifetime [a] |
|---|---|
| (51.08, 259.82) | 57.75 $\downarrow -8.66\%$ |
| (28.27, 267.21) | 52.75 $\downarrow +0.47\%$ |
| (12.87, 267.21) | 53 $\downarrow 0\%$ |
| (6.15, 262.21) | 53 $\downarrow +0.94\%$ |
| (2.94, 262.18) | 53.5 |

Finally, we use the information gained by the previous studies to set up and simulate an appropriate doublet configuration. Figures 7.6 and 7.7 show the resulting velocity, pressure, and temperature fields (at the end of the *specific lifetime*) in the active subregion of size $2.4 \times 1.4\,\text{km}^2$ with mesh size $h \approx 6.15$

($h \approx 262.21$ else), time step $\Delta t = 3$ months, and domain size $10 \times 10\,\mathrm{km}^2$. A cutout of the mesh is shown in Figure 7.5.

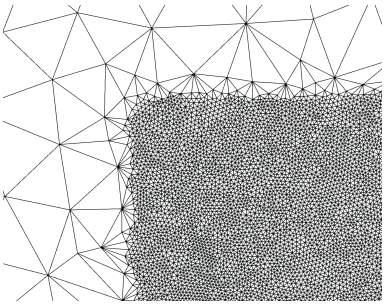


Figure 7.5: Cutout of the mesh for the $10 \times 10\,\mathrm{km}^2$ domain with $h_{\min} \approx 6.15\,\mathrm{m}$ (in the active subregion) and $h_{\max} \approx 262.21\,\mathrm{m}$ (in the inactive region).

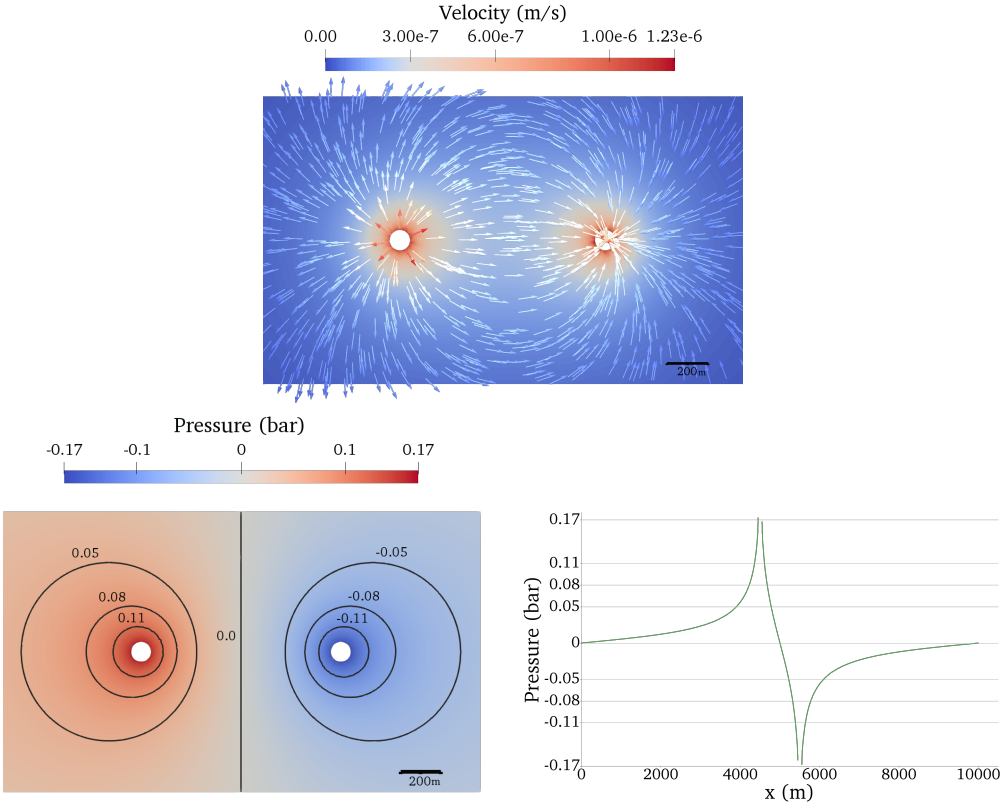



Figure 7.6: Magnitude of the velocity [m/s] with arrows (top) and pressure [bar] with isolines (bottom left) and pressure along the line through the well centers (bottom right) for an injection and production rate equal to 100 [l/s]. (Artificial well regions with radius $r_\varepsilon$ are not shown.)

The arrows in the visualization of the velocity (Figure 7.6, top) indicate that the flow is directed from the injection well (left) towards the production well (right). Due to the homogeneity of the domain and the same flow rates at both

wells, the pressure is symmetric with respect to the isoline 0 bar (except for the sign), see Figure 7.6 (bottom left). The pressure build-up and draw-down (also called *groundwater impression and depression cones*) are shown in Figure 7.6 (bottom right) and reveal a steep gradient in the vicinity of the wells (centered at $(4.5, 5)$ km and $(5.5, 5)$ km) only.

The *cold water plume* after 53 years of doublet operation with the typical tear drop shape towards the production well is shown in Figure 7.7.

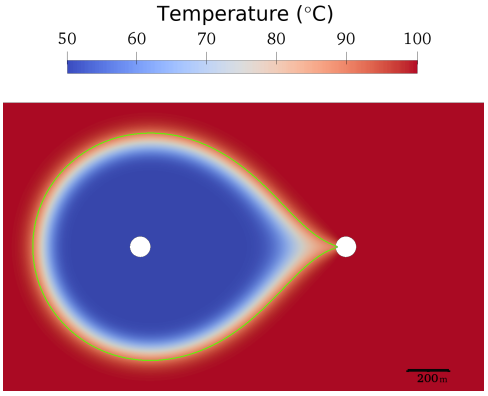



Figure 7.7: Temperature distribution (for the injection temperature 50 °C) in the active subregion after 53 years of operation (specific lifetime) with an isoline (green) highlighting 90 °C. (Artificial well regions with radius $r_\varepsilon$ are not shown.).

### 7.2.3 Optimization of Energy Production

Based on the computational model described in Subsection 7.2.1, an optimization algorithm shall be used for computing the parameters describing the placement of wells (e.g., well locations) that maximizes the net energy production (also called doublet/triplet/etc. capacity).

Let us assign to each well a pump efficiency $\varepsilon_{w^{(\star)}} \in (0, 1]$ and a discharge/flow rate $Q_{3D,w^{(\star)}}$ [$m^3/s$]. The net energy $E_{net}$ [J], with respect to the considered lifetime $t_L$ [s] of the geothermal plant, is given by

$$E_{net} := E_{prod} - E_{pump}, \tag{7.2.9a}$$

where

$$E_{prod} = \int_0^{t_L} \rho_f C_f \left( \sum_{l=1}^{N^{prod}} Q_{3D,w_l^{prod}} T_{w_l^{prod}} - \sum_{k=1}^{N^{inj}} Q_{3D,w_k^{inj}} T_{w_k^{inj}} \right) \mathrm{d}t \tag{7.2.9b}$$

stands for the energy gained through the heat transfer, while

$$E_{pump} = \int_0^{t_L} \left( \sum_{l=1}^{N^{prod}} \frac{Q_{3D,w_l^{prod}}}{\varepsilon_{w_l^{prod}}} \left| \Delta p_{w_l^{prod}} \right| + \sum_{k=1}^{N^{inj}} \frac{Q_{3D,w_k^{inj}}}{\varepsilon_{w_k^{inj}}} \left| \Delta p_{w_k^{inj}} \right| \right) \mathrm{d}t \tag{7.2.9c}$$

denotes the energy that has to be invested in the operation of the pumps and represents energy losses. In Equations (7.2.9b)-(7.2.9c), $\Delta p_{w^{(\star)}}$ denotes the pressure difference (with respect to the ambience/reference pressure) induced by the pump at the respective well, while $T_{w^{(\star)}}$ stands for the temperature at the respective well.

Assuming that the lifetime interval $[0, t_L]$ is decomposed in $N_t$ subintervals, with time step increment $\Delta t_i$, we approximate the net energy as

$$E_{net} \approx \sum_{i=1}^{N_t} \Delta t_i \left[ \rho_f C_f \left( \sum_{l=1}^{N^{prod}} Q^i_{3D,w_l^{prod}} T^i_{w_l^{prod}} - \sum_{k=1}^{N^{inj}} Q^i_{3D,w_k^{inj}} T^i_{w_k^{inj}} \right) - \left( \sum_{l=1}^{N^{prod}} \frac{Q^i_{3D,w_l^{prod}}}{\varepsilon_{w_l^{prod}}} \left| \Delta p^i_{w_l^{prod}} \right| + \sum_{k=1}^{N^{inj}} \frac{Q^i_{3D,w_k^{inj}}}{\varepsilon_{w_k^{inj}}} \left| \Delta p^i_{w_k^{inj}} \right| \right) \right].$$

The problem of finding the positioning of wells that maximizes the energy production ($E_{net}$) can then be formulated as a minimization problem for the objective functional $J := -E_{net}$ as

$$\arg\min \ \mathrm{J}\left(\{(\mathrm{x}_{\mathrm{w}^{(\star)}}, \mathrm{y}_{\mathrm{w}^{(\star)}})\}\right), \tag{7.2.10}$$

where $E_{net}$ (and thus $J$) depends on the well locations $\{(x_{w^{(\star)}}, y_{w^{(\star)}})\}$ through the solutions of the problems for the flow field and the temperature (and on the given data), as well as through the total simulation time, i.e., the number of time integration step $N_t$.

The considered optimization problem (7.2.10) for the multiple well placement aims at minimizing the pressure difference between injection and production wells while maximizing the time until the *thermal breakthrough* (the temperature at a production well drops below the formation temperature) takes place.

**Remark 7.2.3** (On Reasonable Prerequisites)**.**
*The injection and production rates imposed in a geothermal installation should be chosen with respect to the permeability of the region, such that the net energy is positive at least at the initial time - otherwise energy cannot be produced at all, but is lost.*

*It is worthwhile noting that the distance, and thus the geometric parameters, between injection and production wells can only be varied in a reasonable range of values since hydraulic connectivity between injection and production wells should be practically guaranteed. Hence, when it comes to the optimization of the geothermal energy production by multiple wells, not only the optimal placement of the wells in the reservoir is searched for, but also the optimal distance relative to each other plays an important role for the different well arrangements.*

**Remark 7.2.4** (Limitations of the Computational Results)**.**
*As noted in Subsection 7.2.1, the non-matching immersed boundary method provides a physically meaningful solution only for distances larger than $r_\varepsilon$ from the center of the well. Hence, well pressures and water temperatures cannot be exactly evaluated at the well boundary, but are extracted at measurement points along the circle of radius $r_\varepsilon$ (enclosing the well), yielding in practice slightly more*

*pessimistic estimates for the net energy and the economic lifetime. However, for reasonable values of $r_\varepsilon$ it is reasonable to assume that this approximation does not have a significant influence on the results for optimal placement.*

*Related to the model dimension reduction, injection and production pressures are evaluated at the aquifer level only. Therefore, the additional pressure needed to pump water through the well itself (between surface and aquifer level) is neglected, which however is expected to be small, compared to the pressure build-up (resp. drawdown) which arises at the injection (resp. production) wells.*

## 7.3 Numerical Study

This study focuses on two types of multi-well configurations: Lattices (doublet arrays) and hexagonal arrangements. These choices are inspired by geothermal multi-well configurations considered in different geothermal projects, see, e.g., [LZH15], [Wil$^+$17b], and [Vör$^+$07].

The two examined classes of configurations are associated with the notation *Setup L* (for Lattice) and *Setup H* (for Hexagon). In both cases, we consider different scenarios characterized by:

- production and injection wells with respective flow rates,
- boundary and initial conditions on the temperature field,
- heterogeneous geological conditions (varying permeability).

The simulation results aim at illustrating how different operational schemes, formation temperatures, and permeability structures can influence the optimal net energy in multi-well configurations.

Let us remark here, that the rest of the numerical study concerns solely the active $6 \times 6\,\mathrm{km}^2$ subregion.

Further, let us introduce here notation that will be used in the numerical study. We will abbreviate the *earliest thermal breakthrough* by ETB in the plots of the well-wise temperature evolution (considered only for the production wells). In case of a non-constant aquifer temperature, the temperature immediately starts to vary in accordance with the flow field, such that here the criterion for the ETB was implemented to detect temperature drops of at least $0.1\,^\circ\mathrm{C}$. The *specific lifetime* will be referred to as the time when a temperature drop of 10% is measured at any production well. It relates to the *economic lifetime* of the geothermal facility, which is commonly defined as a certain limit of production temperature below which it is no longer economic to proceed. Injection wells will be associated with $\oplus$ and production wells will be visualized by $\ominus$. If not otherwise stated, the final configuration obtained from the optimization (i.e., for step 40) is an optimal solution and therefore presented in detail. For the sake of clarity, the permeability heterogeneities will be indicated in the temperature visualizations by thin pink lines.

The simulations show the result solving the finite element problems (7.2.2), (7.2.6), and the optimization problem (7.2.10).

### 7.3.1 The Computational Domain

The size of the conceptual 2D model of the confined and saturated aquifer is defined according to the results of the preliminary numerical study conducted in Subsection 7.2.2. We consider an *active* region (with increased mesh resolution) of $6 \times 6$ km$^2$, centrally embedded in $\widehat{\mathbf{\Omega}} := [0, 14] \times [0, 14]\,\mathrm{km}^2$, so that a distance of 4 km is assured between the active subregion and the boundary.

As spatial discretization, an unstructured, non-uniform triangular mesh is used. Based on the mesh-independence study performed in Subsection 7.2.2, we choose a characteristic mesh size of approximately 6 m in the active area. Outside of the active subregion, the mesh is gradually coarsened, reaching a maximum element size of approximately 282 m near the (outer) boundary, similar to the mesh shown in Figure 7.5 (see also Table 7.3).

Let us highlight the fact that the mesh is generated only once at the beginning of the optimization, and does not have to be updated at each optimization step (i.e., when repositioning of the wells). This aspect constitutes an important advantage of the proposed numerical framework. Otherwise - especially when computing numerous scenarios of multi-well configurations for a well placement optimization with the finite element method - computationally expensive and time consuming re-meshing procedures would be required. The time discretization is chosen in accordance with the preliminary study (see Table 7.1). Mesh parameters are summarized in Table 7.3.

Table 7.3: Mesh parameters used in the simulations.

| | |
|---|---|
| Domain size | $14 \times 14\,\mathrm{km}^2$ |
| Active subregion (enhanced resolution) | $6 \times 6\,\mathrm{km}^2$ |
| Smallest element diameter ($h_{\min}$) | 5.9 m |
| Largest element diameter ($h_{\max}$) | 281.8 m |
| # Nodes of the spatial mesh | $\approx 5.5 \cdot 10^5$ |
| Time step length | 3 months |

### 7.3.2 Model Parameters

The parameters chosen for the numerical simulations are summarized in Table 7.4. Moreover, in all the considered setups we assume that the production rates $Q$, the injection and production pressures ($\Delta p^{\mathrm{inj}}$, $\Delta p^{\mathrm{prod}}$) as well as the injection temperature $T^{\mathrm{inj}}$ do not depend on time.

Table 7.4: Petrophysical properties of reservoir rock, fluid mechanical properties, reservoir dimensional and operational parameters (top table), partly taken from [Wil+17b] and [Cro+16], and numerical parameters (bottom table) used in the simulations.

| Notation | Unit | Parameter Name | Lattice | Hexagon |
|---|---|---|---|---|
| $\mu_{\text{eff}} = \mu$ | $^{\text{kg}}/_{\text{m s}}$ | viscosities | 0.0003 | |
| $k$ | $\text{m}^2$ | permeability | $\in 3 \cdot \{10^{-11}, 10^{-12}, 10^{-15}\}$ | |
| $H$ | m | aquifer thickness | 300 | |
| $Q_{inj}, Q_{prod}$ | $^{\text{l}}/_{\text{s}}$ | injection, production flow rates | 100 | $\in \{50, 100\}$ |
| $r_w$ | m | well radius | 0.2 | |
| $t_L$ | a | maximum lifetime | 80 | |
| $\phi$ | - | porosity | 0.28 | |
| $\rho_f$ | $^{\text{kg}}/_{\text{m}^3}$ | fluid/brine density | 1050 | |
| $C_f$ | $^{\text{J}}/_{\text{kg K}}$ | fluid/brine heat capacity | 4200 | |
| $\rho_s$ | $^{\text{kg}}/_{\text{m}^3}$ | rock density | 2650 | |
| $C_s$ | $^{\text{J}}/_{\text{kg K}}$ | rock heat capacity | 730 | |
| $\lambda_s$ | $^{\text{J}}/_{\text{K m s}}$ | thermal conductivity (rock) | 2.7 | |
| $\lambda_f$ | $^{\text{J}}/_{\text{K m s}}$ | thermal conductivity (fluid/brine) | 0.7 | |
| $\alpha_L$ | m | longitudinal dispersion coefficient | 5 | |
| $\alpha_T$ | m | transversal dispersion coefficient | 0 | |
| $T_{inj}$ | K | injection temperature | 323.15 ($= 50\,°C$) | |
| $T_0$ | K | aquifer/formation temperature | 373.15 ($=100\,°C$)/linear | |
| $T_{quit}$ | K | minimum production temperature | 363.15 ($= 90\,°C$)/$-10\%$ | |
| $\varepsilon$ | - | pump efficiency | 0.6 | |
| $\alpha$ | - | stabilization parameter (GLS) | 1 | |
| $\delta$ | - | stabilization parameter (grad-div) | 0.1 | |
| $\ell_{\mathbf{\Omega}}$ | m | characteristic length | 300 | |
| $\gamma$ | - | penalty for temperature BC | 100 | |
| $r_\varepsilon$ | m | artificial well radius (IBM) | 50 | |
| $N_O$ | - | optimization steps | 40 | |

### Initial and Boundary Conditions

In the computations, all pressure values are computed with respect to the pressure at the depth of the aquifer, where the reference (datum) is set.

For the heat equation, which describes the evolution of the temperature field, we consider aquifers that extend horizontally, assuming a constant depth of approximately 3 km below surface. For a surface temperature of 10 °C, we set the initial and boundary conditions

$$\begin{aligned} T(0, \boldsymbol{x}) &= 100\,°\text{C}, \quad \text{for } \boldsymbol{x} \in \widehat{\mathbf{\Omega}}, \\ T(t, \boldsymbol{x}) &= 100\,°\text{C}, \quad \text{for } t > 0,\ \boldsymbol{x} \in \partial\widehat{\mathbf{\Omega}}\,. \end{aligned} \tag{7.3.1}$$

This is in line with the natural temperature gradient of approximately 30 $^{\text{°C}}/_{\text{km}}$ as described in Subsection 7.1.1.

For the case of a dipping/declined aquifer, we consider a temperature distribution varying linearly along the boundary and constant in time, such that

$$T(t,x,y) = \begin{cases} 100\,^\circ\mathrm{C}, & \text{for } x = 0\,\mathrm{km}, \\ 130\,^\circ\mathrm{C}, & \text{for } x = 14\,\mathrm{km}\,. \end{cases} \tag{7.3.2}$$

Note that this corresponds to an increase in depth of about 1 km in the horizontal direction.

#### Optimization Parameters

The library NLopt, [Nlo], supports a large variety of derivative-free optimization algorithms and, a priori, it can be hard to select the best choice for the considered problem, [RS13]. In this numerical investigation, a detailed study has been performed comparing two local (COBYLA, NEWUOA) with one global (DIRECTL, DIviding RECTangles algorithm Locally-biased, [GK01]) optimization algorithm. The numerical study revealed that the global optimization outperforms the local variants, which might get trapped in local minima of (7.2.10). Therefore, numerical results obtained with the local approaches will be omitted, presenting in the rest of this section only the outcome of the DIRECTL method.

Therefore, the numerical results presented in this section show the outcome of the global optimization algorithm DIRECTL. This is a deterministic-search algorithm based on decomposing the search domain into hypercubes (hyperrectangles) and successive refinement.

For each considered case, we allow a maximum of 40 optimization steps. The so-called *control variables* can be altered during the optimization in order to determine optimized well positions. Here we consider

- the (rigid) rotation angle (counter clockwise) in $[0, 180]\,^\circ\mathrm{C}$,
- the distance between well centers, (lattice size in $[500, 1414]\,\mathrm{m}$, hexagon radius in $[500, 2800]\,\mathrm{m}$) and
- the $x$-translation (in $[-1221, 1221]\,\mathrm{m}$) of the whole geometric structure.

For each setup, the configuration corresponding to the optimization step yielding the maximum net energy will be chosen for a detailed presentation.

The stopping criterion for the temperature simulation (in each optimization step) is twofold: On the one hand, the maximum production time is restricted to 80 years. On the other hand, the simulation stops whenever the *specific lifetime* is reached, i.e., the minimum temperature measured at a production well (see Rem. 7.2.4) reveals 10% temperature reduction.

### 7.3.3 Lattice-Type Configurations

The lattice (checkerboard) configuration represents geothermal doublet arrays and consists of 8 injection wells and 8 production wells, which are arranged in form of a 4 rows × 4 columns lattice, where the same flow rate is imposed at all wells (see Table 7.4). In each row, two doublets are placed, where injection and production sites appear staggered: the rows one and three start with a production well, whereas the rows two and four start with an injection well, see Figure 7.8.

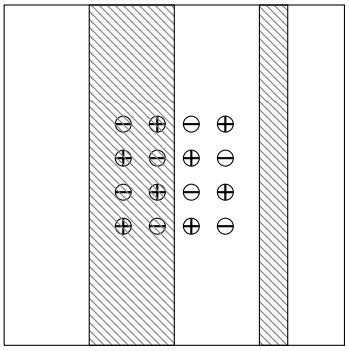


Figure 7.8: Lattice: Sketch of the $6 \times 6\,\mathrm{km}^2$ (centered) subdomain in the case of rotation angle $0°$, lattice size 600 m, and translation 0 m. Injection wells are associated with $\oplus$, $\ominus$ indicates production wells, patterned and plain regions indicate two different permeabilities (laterally varying).

In order to assess the impact of geometrical parameters of the lattice arrangement on the net energy, different scenarios with varying control variables in homogeneous and heterogeneous temperature and permeability fields are examined.

Firstly, we consider a horizontally varying permeability structure with two different permeabilities. This is intended to resemble two different carbonate facies as, e.g., typically encountered in the Upper Jurassic carbonates in the Munich region. We set

$$k = \begin{cases} 3 \cdot 10^{-15}\,\mathrm{m}^2, & \text{for } x \in (5500, 7000) \cup (8500, 9000), \\ 3 \cdot 10^{-12}\,\mathrm{m}^2, & \text{elsewhere,} \end{cases} \tag{7.3.3}$$

as indicated in Figure 7.8.

Secondly, we distinguish between (i) constant initial and boundary temperature conditions and a heterogeneous permeability structure, and (ii) a dipping aquifer with linearly varying initial and boundary conditions for the temperature.

For each case, we are searching for the optimal position of the wells with respect to the produced net energy (7.2.9). In a first scenario, we fix the position of the lattice center at the center of the domain and consider a variation of the *lattice size* (i.e., the distance between closest wells) and rigid rotation. Thereby, the lattice size is allowed to vary between 500 m and 1414 m, while the rotation cannot exceed $180°$ (due to symmetry). In a second scenario, we fix the lattice size equal to 600 m, allowing a rigid rotation in $[0°, 180°]$ and a rigid translational displacement in $x$-direction with a maximal value of 1221 m with respect to the initial position.

The considered setups for the lattice-type configuration are summarized in Table 7.5.

Table 7.5: Considered lattice setups with fixed flow rate of 100 l/s per well (i.e., in total 1600 l/s), varying in the permeability structure, the formation temperature, and the control variables.

| Setup | Permeability | Temperature | Control variables |
|---|---|---|---|
| LC | Heterog. (7.3.3) | Constant (7.3.1) | Lattice size, rotation angle |
| L1 | Heterog. (7.3.3) | Linear (7.3.2) | Lattice size, rotation angle |
| L2 | Homog. ($k = 3 \cdot 10^{-12}\,\mathrm{m}^2$) | Linear (7.3.2) | $x$-translation, rotation angle |
| L3 | Heterog. (7.3.3) | Linear (7.3.2) | $x$-translation, rotation angle |

### Optimal Placement and Net Energy

In Figure 7.9 we plot for each of the lattice scenarios the results for the net energy over the optimization step.

The curves show the strong impact of the inter-well distance (lattice size) on the extracted energy. In fact, including the variation of the lattice size as optimization variable (enabling larger sizes as in setups (LC) and (L1)) allows to increase the obtained energy by approximately up to 300% with respect to the cases (L2) and (L3), where the lattice size is kept fixed to a smaller value. The setups (LC) and (L1) result in strong oscillations of the collected net energy, which are related to the variation of the influential lattice size. Further, we observe that the net energy for (LC) is generally smaller than the net energy for (L1), which can be explained by the higher average temperature of the considered aquifer domain in (L1). Note that the net energy for the setups with fixed lattice size ((L2) and (L3)) is more robust with respect to a variation of their controls (rotation angle and $x$-translation). In particular, the homogeneous setup (L2) has an almost constant energy level which is larger than the slightly varying energy associated with its heterogeneous (with partly lower permeability) version (L3).

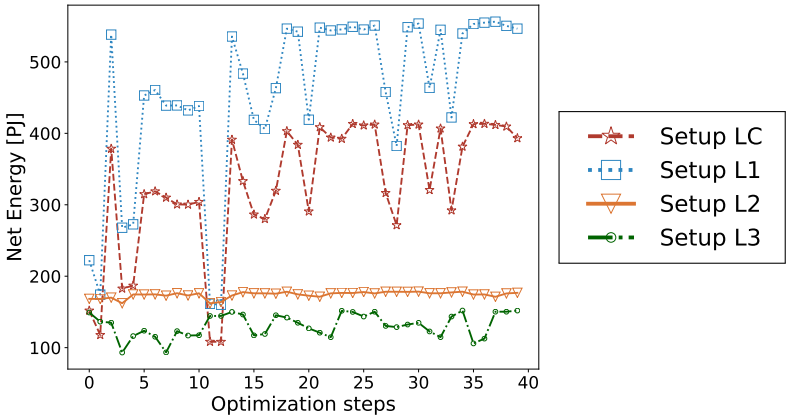



Figure 7.9: Lattice Setups: Net energy collected over the simulation time in each of the 40 optimization steps with the global optimization algorithm DIRECTL for the four different scenarios described in Table 7.5.

### Simulation Results in the Optimal Configurations

In the following, the computational results for the velocity fields, the pressure fields, the temperature fields, and the well-wise temperature evolutions in the optimized geometrical configuration - either at the end of the *specific lifetime*, or after 80 years of operation - are discussed.

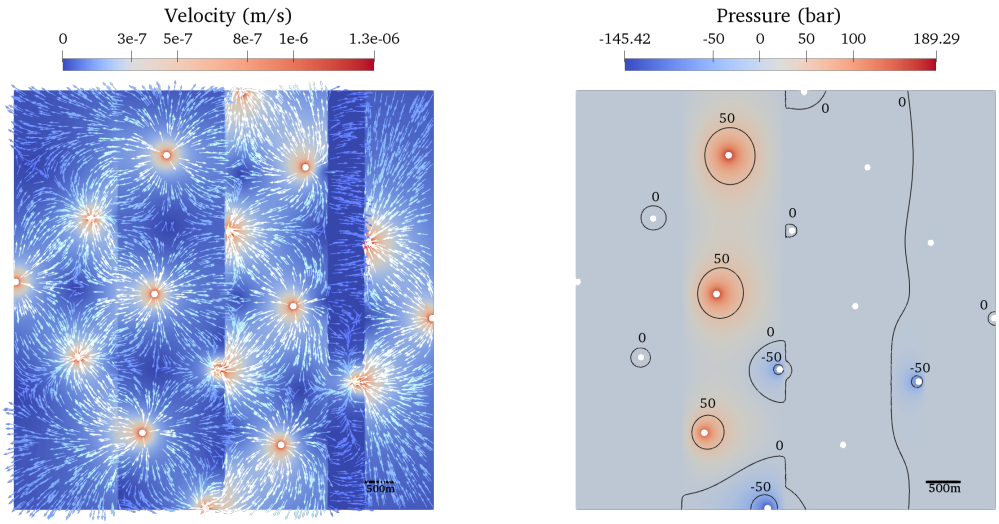



Figure 7.10: Setup (LC): Velocity with unscaled arrows (left) and pressure with contours (right) in the optimal configuration. Control variables: rotation angle $\approx 2.27\,\text{rad}$, lattice size $\approx 1408.4\,\text{m}$ (translation is fixed to $0\,\text{m}$).

Figures 7.10 and 7.11 show the velocity, pressure, and temperature corresponding to the optimal positioning (in step 37, see Figure 7.9) for the **setup (LC)**. The heterogeneous permeability structure yields a reduced velocity in the low permeability regions, which decelerates the evolution of the cold water front in one direction, while accelerating it in another direction. The maximal energy is reached for an almost maximal lattice size (in the allowed range $[500, 1414]\,\text{m}$) and a rotation angle that places the majority of the wells in the high permeability regions, where lower pressures are required to introduce a certain flow rate. Only three injection and three production wells are located in the low permeability regions.

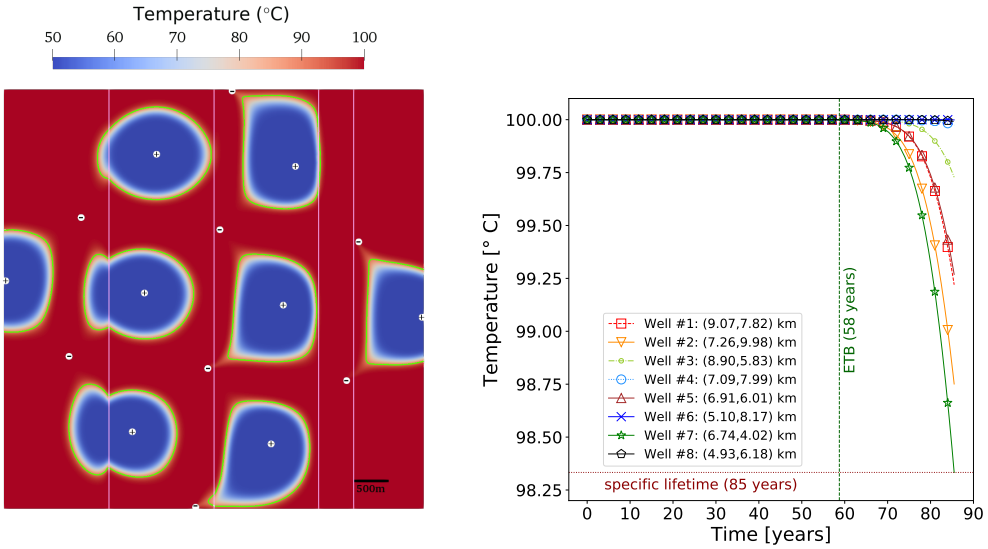



Figure 7.11: Setup (LC): Temperature distribution in the optimal configuration after 80 years of operation (left) and well-wise temperature evolution (right). Control variables: rotation angle $\approx 2.27$ rad, lattice size $\approx 1408.4$ m (translation is fixed to 0 m). The green line (left) is a 90 °C isoline.

Maximizing the net energy implies minimizing the pressure difference between injectors and producers while maximizing the *specific lifetime*. In particular the latter is very sensitive to the inter-well distance. If the *thermal breakthrough* is not established within 80 years of operation, then the net energy (7.2.9) is controlled by the pressure difference between injectors and producers. In heterogeneous permeability structures this relates to the positioning of the wells. As soon as the *earliest thermal breakthrough* occurs, as illustrated in Figure 7.11 (right, vertical green line), both, the pressure difference between injection and production wells and the production temperature determine the net energy (7.2.9) and thus the optimized positioning.

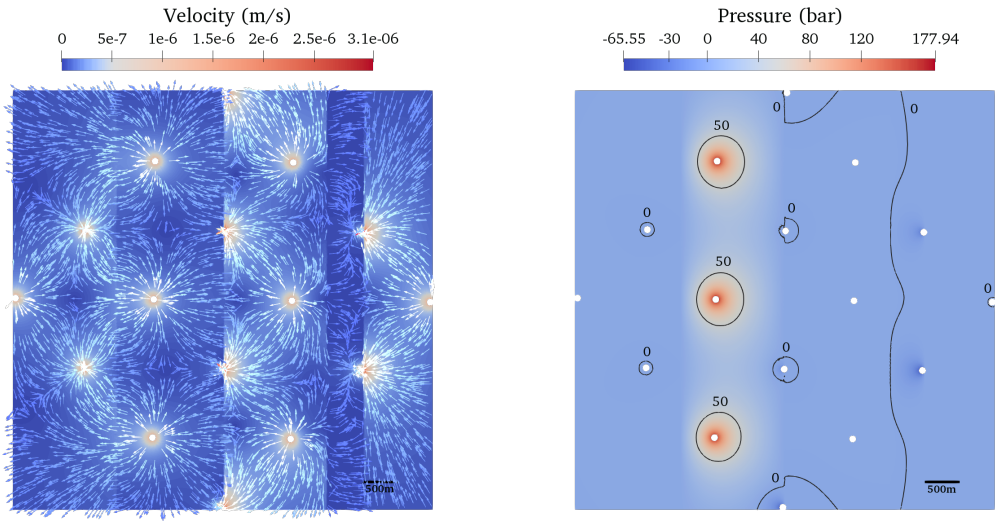



Figure 7.12: Setup (L1): Velocity with unscaled arrows (left) and pressure with contours (right) in the optimal configuration. Control variables: rotation angle $\approx 2.35\,\mathrm{rad}$, lattice size $= 1397\,\mathrm{m}$ (translation is fixed to $0\,\mathrm{m}$).

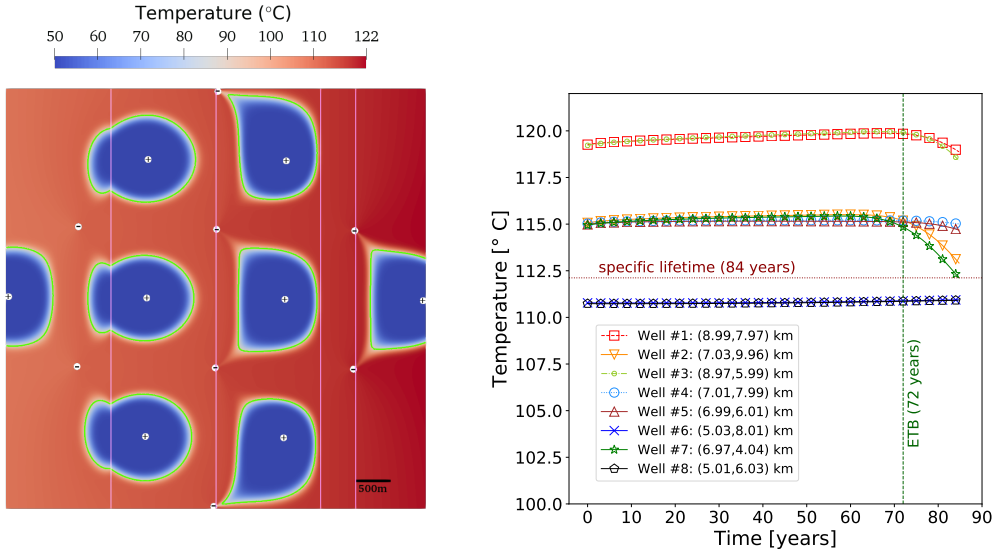



Figure 7.13: Setup (L1): Temperature distribution in the optimal configuration after 80 years of operation (left) and well-wise temperature evolution (right). Control variables: rotation angle $\approx 2.35\,\mathrm{rad}$, lattice size $= 1397\,\mathrm{m}$ (translation is fixed to $0\,\mathrm{m}$). The green line (left) is a $90\,^{\circ}\mathrm{C}$ isoline.

Introducing the linearly varying formation temperature (7.3.2) as in **setup (L1)**, similar simulation results as for (LC) are obtained. The maximum net energy in the case of (L1) is gained at optimization step 38 and the corresponding

configuration is shown in Figures 7.12 and 7.13. We observe that the majority of the production wells is located in a zone of intermediate temperature. For the wells placed in the coolest parts, the temperature slightly increases over time due to the fluid coming from hotter sections of the reservoir, see Figure 7.13 (right).

For the considered geological conditions and the control variables, this result suggests that in the absence of a significant temperature anomaly, a moderately linearly varying reservoir temperature in the order of magnitude as described in (7.3.2) has almost no impact on the optimization.

The cold water front reaches the production wells located north (#2) and south (#7) in the scenarios (LC) and (L1) first, and the southernmost well restricts the *specific lifetime*, see Figure 7.11 and 7.13. Within a lattice structure, an internal injection well distributes the flow rate to 4 surrounding production wells in different directions. Consequently, the velocity of the fluid along the connecting line is lower than in the case of fewer surrounding production wells. Production wells located at the border of the lattice are surrounded by two or three injectors also positioned at the border, such that the progress of the cooling front is faster here, which might be even promoted by unfavorable permeability structures.

For a fixed lattice size of 600 m, the **setup (L2)** allows for the $x$-translation and rotation of the lattice configuration within a homogeneous aquifer with linearly varying formation temperature. In Figures 7.14 and 7.15 we observe a significant translation towards higher temperatures and note the different speeds of temperature drop for the respective production wells. Tied to the fixed lattice size is the *specific lifetime* of 24 years only.

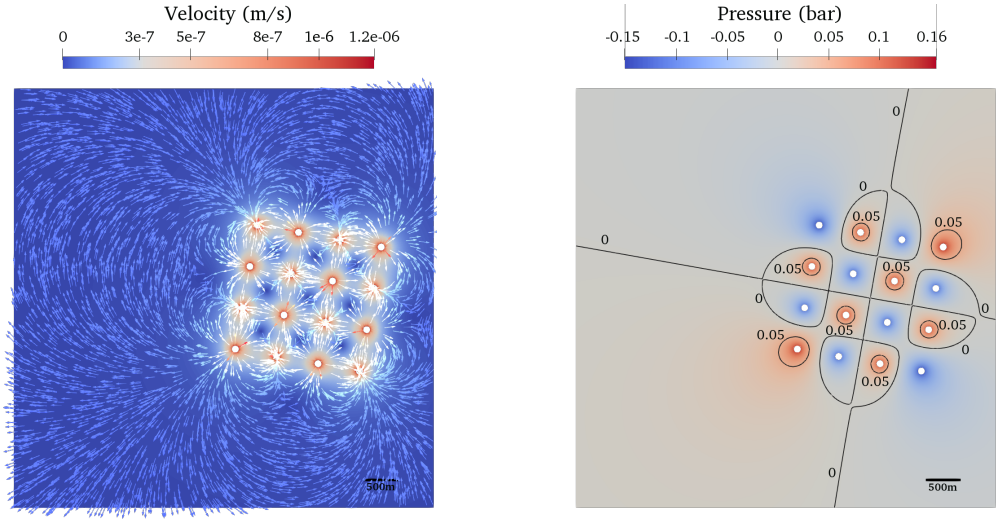



Figure 7.14: Setup (L2): Velocity with unscaled arrows (left) and pressure with contours (right) in the optimal configuration. Control variables: rotation angle $\approx 2.97$ rad, translation $\approx 1206$ m (lattice size is fixed to 600 m).

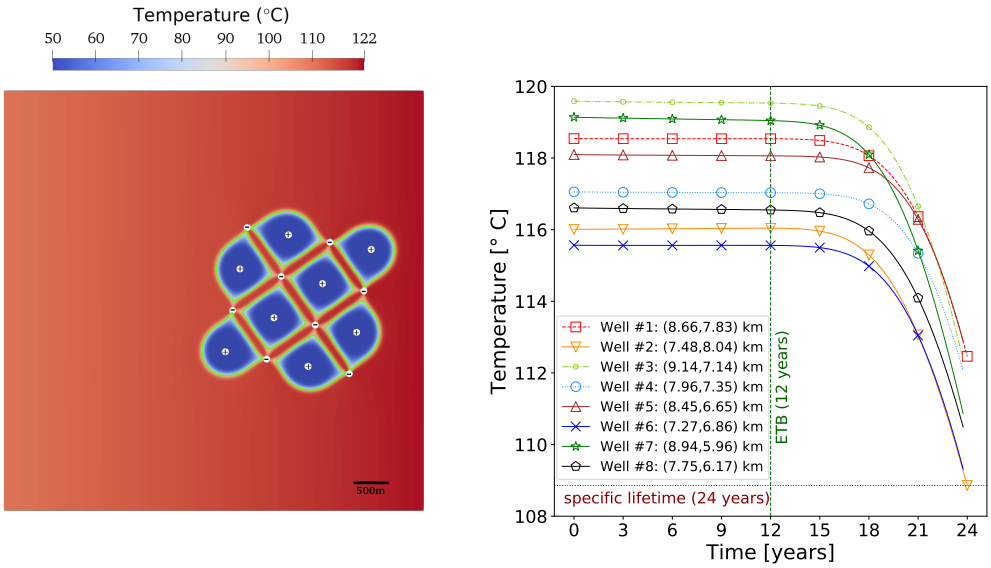



Figure 7.15: Setup (L2): Temperature distribution in the optimal configuration after the *specific lifetime* of 24 years (left) and well-wise temperature evolution (right). Control variables: rotation angle $\approx 2.97$ rad, translation $\approx 1206$ m (lattice size is fixed to 600 m). The green line (left) is a 90 °C isoline.

In contrast to the simulation results concerning the previous setups (LC) and (L1), the pressure range of values is much lower for the setup (L2) due to the homogeneous permeability that characterizes this setup. Since the pressure field remains constant over the entire simulation (independent of temperature, rotation, and translation), only the production temperature influences the optimal placement of the lattice configuration concerning (7.2.10). As long as the *thermal breakthrough* is not established, the maximum net energy of the system is solely controlled by the initial temperature difference of the respective doublets.

The only difference between **setup (L3)** and the previous setup (L2) is the permeability which is now heterogeneous. The simulation results for this case are visualized in Figures 7.16 and 7.17. The *specific lifetime* of 22.75 years is slightly shorter than for (L2).

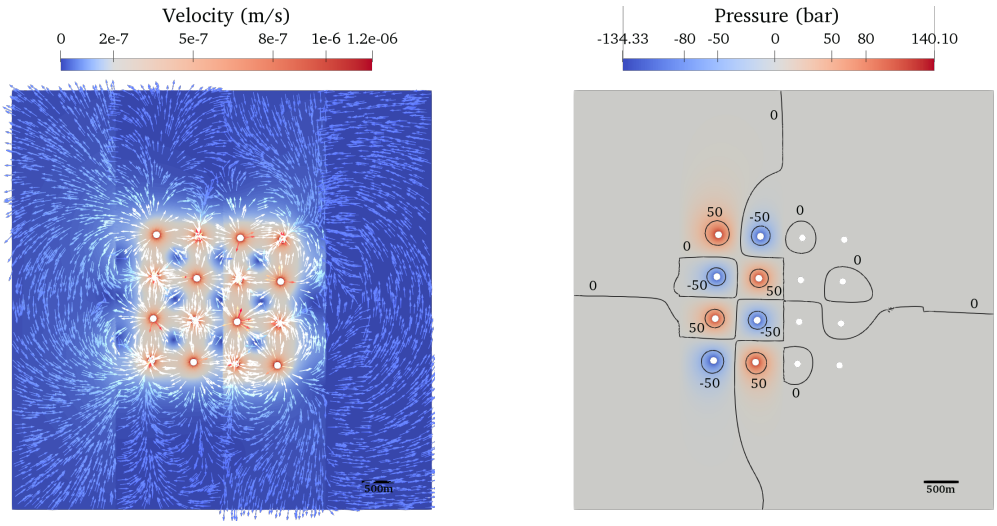



Figure 7.16: Setup (L3): Velocity with unscaled arrows (left) and pressure with contours (right) in the optimal configuration. Control variables: rotation angle $\approx 1.53\,\text{rad}$, translation $\approx -70.35\,\text{m}$ (lattice size is fixed to $600\,\text{m}$).

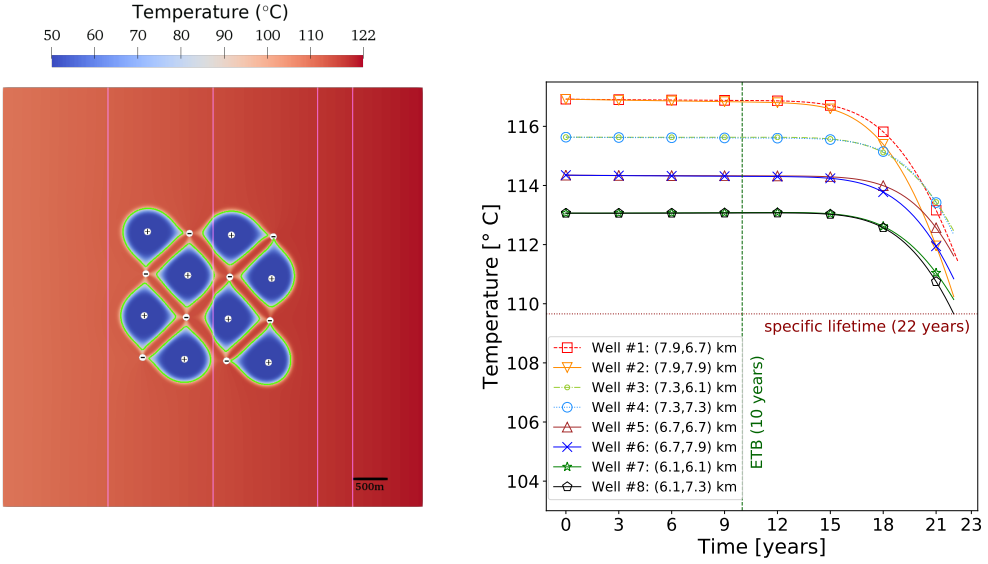



Figure 7.17: Setup (L3): Temperature distribution in the optimal configuration after the *specific lifetime* of 22.75 years (left) and well-wise temperature evolution (right). Control variables: rotation angle $\approx 1.53\,\text{rad}$, translation $\approx -70.35\,\text{m}$ (lattice size is fixed to $600\,\text{m}$). The green line (left) is a $90\,°\text{C}$ isoline.

As can be seen in the velocity and pressure fields illustrated in Figure 7.16, the linear aquifer temperature distribution has no significant effect on the positioning.

The lattice is even slightly moved in the direction of lower temperatures. This suggests that under the given conditions, the permeability structure (with contrasts of several orders of magnitude) mainly controls the energy.

In the optimal configuration, half of the lattice structure is positioned in a low permeability zone and the other half of the lattice in a high permeability zone. In Figure 7.17 (left), we observe that this effects a deceleration of the cooling fronts towards the permeability contrast.

### 7.3.4 Hexagonal Configurations

The hexagonal setup concerns arrangements with 6 wells located at the corners of an equilateral hexagon. The corresponding flow rates might vary by half, see Table 7.4. For this configuration we will also consider different scenarios. Keeping the formation temperature constant in all cases (as in (7.3.1)), we consider heterogeneous permeability structures characterized by *faults/damage zones*, see Figure 7.18.

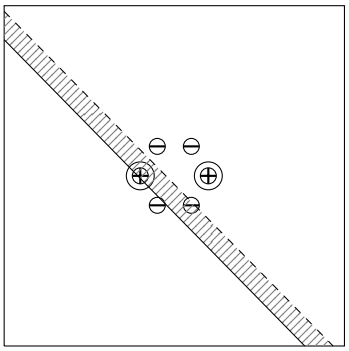


Figure 7.18: Hexagon: Sketch of the $6 \times 6\,\mathrm{km}^2$ (centered) subdomain for the hexagon configuration ($\mathrm{H}_{I2P4}$) in the case of rotation angle $0^\circ$ and hexagon radius 600 m. The dashed line confining the damage zone is variable in its distance (between 100 m and 400 m) to the solid line. Injection wells are associated to $\oplus$, $\ominus$ indicates production wells, and patterned and plain regions indicate two different permeabilities.

We distinguish two different permeability structures: (I) with a hydraulically active damage zone (i.e., *leak fault*)

$$k = \begin{cases} 3 \cdot 10^{-12}\,\mathrm{m}^2, & \text{in the interior of the damage zone,} \\ 3 \cdot 10^{-15}\,\mathrm{m}^2, & \text{elsewhere,} \end{cases} \tag{7.3.4}$$

and (II) with a slightly healed damage zone:

$$k = \begin{cases} 3 \cdot 10^{-12}\,\mathrm{m}^2, & \text{in the interior of the damage zone,} \\ 3 \cdot 10^{-11}\,\mathrm{m}^2, & \text{elsewhere.} \end{cases} \tag{7.3.5}$$

Throughout the simulations, we additionally vary the width of the damage zone $\mathrm{width}^{\mathrm{dz}}$ as

$$\mathrm{width}^{\mathrm{dz}} \in \{100, 200, 300, 400\}\ \mathrm{m}.$$

The damage zones are obtained considering the confining lines passing through the points

$$(a_1, 0), \text{ and } (0, a_1),$$
$$(a_1 + \sqrt{2}\,\text{width}^{\text{dz}}, 0), \text{ and } (0, a_1 + \sqrt{2}\,\text{width}^{\text{dz}}),$$

where $a_1 = 13400\,\text{m}$. Let us mention that in all of these cases, the center of the hexagon (which coincides with the center of the active region) is located outside the damage zone. The smaller the damage zone width gets, the further away is the fault center from the hexagon center.

Furthermore we consider a scenario, in which the center of the hexagon approximately coincides with the center of the damage zone of $\text{width}^{\text{dz}} = 400\,\text{m}$ width, defined by the lines passing through the two points:

$$P_1 = (a_2, 0),\ P_2 = (0, a_2),\ \text{(first line)},$$
$$Q_1 = (a_3, 0),\ Q_2 = (0, a_3),\ \text{(second line)}\,,$$

using $a_2 = 13717.66\,\text{m}$ and $a_3 = 14282.3\,\text{m}$.

In addition, the following exploitation schemes will be used:

- $\text{H}_{I3P3}$: 3 injection and 3 production wells (alternated) along the hexagon, all wells are assigned the same flow rate of $100\,\text{l/s}$,
- $\text{H}_{I4P2}$: 4 injection wells and 2 production wells, $Q_{w^{\text{inj}}} = 50\,\text{l/s}$, $Q_{w^{\text{prod}}} = 100\,\text{l/s}$,
- $\text{H}_{I2P4}$: 2 injection wells and 4 production wells, $Q_{w^{\text{inj}}} = 100\,\text{l/s}$, $Q_{w^{\text{prod}}} = 50\,\text{l/s}$ (shown in Figure 7.18).

In all of these cases, we seek for the optimal configuration varying the radius of the hexagon between 500 m and 2800 m and rigidly rotating the hexagon up to 180° with respect to its center (i.e., translation is not considered for the hexagonal configurations).

In the following simulations, the five essentially different hexagonal scenarios, summarized in Table 7.6, are considered.

Table 7.6: Main hexagon setups with $\oplus$ and $\ominus$ associated with injection and production wells, respectively (counter clockwise, starting from the positive $x$-axis in the start configuration). Control variables: radius and rotation angle in all cases.

| **Setup** | **Permeability** | **width$^{dz}$** | **Wells** | **Total Flow Rate** |
|---|---|---|---|---|
| $\text{H}_{I3P3}$ | Eq. (7.3.4) | 400 m | $\oplus\ \ominus\ \oplus\ \ominus\ \oplus\ \ominus$ | 600 l/s |
| $\text{H}_{I4P2}$ | Eq. (7.3.4) | 400 m | $\ominus\ \oplus\ \oplus\ \ominus\ \oplus\ \oplus$ | 400 l/s |
| $\text{H}_{I2P4}$(i) | Eq. (7.3.4) | 400 m | $\oplus\ \ominus\ \ominus\ \oplus\ \ominus\ \ominus$ | 400 l/s |
| $\text{H}_{I2P4}$(ii) | Eq. (7.3.4) | 400 m (centered) | $\oplus\ \ominus\ \ominus\ \oplus\ \ominus\ \ominus$ | 400 l/s |
| $\text{H}_{I2P4}$(iii) | Eq. (7.3.5) | 200 m (healed) | $\oplus\ \ominus\ \ominus\ \oplus\ \ominus\ \ominus$ | 400 l/s |

The presentation of the computational results includes the net energy during the optimization comparing different damage zone widths for setups ($H_{I3P3}$), ($H_{I4P2}$), and ($H_{I2P4}$(i)). Here, only the results corresponding to the optimal positioning for $\text{width}^{\text{dz}} = 400\,\text{m}$ are presented in detail.

Let us recall that, except for the scenario ($H_{I2P4}$(ii)), where the damage zone centrally passes through the center of the hexagon, the optimization is not able to place two wells entirely inside the damage zone. Further, for decreasing damage zone width, the damage zone moves away from the center of the hexagon.

### Setup ($H_{I3P3}$) – 3 Injection and 3 Production Wells (Alternating)

During the first optimization steps, we observe a strong variation of almost factor 3 in the resulting net energy for the setup ($H_{I3P3}$), see Figure 7.19. Here the three smallest values of the net energy correspond to small hexagon radii of less than 1 km, i.e., also for this hexagonal case the inter-well distance has a much stronger impact on the net energy than the rotation angle. The optimized positioning is very similar for the different damage zone widths, however, the largest ($\text{width}^{\text{dz}} = 400\,\text{m}$) yields the maximal net energy, which decreases with decreasing damage zone width.

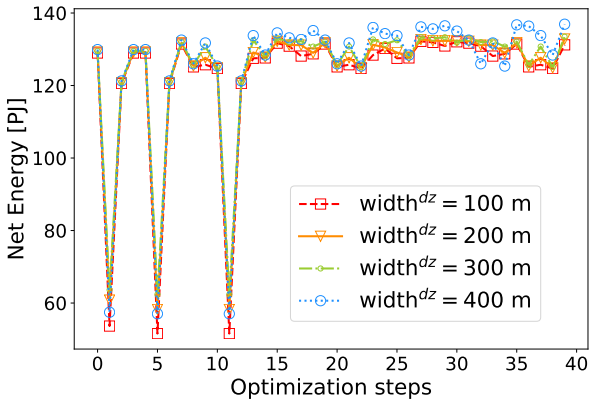



Figure 7.19: Setup ($H_{I3P3}$): Net energy collected over the simulation time in each of the 40 optimization steps with the global optimization algorithm DIRECTL for different damage zone widths.

Figures 7.20 and 7.21 illustrate the simulation results for the optimal configuration with $\text{width}^{\text{dz}} = 400\,\text{m}$. There, one injection well and one production well are positioned as close as possible to the damage zone (at the border), which is of enhanced permeability such that the pressure difference between these two wells is small, see Figure 7.20 (right).

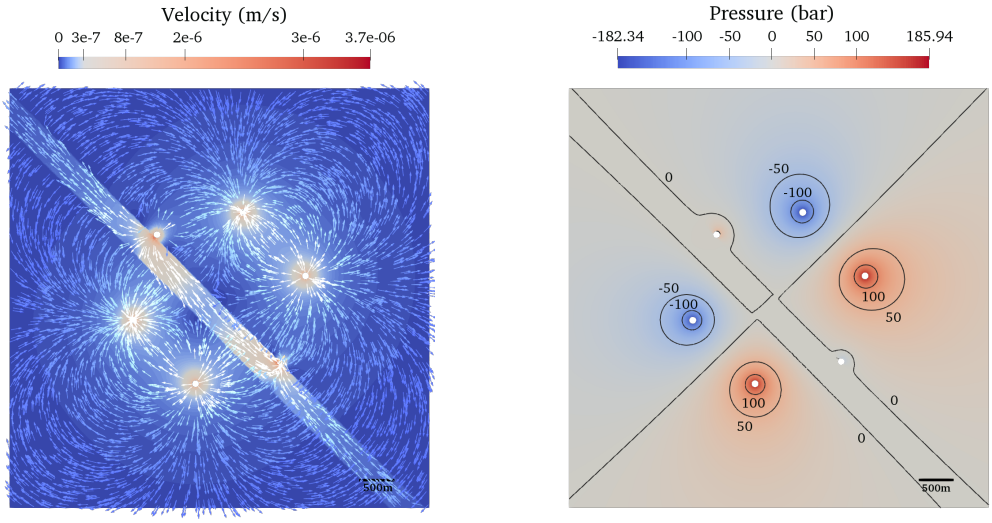



Figure 7.20: Setup ($H_{I3P3}$): Velocity with unscaled arrows (left) and pressure with contours (right) in the optimal configuration for the fault width 400 m. Control variables: rotation angle ≈ 2.35 rad, hexagon radius = 1271 m.

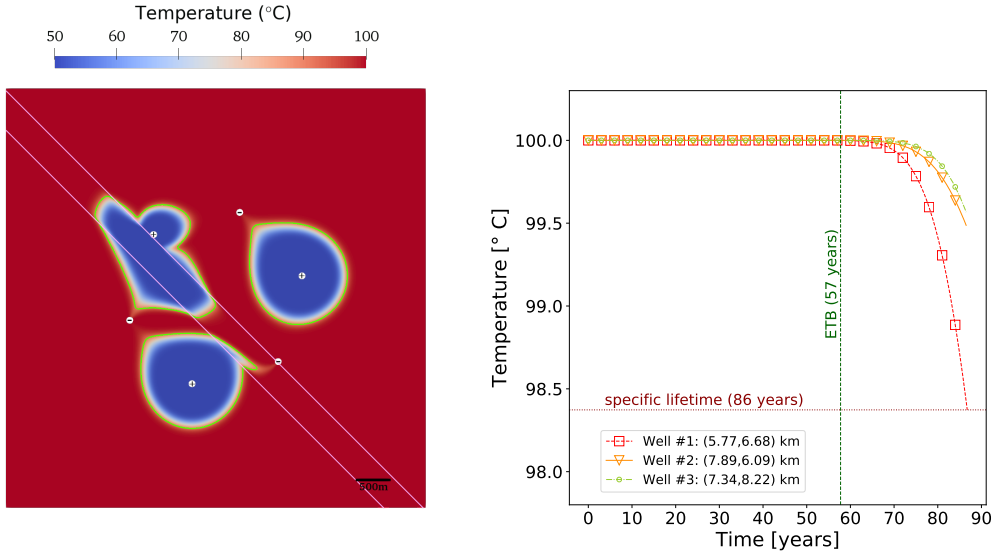



Figure 7.21: Setup ($H_{I3P3}$): Temperature distribution in the optimal configuration for the fault width 400 m after 80 years of operation (left) and well-wise temperature evolution (right). Control variables: rotation angle ≈ 2.35 rad, hexagon radius ≈ 1271 m. The green line (left) is a 90 °C isoline.

In addition, the progress of the cooling front emanating from the injection well at the fault is channeled in the direction of the main axis of the damage zone (where the inter-well distance of the doublet is largest) and decelerated in

the direction of the other neighboring production wells (see Figure 7.21).

The *specific lifetime* is reached shortly after 80 years of operation at the westernmost production well (#1 in Figure 7.21). The velocity, pressure, and temperature fields associated to the four wells away from the damage zone are not significantly influenced by the other two wells (as long as the cold water front stays away from the damage zone). They show a doublet-like interaction such that the typical tear drop shape of the cooling front can be observed on each side of the fault – similar to a single doublet simulation (see Figure 7.7).

### Setup ($H_{I4P2}$) – 4 Injection and 2 Production Wells

Figure 7.22 compares the optimization processes for the different widths of the damage zone and the setup ($H_{I4P2}$). Similar as for the setup ($H_{I3P3}$), the width of 400 m yields the largest net energy. We also recognize the strong influence of the hexagon radius on the resulting net energy.

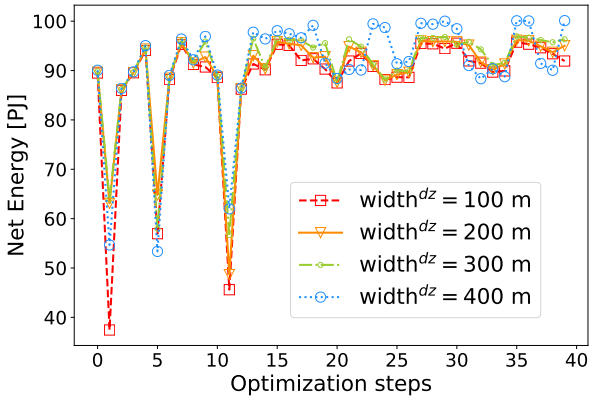



Figure 7.22: Setup ($H_{I4P2}$): Net energy collected over the simulation time in each of the 40 optimization steps, with the global optimization algorithm DIRECTL for different damage zone widths.

The velocity, pressure, and temperature for the optimal configuration in case of 400 m fault width are presented in Figures 7.23 - 7.25. The control variables are very similar to those for the setup ($H_{I3P3}$), i.e., also here two wells are placed as near as possible to the fault zone which has a higher permeability. This time, the two production wells which have higher (doubled) flow rates than the injection wells are chosen. We conclude that, although different competing mechanisms at different stages of the thermo-hydraulic interaction between the wells are involved, placing the wells with highest flow rates as near as possible to the high permeability zone is decisive for the optimization of the net energy.

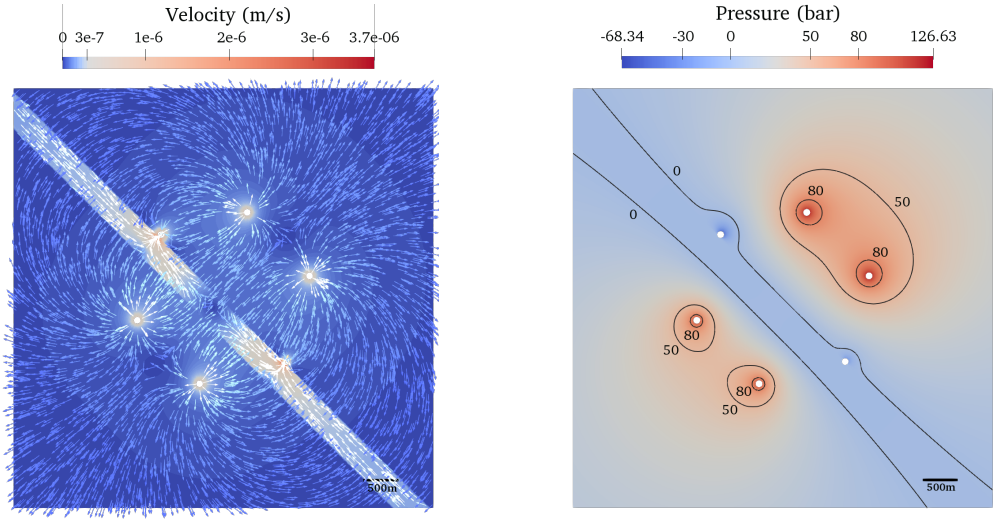



Figure 7.23: Setup ($H_{I4P2}$): Velocity with unscaled arrows (left) and pressure with contours (right) in the optimal configuration for the fault width 400 m. Control variables: rotation angle $\approx 2.35$ rad, hexagon radius $\approx 1271.4$ m.

During the first 80 years, the four cooling fronts develop essentially circular with a slight tendency towards the damage zone (Figure 7.24 left). After 133.55 years (Figure 7.24 right), the cold water fronts, associated with the two injection wells with smallest distance to the fault, have entered the fault zone and reached the production wells in form of channeled evolutions. This happens almost simultaneously, see Figure 7.25.

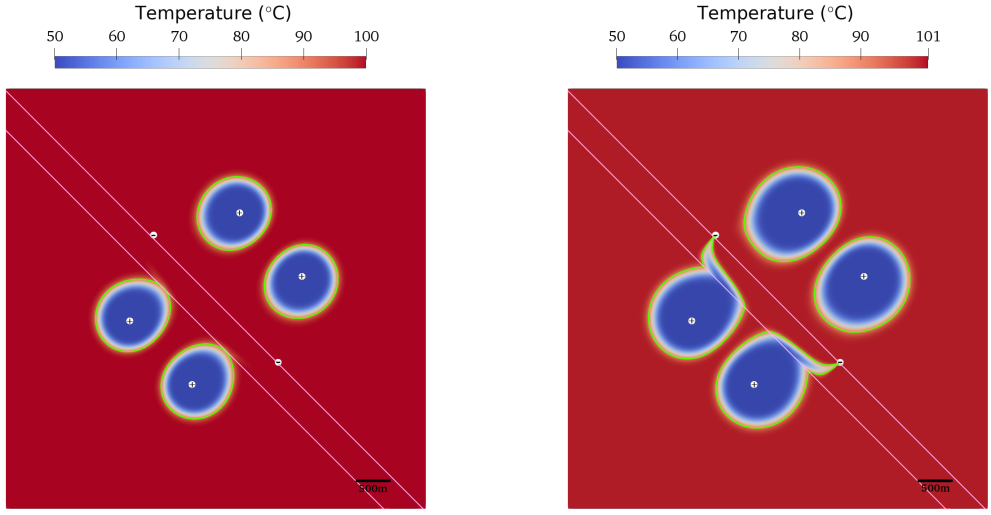



Figure 7.24: Setup ($H_{I4P2}$): Temperature distribution in the optimal configuration for the fault width 400 m after 80 years of operation (left) and after the *specific lifetime* 133.55 years (right). Control variables: rotation angle $\approx 2.35$ rad, hexagon radius $\approx 1271.4$ m. The green lines are 90 °C isolines.

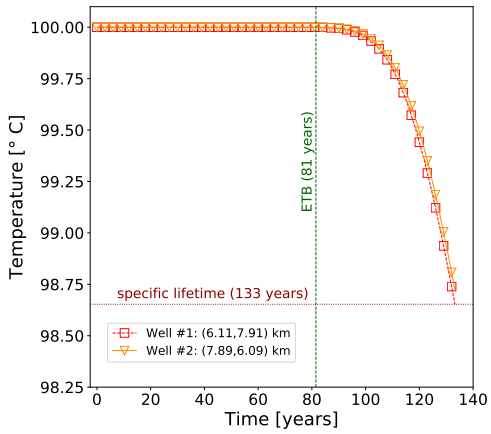



Figure 7.25: Setup ($H_{I4P2}$): Well-wise temperature evolution in the optimal configuration for the fault width 400 m. Control variables: rotation angle $\approx 2.35$ rad, hexagon radius $\approx 1271.4$ m.

### Setups ($H_{I2P4}$(i)) and ($H_{I2P4}$(ii)) – 2 Injection and 4 Production Wells

In this paragraph, we consider setup ($H_{I2P4}$(i)) for different damage zone widths. Additionally, we examine the setup ($H_{I2P4}$(ii)), in which the damage zone ($\text{width}^{\text{dz}} = 400$ m) is translated eastwards, such that it passes through the center of the hexagon. Thus, the optimization procedure is able to place two wells entirely inside the damage zone.

As in the hexagonal configurations considered so far, the hexagon radius primary controls the net energy. In particular, we see in Figure 7.26 that the setup with centered damage zone yields a higher net energy value (optimal at optimization step 36) than any of the other cases during the optimization.

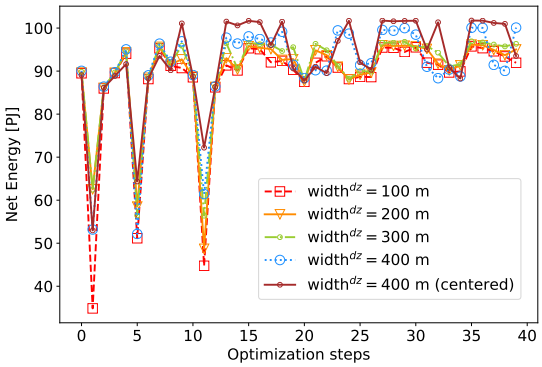



Figure 7.26: Setups ($H_{I2P4}$(i)) and ($H_{I2P4}$(ii)): Net energy collected over the simulation time in each of the 40 optimization steps, with the global optimization algorithm DIRECTL for different damage zone widths and a centered fault of width 400 m.

Similar as for the previous setup ($H_{I4P2}$) but with well tasks and flow rates interchanged, for **setup ($H_{I2P4}$(i))** the two injection wells, which have higher flow rates (doubled) than the production wells, are positioned near the damage zone, see Figure 7.27. Placing them close to or within a region of higher permeability allows to reduce the pressure difference between injectors and producers the most. In fact the control variables are very similar to the setup ($H_{I4P2}$).

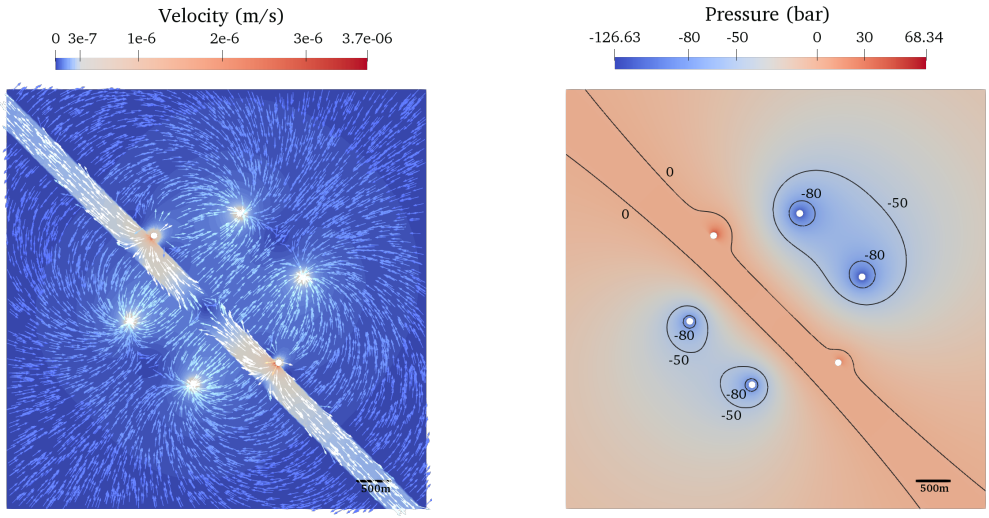



Figure 7.27: Setup ($H_{I2P4}$(i)): Velocity with unscaled arrows (left) and pressure with contours (right) in the optimal configuration for the fault width 400 m. Control variables: rotation angle ≈ 2.35 rad, hexagon radius ≈ 1271.4 m.

The cold water plumes evolve primary along the damage zone which has a higher permeability, see Figure 7.28. After 80 years of operation, the cold water front has passed and left the fault towards the production wells on the other side of the fault. The two production wells located at the same side as the injection wells (with respect to the damage zone) have a very limited influence on the cold water plume. This can be also seen in the well-wise temperature evolutions, depicted in Figure 7.28).

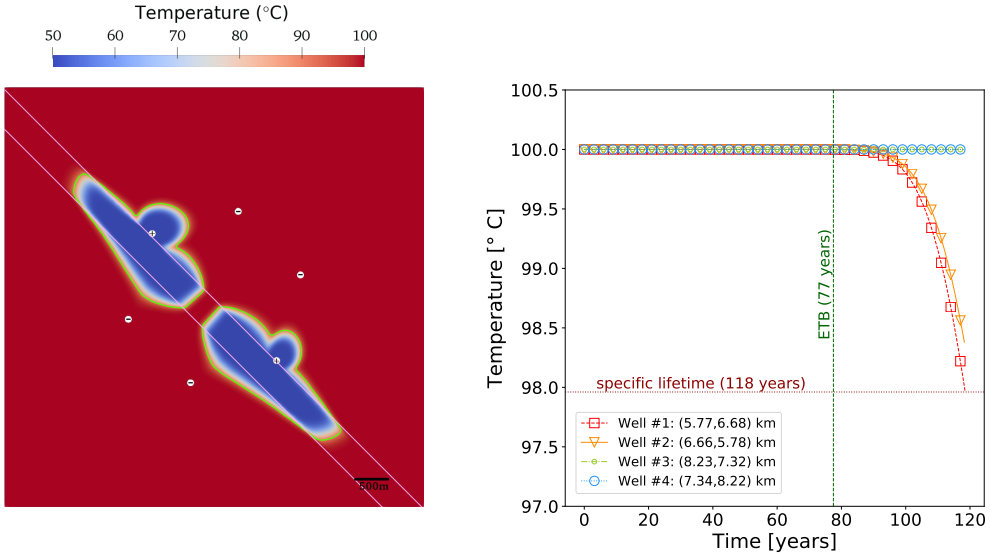



Figure 7.28: Setup ($H_{I2P4}$(i)): Temperature distribution in the optimal configuration for the fault width 400 m after 80 years of operation (left) and well-wise temperature evolution (right). Control variables: rotation angle ≈ 2.35 rad, hexagon radius ≈ 1271.4 m. The green line (left) is a 90 °C isoline.

The results for the **setup ($\mathbf{H_{I2P4}}$(ii))** confirm that the positioning of the wells with highest flow rates in the region with highest permeability maximizes the net energy in the considered setup, see Figures 7.29 - 7.31.

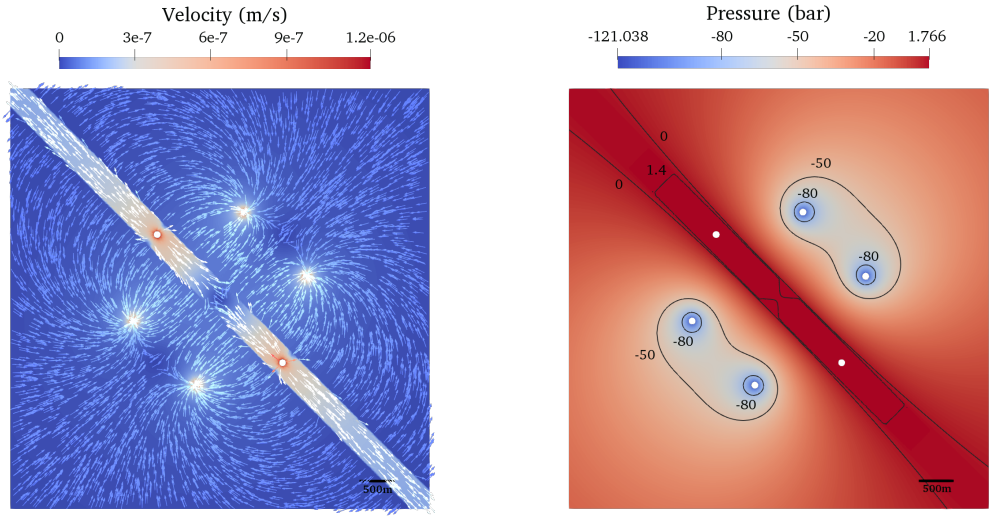



Figure 7.29: Setup ($H_{I2P4}$(ii)): Velocity with unscaled arrows (left) and pressure with contours (right) in the optimal configuration (optimization step 36) for the fault of width 400 m through the center of the hexagon. Control variables: rotation angle ≈ 2.35 rad, hexagon radius ≈ 1280.9 m.

Since the center of the hexagon structure is situated exactly on the main axis of a damage zone, the optimization is able to place two wells entirely inside the fault zone. This results in very small (relative) pressures for these injection wells (in particular smaller than in Figure 7.27, right), although their flow rates are twice as large as for each of the four production wells.

As a consequence of this positioning, the cooling front is channeled in the high permeability zone and the *specific lifetime* is extended to 171 years, see Figure 7.30). After 80 years, the cold water front has started to leave the damage zone towards the production wells, forming two triplets whose cold water fronts finally tend to connect to each other.

The optimized control variables are similar to the previous hexagonal setups, but here the hexagon radius is slightly larger. We see in Figure 7.31 that the *earliest thermal breakthrough* does not happen within the 80 years of net energy gain (in contrast to setup ($H_{I2P4}$(i)), Figure 7.28, right). This explains the energetic superiority of the centered fault, see Figure 7.26. In comparison to the previous setup ($H_{I2P4}$(i)), the *specific lifetime* is around 50 years longer. It is worth mentioning that, since the center of the hexagon is not exactly located in the central axis of the damage zone, a slightly asymmetric evolution of the cooling front is observed. This explains why the *specific lifetime* for two production wells slightly differs from the other two production wells, see Figure 7.31.

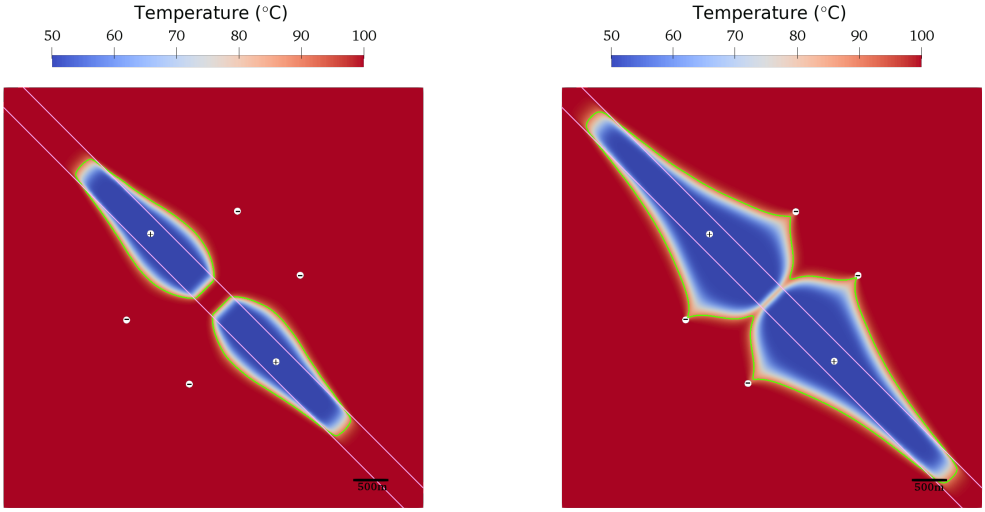



Figure 7.30: Setup ($H_{I2P4}$(ii)): Temperature distribution in the optimal configuration (optimization step 36) for the fault of width 400 m (centered), after 80 years of operation (left) and at the end of the *specific lifetime* of 171 years (right). Control variables: rotation angle $\approx 2.35$ rad, hexagon radius $\approx 1280.9$ m. The green lines are 90 °C isolines.

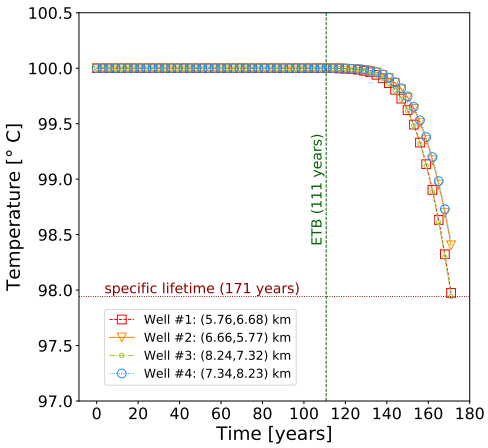



Figure 7.31: Setup ($H_{I2P4}$(ii)): Well-wise temperature evolution in the optimal configuration (optimization step 36) for the fault of width 400 m (centered). Control variables: rotation angle $\approx 2.35$ rad, hexagon radius $\approx 1280.9$ m.

### Setup ($H_{I2P4}$(iii)) – 2 Injection and 4 Production Wells (Healed)

Finally we present the simulation results for the setup ($H_{I2P4}$(iii)) which is characterized by a sightly healed damage zone (lower permeability) of width 200 m.

On the energetic level, this setup is compared to the case of a homogeneous permeability distribution (without the healed fault) in Figure 7.32.

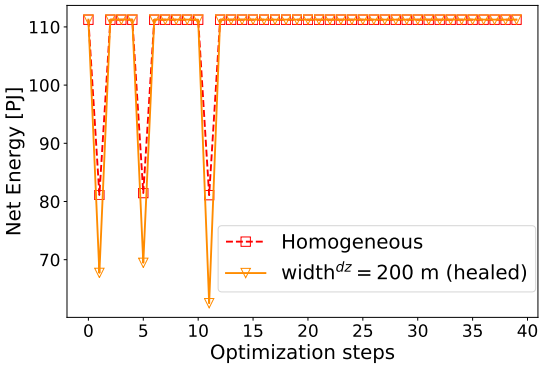



Figure 7.32: Setup ($H_{I2P4}$(iii)) versus its homogeneous equivalent: Net energy collected over the simulation time in each of the 40 optimization steps, with the global optimization algorithm DIRECTL.

During the 40 optimization steps, the maximum net energy for both scenarios is comparable on the considered scale. After the initially oscillating behavior for small radii, the net energy stabilizes to an almost constant value.

The slightly healed damage zone has a lower permeability and thus decomposes the reservoir into two sectors with the same homogeneous permeability. The

optimal placement is realized by placing a triplet on each side of the healed fault, see Figures 7.33 and 7.34.

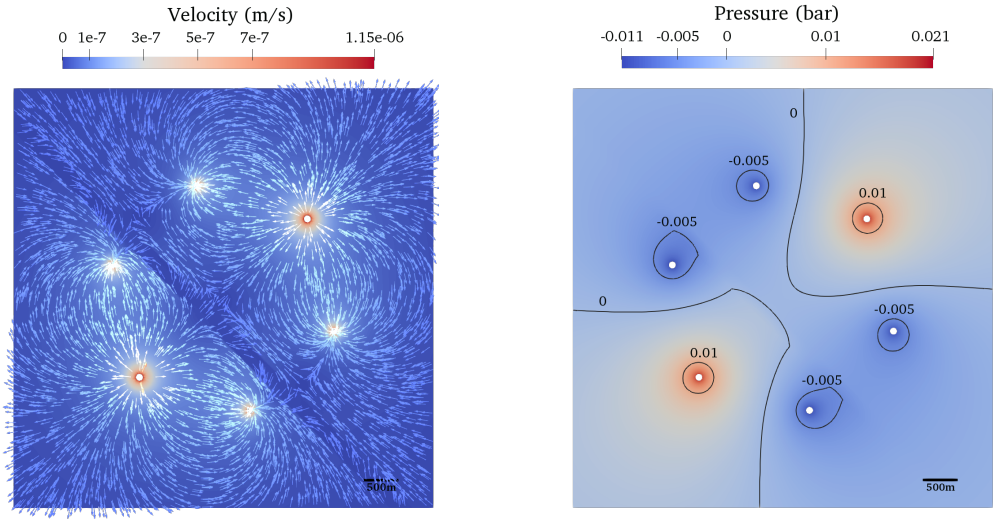



Figure 7.33: Setup ($H_{I2P4}$(iii)): Velocity with unscaled arrows (left) and pressure with contours (right) in the optimal configuration for the damage zone of width 200 m (healed). Control variables: rotation angle $\approx 0.76$ rad, hexagon radius $\approx 1650$ m.

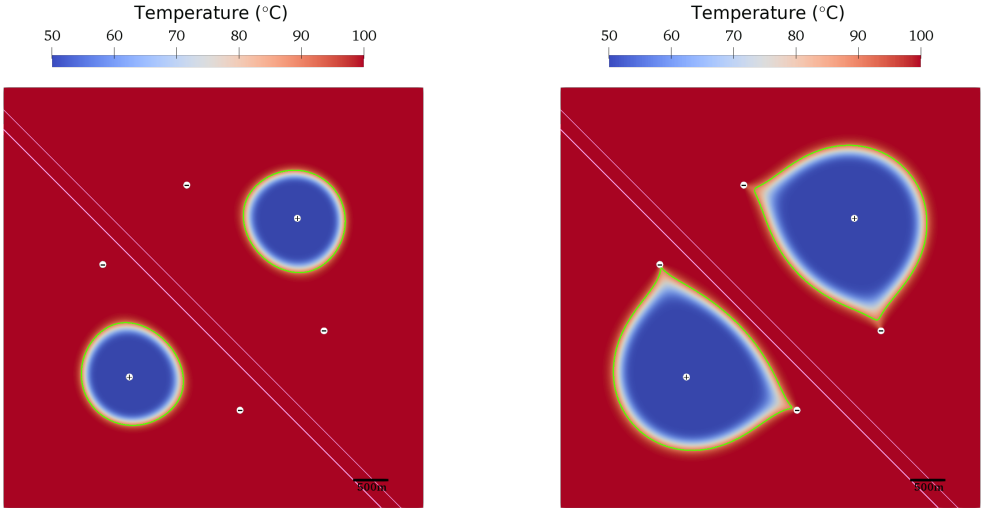



Figure 7.34: Setup ($H_{I2P4}$(iii)): Temperature distribution in the optimal configuration for the damage zone of width 200 m (healed), after 80 years of operation (left) and at the end of the *specific lifetime* of 205.75 years (right). Control variables: rotation angle $\approx 0.76$ rad, hexagon radius $\approx 1650$ m. The green line is a 90 °C isoline.

Due to the geometrical constraints, one triplet is placed closer to the healed damage zone than the other triplet and both are slightly rotated with respect

to the main axis of the damage zone. The *specific lifetime* for this scenario corresponds to 205.75 years, see also Figure 7.35. During this time, the cold water fronts evolve without noticeable interaction, with each other or with the fault.

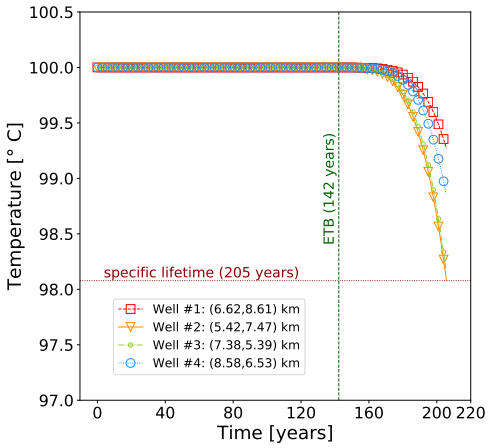



Figure 7.35: Setup ($H_{I2P4}$(iii)): Well-wise temperature evolution in the optimal configuration for the healed damage zone of width 200 m. Control variables: rotation angle $\approx 0.76$ rad, hexagon radius $\approx 1650$ m.

**Remark 7.3.1** (Hexagon Setup with Maximum Net Energy)**.**
*On a first glance, comparing Figures 7.19, 7.22, 7.26, and 7.32, the alternating arrangement might be concluded to yield the highest net energy. This is misleading, since the setups have different exploitation schemes. Note, e.g., that setup ($H_{I3P3}$) has a total flow rate of* 600 l/s *while the other hexagonal setups are subject to only* 400 l/s*. Concerning the gained energy, the flow rates directly enter the net energy formula* (7.2.9)*, but they also have an impact on the pressure and the advection (velocity), which influence both, energy losses and produced energy.*

## 7.4 Summary

The objective of this chapter was to develop a fit for purpose computational framework, aiming at optimizing geothermal energy production of selected multi-well configurations. Here, constraints on the operational time, nontrivial temperature and permeability structures, and thermo-hydraulic interference of neighboring wells are crucial factors that were incorporated.

Based on well accepted model assumptions and realistic conditions in hot sedimentary aquifers, a thermo-hydraulic model approach has been presented. It sequentially couples the Brinkman equations with the heat transport equation (energy balance) in a two-dimensional model of the aquifer.

At the discrete level, we used the proposed finite element method for the Brinkman problem (see Chapter 5), which is capable of a large range of geological conditions. The Brinkman velocity was invoked as advection in the heat transport

equation, discretized in time using an implicit Euler method and in space with linear finite elements.

In order to overcome the difficulties due to the presence of multiple spatial scales (well sizes are much smaller than the domain size), an immersed boundary method was applied. The presented non-matching approach allows to incorporate the wells in terms of sources or sinks of mass and heat, meanwhile separating the mesh generation from the well scale. Thereby the thermo-hydraulic characteristics outside a chosen artificial, enlarged well region are maintained, which has been verified in a dedicated validation study. This methodology is particularly favorable for an (automatic) optimization procedure which spatially relocates wells.

A global, gradient-free optimization algorithm was utilized for the detection of energetically privileged multi-well configurations in a heterogeneous aquifer.

We studied two type of multi-well arrangements: lattice-type and hexagonal configurations. In both cases, homogeneous and heterogeneous geophysical conditions were investigated.

Focusing on the produced net energy, the simulation results advocate specific multi-well placements for the combination of the considered geothermal and hydrogeological conditions and the induced exploitation strategies. In particular, the inter-well distance has been identified in these settings as the most powerful control, which might strongly restrict the lifetime, and thus the produced net energy of a plant. In general, larger inter-well distances lead to longer specific lifetimes which tend to result in more produced energy. In case that the inter-well distance is fixed or large enough, the configurations are visibly adapted to the permeability structure and the formation temperature (if non-constant) by using primarily the remaining controls.

The presented results serve as a starting point for further investigations of smart geothermal multi-well arrangements in heterogeneous formations.

# List of Principal Notations

| Symbol | Description | Section |
|---|---|---|
| $A$; $A_{w^{(\star)}}$ | area; area of $w^{(\star)}$ | 2.2.3; 7.1.3 |
| $\alpha_{BJ}$ | Beavers–Joseph constant (slip coefficient) | 2.3.2 |
| $a(\cdot,\cdot)$; $a_h(\cdot,\cdot)$ | bilinear forms | 3.1.3; 4.1 |
| $\mathcal{A}[(\cdot,\cdot);(\cdot,\cdot)]$ | bilinear form | 3.1.3 |
| $\mathcal{A}_h^{(\pm)}[(\cdot,\cdot);(\cdot,\cdot)]$ | bilinear form with specific signs | 4.2 |
| $\mathcal{A}$ | FE matrix for the Brinkman problem (standard Galerkin formulation) | 4.2 |
| $\mathcal{A}^{\mathrm{nsGLS}}$ | FE matrix for the Brinkman problem, standard Galerkin with non-symmetric GLS | 4.3.2 |
| $\alpha$, $\alpha_{\boldsymbol{T}}$ | GLS parameters | 4.3.2 |
| $\mathcal{A}^{\mathrm{nspf}}$ | FE matrix for the Brinkman problem, only contributions from the penalty-free non-symmetric Nitsche method | 4.4.1 |
| $\alpha_T$, $\alpha_L$ | transversal and longitudinal thermal dispersion | 7.1.6 |
| $b(\cdot,\cdot)$, $b_h(\cdot,\cdot)$ | bilinear forms | 3.1.3, 4.1 |
| $\beta$, $\beta_h$ | (continuous and discrete) inf-sup constants | 3.1.3; 4.1 |
| $C_P$ | constant in the Poincaré inequality | 3.1.2 |
| $C_{SR}$ | shape-regularity constant | 4.1 |
| $C_{QU}$ | quasi-uniformity constant | 4.1 |
| $\hat{c}_I$, $c_I$ | inverse inequality constants | 4.1 |
| $c_{\mathrm{DTI}}, \widehat{c}_{\mathrm{DTI}}, \widetilde{c}_{\mathrm{DTI}}$ | discrete trace-inverse inequality constants | 4.1 |
| $C(\cdot)$ | space of continuous functions | 4.3.1 |
| $c_{La}, c_{SZ}, \overline{c}_{\mathrm{SZ}}$ | constants associated with interpolation error estimates | 5.1 |
| $\mathcal{C}$ | set of corner nodes of the discrete domain | 5.2.1 |
| $C_s$, $C_f$ | material specific heat capacities of the solid and the fluid | 7.1.6 |
| $\delta$ | grad-div stabilization parameter | 5.2.2 |
| $\delta_{w^{(\star)}}$ | Dirac delta distribution w.r.t. the center of the well $w^{(\star)}$ | 7.1.5 |
| $\delta^{r_\varepsilon}_{w^{(\star)}}$ | approximate (continuous) delta function with support area $\pi r_\varepsilon^2$ associated with the well $w^{(\star)}$ | 7.2.1 |
| $E$ | $(n-1)$-dimensional subsets of $\boldsymbol{\Omega} \subset \mathbb{R}^n$, e.g., edges for $n=2$ | 4.1 |
| $\eta_0$ | bound in assumption $(\boldsymbol{M}_1)$ | 5.2.1 |

| | | |
|---|---|---|
| $\varepsilon, \varepsilon_{w^{(\star)}}$ | pump efficiency (associated with $w^{(\star)}$), $\in (0,1]$ | 7.2.3 |
| $E_{net}$ | net energy | 7.2.3 |
| $E_{prod}$, $E_{pump}$ | produced respectively pump energy | 7.2.3 |
| $\boldsymbol{f}$ | external forces | 2.3.1 |
| $\boldsymbol{f}_{\text{int}}, \boldsymbol{f}_{\text{ext}}$ | internal, external forces | 2.3.1 |
| $f(\cdot)$, $f_h(\cdot)$ | linear forms | 3.1.3; 4.1 |
| $\mathcal{F}_h^{(\pm)}[(\cdot,\cdot)]$ | linear form with specific signs | 4.2 |
| $\Gamma$ | boundary of $\boldsymbol{\Omega}$, $\Gamma := \partial\boldsymbol{\Omega}$ | 2.3 |
| $g$, $\boldsymbol{g}$ | gravitational acceleration; source/sink of mass | 2.3; 3.2 |
| $g(\cdot)$, $g_h(\cdot)$ | linear forms | 3.1.3, 4.1 |
| $\Gamma_E$, $\Gamma_N$ | boundary parts associated with essential and natural boundary conditions | 3.2 |
| $\mathcal{G}_h$ | set of boundary facets in a triangulation | 4.1 |
| $\gamma_N$ | Nitsche penalty parameter | 4.4.1 |
| $\Gamma_{w^{(\star)}}$, $\Gamma_{w_k^{inj}}$ | boundaries of $w^{(\star)}$ and $w_k^{inj}$ | 7.1.3, 7.1.6 |
| $g_{w^{(\star)}}$ | constant singular (point) force associated with the well $w^{(\star)}$ (reduced model) | 7.1.5 |
| $\widetilde{\gamma}_k$, $\gamma_k$ | penalty parameters associated with a well $w_k^{inj}$ | 7.1.7 |
| $H$ | hydraulic head; thickness of the aquifer model | 2.3.2; 7.1.3 |
| $H^k(\cdot)$ | Sobolev spaces $W^{k,2}(\cdot)$ | 3.1.1 |
| $H_0^1(\boldsymbol{\Omega})$ | functions from $H^1(\boldsymbol{\Omega})$, vanishing at the boundary | 3.1.1 |
| $H^{-1}(\boldsymbol{\Omega})$ | dual space of $H_0^1(\boldsymbol{\Omega})$ | 3.1.1 |
| $\boldsymbol{H}^{\text{div}}(\boldsymbol{\Omega})$ | $L^2$-vector fields with divergence in $L^2(\boldsymbol{\Omega})$ | 3.1.1 |
| $\boldsymbol{H}_0^{\text{div}}(\boldsymbol{\Omega})$ | vector fields in $\boldsymbol{H}^{\text{div}}(\boldsymbol{\Omega})$ with vanishing normal component at the boundary | 3.1.1 |
| $h$ | mesh-size | 4.1 |
| $h_{\boldsymbol{T}}$, $h_E$ | diameter of $\boldsymbol{T}$ and $E$ | 4.1 |
| $h_{\min}$, $h_{\max}$ | minimum, maximum element-size in a triangulation | 6.3 |
| $\mathbb{I}$ | identity tensor | 2.2.2 |
| $\mathbb{K}$, $k$ | tensorial and scalar permeability | 2.2.2 |
| $K$ | hydraulic conductivity | 2.3.2 |
| $L^p(\boldsymbol{\Omega})$, $L^\infty(\boldsymbol{\Omega})$ | Lebesgue spaces, $1 \leq p < \infty$ | 3.1.1 |
| $L_0^2(\boldsymbol{\Omega})$ | subspace of $L^2$-functions with mean value zero | 3.1.1 |
| $I_{h,k}^{\text{La}}$ | Lagrange (nodal) interpolation operator | 5.1 |
| $\ell_{\boldsymbol{\Omega}}$ | characteristic length | 5.2.1 |
| level | mesh refinement | 6.1 |
| $\lambda_s, \lambda_f$ | scalar thermal conductivities (s = solid, f = fluid) | 7.1.6 |
| $\lambda$ | total thermal conductivity tensor | 7.1.6 |
| $\mu$ | dynamic viscosity | 2.2.3 |
| $\mu_{\text{eff}}$ | effective viscosity | 2.2.3 |
| $m$ | coercivity constant | 3.1.3 |
| $M$ | continuity/boundedness constant | 3.1.3 |
| $\widehat{\mu}$ | auxiliary viscosity, $\widehat{\mu} := \mu_{\text{eff}} + \sigma\ell_{\boldsymbol{\Omega}}^2$ | 5.2.1 |

| | | |
|---|---|---|
| $\mathbb{N}_{>0}$ | natural numbers (excluding 0) | 2.1 |
| $n$ | spatial dimension (usually of $\boldsymbol{\Omega}$) | 2.1 |
| $\nu$ | kinematic viscosity | 2.2.3 |
| $\boldsymbol{n}$; $\boldsymbol{n}_{w^{(\star)}}$ | unit outer normal vector on $\Gamma$; unit normal vetor on $\Gamma_{w^{(\star)}}$ (directed into $\boldsymbol{\Omega}$) | 2.3; 7.1.4 |
| $N_U$, $N_P$ | number of velocity and pressure dofs | 4.2 |
| $N^{inj}, N^{prod}$ | number of injection and production wells | 7.1.3 |
| $\boldsymbol{\Omega}$ | domain in $\mathbb{R}^n$, (often $n = 2, 3$); horizontal slice of the perforated aquifer | 2.1; 7.1.3 |
| $\omega$ | bound in assumption $(\boldsymbol{M}_2)$ | 5.2.1 |
| $\widehat{\boldsymbol{\Omega}}$ | extended domain including the wells $w^{(\star)}$ | 7.1.5 |
| $\phi$ | porosity in $[0, 1]$ | 2.2.2, 7.1.6 |
| $p$ | pressure field | 2.3.1 |
| $\mathbb{P}_k$ | polynomials of degree $k$ and smaller | 4.1 |
| $\boldsymbol{P}$ | vector with pressure dofs | 4.2 |
| $Q$ | volumetric flow rate; | 2.3.2; |
| | abstract Hilbert space, weak pressure space | 3.1.3, 3.3 |
| $Q_h$ | discrete pressure space | 4.1 |
| $q$, $q_h$ | pressure test functions | 3.3, 4.1 |
| $Q_h^k$ | discrete pressure space with continuous, element-wise polynomials of degree $k$ | 4.3.1 |
| $q_h^{\boldsymbol{u}_h}$ | special pressure test function | 5.3.1 |
| $Q_{w_{3D}^{(\star)}}$ | flow rate associated to the well | 7.1.4 |
| $\mathbb{R}$ | real numbers | 2.1 |
| $\rho$; $\rho_s, \rho_f$ | density; densities of the solid and the fluid | 2.2.3; 7.1.6 |
| $\rho_{\boldsymbol{T}}$ | radius of largest inscribed sphere in $\boldsymbol{T}$ | 4.1 |
| $\rho$ | corner stabilization parameter | 5.2.2 |
| $r_{w^{(\star)}}$ | radius of the well $w^{(\star)}$ | 7.1.4 |
| $\rho\, C$ | volumetric (macroscopic) heat capacity | 7.1.6 |
| $r_\varepsilon$ | artificial well radius (IBM) | 7.2.1 |
| $S_f$ | saturation w.r.t. the fluid f | 2.2.2 |
| $\mathbb{S}$ | total/Cauchy stress tensor | 2.3.1 |
| $\sigma$ | ratio of dyn. viscosity and permeability $\sigma := \mu\mathbb{K}^{-1}$ | 3.2 |
| $\mathcal{S}_h^\star$ | generic residual-based pressure stabilization | 4.3.2 |
| $\mathcal{S}_h^{GLS}$, $S_{h,\alpha}^{GLS}$ | GLS stabilization | 4.3.2, 5.2.2 |
| $\mathcal{S}_h^{PSPG}$ | pressure-stabilized Petrov–Galerkin stabilization | 4.3.2 |
| $I_{h,k}^{\mathrm{SZ}}$, $I_{h,k}^{\mathrm{SZ}^{\int^0}}$ | Scott–Zhang interpolation operators | 5.1 |
| $S(\boldsymbol{T})$ | union of all cells in $\mathcal{T}_h$ sharing a vertex with $\boldsymbol{T}$ | 5.1 |
| $S_{h,\delta}^{GD}$ | grad-div stabilization | 5.2.2 |
| $S_{h,\rho}^{\mathcal{C}}$ | corner stabilization | 5.2.2 |
| $t$ | time variable | 2.3.1 |
| $T$, $T(t, \boldsymbol{x})$ | time interval bound; temperature field | 2.3.1; 7.1.6 |
| $\boldsymbol{t}$ | tangential vector | 2.3.2 |

| | | |
|---|---|---|
| $\mathcal{T}_h$ | triangulation | 4.1 |
| $\boldsymbol{T}$ | mesh cell, $n$-dimensional subsets of $\boldsymbol{\Omega} \subset \mathbb{R}^n$, e.g., triangles for $n = 2$ | 4.1 |
| $\widehat{\boldsymbol{T}}$ | reference triangle | 5.1 |
| $\theta$ | ratio of $\mu_{\text{eff}}$ and $\widehat{\mu}$ | 5.2.3 |
| $t_L$ | lifetime (of the geothermal plant) | 7.1.6 |
| $T_{w_k^{inj}}$ | injection temperature associated with the well $w_k^{inj}$ | 7.1.6 |
| $T_0(\boldsymbol{x})$ | initial (aquifer/formation) temperature field | 7.1.6 |
| $t^n$ | time at time step $n$ | 7.2.1 |
| $\Delta t_n$, $\Delta t$ | time step lengths | 7.2.1, 7.2.3 |
| $N_t$ | number of time steps | 7.2.3 |
| $\mathbb{T}_h$ | discrete temperature space with continuous, element-wise linear polynomials | 7.2.1 |
| $T_{quit}$ | production temperature, at which the operation of the geothermal plant is stopped | 7.2.2 |
| $\boldsymbol{u}$ | vector field, velocity | 2.1, 2.3.1 |
| $\boldsymbol{U}$ | vector with velocity dofs | 4.2 |
| $U_{w^{(\star)}}$ | velocity magnitude at the boundary of the well $w^{(\star)}$ | 7.1.4 |
| $\boldsymbol{u}_D$, $\boldsymbol{u}_S$ | velocity fields (D = Darcy region, S = Stokes region) | 2.3.2 |
| $\mathbb{V}$ | viscous stress tensor | 2.3.1 |
| $\boldsymbol{V}_0$ | abstract volume with boundary $\partial \boldsymbol{V}_0$ | 2.3.1 |
| $V$ | abstract Hilbert space | 3.1.3 |
| $\boldsymbol{V}$; $\boldsymbol{V}_h$ | weak velocity space; discrete/finite-dimensional velocity space | 3.3; 4.1 |
| $\boldsymbol{v}$, $\boldsymbol{v}_h$ | velocity test function | 3.3, 4.1 |
| $\boldsymbol{V}_0$ | space of weakly divergence-free velocities | 3.1.3 |
| $\boldsymbol{V}_{0,h}$ | space of discrete divergence-free velocities | 4.1 |
| $\boldsymbol{V}_h^k$ | discrete velocity space with continuous, element-wise polynomials of degree $k$ | 4.3.1 |
| $\boldsymbol{w}_h^{\boldsymbol{u}_h}$ | special velocity test function | 5.3.1 |
| $w^{(\star)}, w_{2D}^{(\star)}$ | generic horizontal slice of a cylindrical well (2D) | 7.1.3 |
| $w_{3D}^{(\star)}$ | generic cylindrical well (3D) | 7.1.3 |
| $w_k^{inj}$ | injection well (2D) | 7.1.7 |
| $x$ | space variable | 2.3.1 |
| $\Delta(\cdot)$ | Laplace operator | 2.1 |
| $\nabla(\cdot)$ | gradient operator | 2.1 |
| $\nabla \cdot (\cdot)$ | divergence operator | 2.1 |
| $\mathbb{D}(\cdot)$ | symmetric part of the gradient | 2.1 |
| $(\cdot)^T$ | transposed | 2.1 |
| $\|\cdot\|_{L^2(\boldsymbol{\Omega})}$, $\|\cdot\|_0$ | $L^2$-norm | 3.1.1 |
| $(\cdot,\cdot)$, $(\cdot,\cdot)_{L^2(\boldsymbol{\Omega})}$ | $L^2$-product | 3.1.1 |
| $\langle\cdot,\cdot\rangle$, $\langle\cdot,\cdot\rangle_{L^2(\Gamma)}$ | $L^2$-product on the boundary $\Gamma$ | 3.1.1 |
| $\|\cdot\|_k$, $\|\cdot\|_k$ | norm and semi norm on the Sobolev space $H^k(\cdot)$ | 3.1.1 |

| | | |
|---|---|---|
| $[\cdot,\cdot]$ | dual pairing | 3.1.1 |
| $\|\cdot\|, \|\cdot\|$ | Euclidean norm | 3.1.2 |
| $(\cdot)'$ | dual space | 3.1.1 |
| $\mathring{(\cdot)}$ | interior | 4.1 |
| $\overline{(\cdot)}$ | closure | 4.1 |
| $\operatorname{diam}(\cdot)$ | diameter | 4.1 |
| $\operatorname{rank}(\cdot)$ | rank | 4.2 |
| $\ker(\cdot)$ | kernel | 4.2 |
| $\operatorname{Im}(\cdot)$ | range | 4.2 |
| $[\![\cdot]\!]$ | jump | 4.3.2, 5.2.2 |
| $\|\|\|(\cdot,\cdot)\|\|\|_h$ | (mesh-dependent) energy norm | 5.2.3 |

| | |
|---|---|
| 1D, 2D, 3D | one-, two-, three-dimensional |
| FE | finite element |
| dof, DOF | degree of freedom |
| GLS | Galerkin least squares |
| PSPG | pressure-stabilized Petrov–Galerkin |
| nsGLS | non-symmetric GLS |
| GD | grad-div |
| C | corner |
| pfns | penalty-free non-symmetric |
| IBM | immersed boundary method |
| ETB | earliest thermal breakthrough |
| lhs, rhs | left-hand side, right-hand side |
| disc | discontinuous |

# Bibliography


[Ach90] D. Acheson. *Elementary Fluid Dynamics*. Oxford Applied Mathematics and Computing Science Series. The Clarendon Press, Oxford University Press, New York, 1990, pp. x+397.

[Ahm+18] N. Ahmed, C. Bartsch, V. John, and U. Wilbrandt. "An Assessment of some Solvers for Saddle Point Problems emerging from the Incompressible Navier–Stokes Equations". In: *Comput. Methods Appl. Mech. Engrg.* 331 (2018), pp. 492–513.

[AGL05] J. Ahrens, B. Geveci, and C. Law. "ParaView: An End-User Tool for Large Data Visualization". In: *Visualization Handbook* (2005).

[All91a] G. Allaire. "Homogenization of the Navier–Stokes Equations and Derivation of Brinkman's Law". In: *Mathématiques appliquées aux sciences de l'ingénieur (Santiago, 1989)*. Cépaduès, Toulouse, 1991, pp. 7–20.

[All90a] G. Allaire. "Homogenization of the Navier–Stokes Equations in Open Sets Perforated with Tiny Holes. I. Abstract Framework, a Volume Distribution of Holes". In: *Arch. Rational Mech. Anal.* 113.3 (1990), pp. 209–259.

[All90b] G. Allaire. "Homogenization of the Navier–Stokes Equations in Open Sets Perforated with Tiny Holes. II. Noncritical Sizes of the Holes for a Volume Distribution and a Surface Distribution of Holes". In: *Arch. Rational Mech. Anal.* 113.3 (1990), pp. 261–298.

[All91b] G. Allaire. "Homogenization of the Navier–Stokes Equations with a Slip Boundary Condition". In: *Comm. Pure Appl. Math.* 44.6 (1991), pp. 605–641.

[Ang99] P. Angot. "Analysis of Singular Perturbations on the Brinkman Problem for Fictitious Domain Models of Viscous Flows". In: *Math. Methods Appl. Sci.* 22.16 (1999), pp. 1395–1412.

[Ang11] P. Angot. "On the Well-Posed Coupling between Free Fluid and Porous Viscous Flows". In: *Appl. Math. Lett.* 24.6 (2011), pp. 803–810.

[ABF99] P. Angot, C.-H. Bruneau, and P. Fabrie. "A Penalization Method to take into account Obstacles in Incompressible Viscous Flows". In: *Numer. Math.* 81.4 (1999), pp. 497–520.

[ABS16] M. Antics, R. Bertani, and B. Sanner. "Summary of EGC 2016 Country Update Reports on Geothermal Energy in Europe". In: Proceedings European Geothermal Congress. France. 2016.

[Ant13] F. Antonov. "Numerical Determination of Permeability Tensor Components for 3D-braided Composites Using RVC Approach". In: *9th European LS-DYNA Conference* (2013).

[Ape99] T. Apel. "Interpolation of Non-Smooth Functions on Anisotropic Finite Element Meshes". In: *ESAIM: M2AN* 33.6 (1999), pp. 1149–1185.

[AL06] T. Arbogast and H. Lehr. "Homogenization of a Darcy–Stokes System modeling Vuggy Porous Media". In: *Comput. Geosci.* 10.3 (2006), pp. 291–302.

[Arn82] D. Arnold. "An Interior Penalty Finite Element Method with Discontinuous Elements". In: *SIAM J. Numer. Anal.* 19.4 (1982), pp. 742–760.

[ABF84] D. Arnold, F. Brezzi, and M. Fortin. "A Stable Finite Element for the Stokes Equations". In: *Calcolo* 21.4 (1984), pp. 337–344.

[AC03] M. Aubertin and R. Chapuis. *Predicting the Coefficient of Permeability of Soils using the Kozeny-Carman Equation*. Rapport technique. École polytechnique de Montréal, 2003.

[Aur09a] J.-L. Auriault. "About the Beavers and Joseph Boundary Condition". In: *Transport in Porous Media* 83.2 (2009), pp. 257–266.

[Aur09b] J.-L. Auriault. "On the Domain of Validity of Brinkman's Equation". In: *Transport Porous Media* 79.2 (2009), pp. 215–223.

[BO80] I. Babuška and J. Osborn. "Analysis of Finite Element Methods for Second Order Boundary Value Problems using Mesh Dependent Norms". In: *Numer. Math.* 34.1 (1980), pp. 41–62.

[Bab72] I. Babuška. "The Finite Element Method with Lagrangian Multipliers". In: *Numer. Math.* 20 (1972/73), pp. 179–192.

[Bab73] I. Babuška. "The Finite Element Method with Penalty". In: *Math. of Comp.* 27 (1973), pp. 221–228.

[BC10] S. Badia and R. Codina. "Stabilized Continuous and Discontinuous Galerkin Techniques for Darcy Flow". In: *Comput. Methods Appl. Mech. Engrg.* 199.25-28 (2010), pp. 1654–1667.

[BC12] S. Badia and R. Codina. "Stokes, Maxwell and Darcy: a Single Finite Element Approximation for Three Model Problems". In: *Appl. Numer. Math.* 62.4 (2012), pp. 246–263.

[BC09] S. Badia and R. Codina. "Unified Stabilized Finite Element Formulations for the Stokes and the Darcy Problems". In: *SIAM J. Numer. Anal.* 47.3 (2009), pp. 1971–2000.

[BH92] H. Barbosa and T. Hughes. "Boundary Lagrange Multipliers in Finite Element Methods: Error Analysis in Natural Norms". In: *Numer. Math.* 62.1 (1992), pp. 1–15.

[BV02] G. Barrenechea and F. Valentin. "An Unusual Stabilized Finite Element Method for a Generalized Stokes Problem". In: *Numerische Mathematik* 92.4 (2002), pp. 653–677.

[BE86] J. Barrett and C. Elliott. "Finite Element Approximation of the Dirichlet Problem using the Boundary Penalty Method". In: *Numer. Math.* 49.4 (1986), pp. 343–366.

[Bar94] W. Bartz. *Viskosität und Fliessverhalten: Glossary*. Handbuch der Tribologie und Schmierungstechnik. Expert, 1994.

[Bat99] G. K. Batchelor. *An Introduction to Fluid Dynamics*. paperback. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1999, pp. xviii+615.

[BH07a] Y. Bazilevs and T. Hughes. "Weak Imposition of Dirichlet Boundary Conditions in Fluid Mechanics". In: *Computers & Fluids* 36.1 (2007), pp. 12–26.

[Bea72] J. Bear. *Dynamics of Fluids in Porous Media*. Dover Civil and Mech. Engrg. Dover Publications, 1972.

[BJ67] G. Beavers and D. Joseph. "Boundary Conditions at a Naturally Permeable Wall". In: *Journal of Fluid Mech.* 30.1 (1967), pp. 197–207.

[BSM70] G. Beavers, E. Sparrow, and R. Magnuson. "Experiments on Coupled Parallel Flows in a Channel and a Bounding Porous Medium". In: *Journal of Fluids Engrg.* 92.4 (1970), pp. 843–848.

[BB01] R. Becker and M. Braack. "A Finite Element Pressure Gradient Stabilization for the Stokes Equations based on Local Projections". In: *Calcolo* 38.4 (2001), pp. 173–199.

[BHS03] R. Becker, P. Hansbo, and R. Stenberg. "A Finite Element Method for Domain Decomposition with Non-Matching Grids". In: *M2AN Math. Model. Numer. Anal.* 37.2 (2003), pp. 209–225.

[Ber14] C. Berg. "Permeability Description by Characteristic Length, Tortuosity, Constriction and Porosity". In: *Transport in Porous Media* 103.3 (2014), pp. 381–400.

[BS90] L. Bergmann and C. Schaefer. *Lehrbuch der Experimentalphysik: Mechanik, Akustik, Wärme. Band I*. Vol. 10. De Gruyter, 1990.

[BGR08] C. Bernardi, V. Girault, and K. Rajagopal. "Discretization of an Unsteady Flow Through Porous Solid modeled by Darcy's Equations". In: *Math. Models Methods Appl. Sci.* 18.12 (2008), pp. 2087–2123.

[Ber+17] R. Bertani, P. Dumas, J. Bonafin, O. Flòvenz, B. Jónsdóttir, A. Manzella, A. Donato, G. Gola, A. Santilano, E. Trumpy, S. Simsek, J.-D. van Wees, M. Pluymaekers, H. Veldkamp, S. van Gessel, D. Bonté, L. Rybach, B. Sanner, and L. Angelino. *Perspectives for Geothermal Energy in Europe*. World Scientific Publishing Europe Ltd., New York, 2017.

[BSL06] R. Bird, W. Stewart, and E. Lightfoot. *Transport Phenomena*. Revised 2nd Edition. John Wiley & Sons, Inc., 2006.

[Bla+18] L. Blank, A. Caiazzo, F. Chouly, A. Lozinski, and J. Mura. "Analysis of a Stabilized Penalty-Free Nitsche Method for the Brinkman, Stokes, and Darcy Problems". In: *ESAIM Math. Model. Numer. Anal.* 52.6 (2018), pp. 2149–2185.

[BDG06] P. Bochev, C. Dohrmann, and M. Gunzburger. "Stabilization of Low-Order Mixed Finite Elements for the Stokes Equations". In: *SIAM J. Numer. Anal.* 44.1 (2006), pp. 82–101.

[Boe12] R. de Boer. *Theory of Porous Media*. Highlights in the Historical Development and Current State. Springer Berlin Heidelberg, 2012.

[BBF13] D. Boffi, F. Brezzi, and M. Fortin. *Mixed Finite Element Methods and Applications*. Vol. 44. Springer Series in Comp. Math. Springer, Heidelberg, 2013, pp. xiv+685.

[BB16] T. Boiveau and E. Burman. "A Penalty-Free Nitsche Method for the Weak Imposition of Boundary Conditions in Compressible and Incompressible Elasticity". In: *IMA J. Numer. Anal.* 36.2 (2016), pp. 770–795.

[Boi+18] T. Boiveau, E. Burman, S. Claus, and M. Larson. "Fictitious Domain Method with Boundary Value Correction using Penalty-Free Nitsche Method". In: *J. Numer. Math.* 26.2 (2018), pp. 77–95.

[BS11] M. Braack and F. Schieweck. "Equal-Order Finite Elements with Local Projection Stabilization for the Darcy–Brinkman Equations". In: *Comput. Methods Appl. Mech. Engrg.* 200.9-12 (2011), pp. 1126–1136.

[Bra+07] M. Braack, E. Burman, V. John, and G. Lube. "Stabilized Finite Element Methods for the Generalized Oseen Problem". In: *Comput. Methods Appl. Mech. Engrg.* 196.4-6 (2007), pp. 853–866.

[BN13] M. Braack and K. Nafa. "A Monolithic Finite Element Discretization for Coupled Darcy–Stokes Flow". In: *PAMM* 13.1 (2013), pp. 243–244.

[Bra07] D. Braess. *Finite Elements*. Third Edition. Theory, Fast Solvers, and Applications in Elasticity Theory, Translated from the German by Larry L. Schumaker. Cambridge University Press, Cambridge, 2007, pp. xviii+365.

[BS08] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Third Edition. Vol. 15. Texts in Appl. Math. Springer, New York, 2008, pp. xviii+397.

[Bre74] F. Brezzi. "On the Existence, Uniqueness and Approximation of Saddle-Point Problems arising from Lagrangian Multipliers". In: *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge* 8.R-2 (1974), pp. 129–151.

[BP84] F. Brezzi and J. Pitkäranta. "On the Stabilization of Finite Element Approximations of the Stokes Equations". In: *Efficient Solutions of Elliptic Systems*. Vol. 10. Notes Numer. Fluid Mech. Friedr. Vieweg, Braunschweig, 1984, pp. 11–19.

[BF91] F. Brezzi and M. Fortin. *Mixed and hybrid finite element methods*. Vol. 15. Springer Series in Computational Mathematics. Springer, New York, 1991, pp. x+350.

[BFM93] F. Brezzi, M. Fortin, and L. Marini. "Mixed Finite Element Methods with Continuous Stresses". In: *Math. Models and Methods in Appl. Sci.* 3.2 (1993), pp. 275–287.

[Bri49] H. C. Brinkman. "A Calculation of the Viscous Force Exerted by a Flowing Fluid on a Dense Swarm of Particles". In: *Appl. Sci. Research* 1.1 (1949), pp. 27–34.

[Bur12] E. Burman. "A Penalty-Free Nonsymmetric Nitsche-Type Method for the Weak Imposition of Boundary Conditions". In: *SIAM J. Numer. Anal.* 50.4 (2012), pp. 1959–1981.

[Bur08] E. Burman. "Pressure Projection Stabilizations for Galerkin Approximations of Stokes' and Darcy's Problem". In: *Numer. Methods Partial Differential Equations* 24.1 (2008), pp. 127–143.

[BH06a] E. Burman and P. Hansbo. "A Stabilized Non-Conforming Finite Element Method for Incompressible Flow". In: *Comput. Methods Appl. Mech. Engrg.* 195.23-24 (2006), pp. 2881–2899.

[BH07b] E. Burman and P. Hansbo. "A Unified Stabilized Method for Stokes' and Darcy's Equations". In: *J. Comput. Appl. Math.* 198.1 (2007), pp. 35–51.

[BH06b] E. Burman and P. Hansbo. "Edge Stabilization for the Generalized Stokes Problem: A Continuous Interior Penalty Method". In: *Comput. Methods Appl. Mech. Engrg.* 195.19-22 (2006), pp. 2393–2410.

[BH12] E. Burman and P. Hansbo. "Fictitious Domain Finite Element Methods using Cut Elements: II. A Stabilized Nitsche Method". In: *Appl. Numer. Math.* 62.4 (2012), pp. 328–341.

[BH05] E. Burman and P. Hansbo. "Stabilized Crouzeix–Raviart Element for the Darcy-Stokes Problem". In: *Numer. Methods Partial Differential Equations* 21.5 (2005), pp. 986–997.

[BZ12] E. Burman and P. Zunino. "Numerical Approximation of Large Contrast Problems with the Unfitted Nitsche Method". In: *Frontiers in Numerical Analysis–Durham 2010*. Vol. 85. Lect. Notes Comput. Sci. Eng. Springer, Heidelberg, 2012, pp. 227–282.

[CJW14] A. Caiazzo, V. John, and U. Wilbrandt. "On Classical Iterative Subdomain Methods for the Stokes–Darcy Problem". In: *Comput. Geosci.* 18.5 (2014), pp. 711–728.

[Cao+10] Y. Cao, M. Gunzburger, F. Hua, and X. Wang. "Coupled Stokes–Darcy Model with Beavers–Joseph Interface Boundary Condition". In: *Commun. Math. Sci.* 8.1 (2010), pp. 1–25.

[CW07a] M. Cardenas and J. Wilson. "Dunes, Turbulent Eddies, and Interfacial Exchange with Permeable Sediments". In: *Water Resources Research* 43.8 (2007).

[CW07b] M. Cardenas and J. Wilson. "Hydrodynamics of Coupled Flow Above and Below a Sediment-Water Interface with Triangular Bedforms". In: *Advances in Water Resources* 30.3 (2007), pp. 301–313.

[Car97] P. Carman. "Fluid Flow Through Granular Beds". In: *Chem. Engrg. Research and Design* 75 (1997), pp. 32–48.

[CGW10] N. Chen, M. Gunzburger, and X. Wang. "Asymptotic Analysis of the Differences between the Stokes–Darcy System with Different Interface Conditions and the Stokes–Brinkman System". In: *Jounal of Math. Anal. and Appl.* 368.2 (2010), pp. 658–676.

[CWW16] W. Chen, F. Wang, and Y. Wang. "Weak Galerkin Method for the Coupled Darcy–Stokes Flow". In: *IMA J. Numer. Anal.* 36.2 (2016), pp. 897–921.

[Cia02] P. Ciarlet. *The Finite Element Method for Elliptic Problems*. Vol. 40. Classics in Applied Mathematics. Reprint of the 1978 original. Soc. for Ind. and Appl. Math., Philadelphia, PA, 2002, pp. xxviii+530.

[CL02] P. Ciarlet and J. Lions. *Handbook of Numerical Analysis: Solution of Equations in Rn (Part 4), Techniques of Scientific Computer (Part 4), Numerical Methods for Fluids*. Handbook of Numerical Analysis Teil 2. Elsevier Science, 2002.

[Clé75] P. Clément. "Approximation by Finite Element Functions Using Local Regularization". In: *Rev. Française Automat. Informat. Recherche Opérationnelle Sér.* 9.R-2 (1975), pp. 77–84.

[CKS12] B. Cockburn, G. Karniadakis, and C. Shu. *Discontinuous Galerkin Methods: Theory, Computation and Applications*. Lecture Notes in Comput. Sci. and Engrg. Springer Berlin Heidelberg, 2012.

[CKS07] B. Cockburn, G. Kanschat, and D. Schötzau. "A Note on Discontinuous Galerkin Divergence-Free Solutions of the Navier–Stokes Equations". In: *Journal of Sci. Comput.* 31.1-2 (2007), pp. 61–73.

[Cod15] R. Codina. "On some Mathematical Aspects of the Finite Element Approximation of Darcys Problem". In: *MAMERN VI2015: 6th International Conference on Approximation Methods and Numerical Modelling in Environment and Natural Resources* (2015).

[Cod01] R. Codina. "A Stabilized Finite Element Method for Generalized Stationary Incompressible Flows". In: *Comput. Methods Appl. Mech. Engrg.* 190.20-21 (2001), pp. 2681–2706.

[Cos06] A. Costa. "Permeability-Porosity Relationship: A Reexamination of the Kozeny-Carman Equation based on a Fractal Pore-Space Geometry Assumption". In: *Geophysical Research Letters* 33.2 (2006).

[Cro+16] R. Crooijmans, C. Willems, H. Nick, and D. Bruhn. "The Influence of Facies Heterogeneity on the Doublet Performance in Low-Enthalpy Geothermal Sedimentary Reservoirs". In: *Geothermics* 64 (2016), pp. 209–219.

[CR73] M. Crouzeix and P.-A. Raviart. "Conforming and Nonconforming Finite Element Methods for Solving the Stationary Stokes Equations. I". In: *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge* 7 (1973), pp. 33–75.

[D'A12] C. D'Angelo. "Finite Element Approximation of Elliptic Problems with Dirac Measure Terms in Weighted Spaces: Applications to one- and Three-Dimensional Coupled Problems". In: *SIAM J. Numer. Anal.* 50.1 (2012), pp. 194–215.

[DZ09] C. D'Angelo and P. Zunino. "A Finite Element Method Based on Weighted Interior Penalties for Heterogeneous Incompressible Flows". In: *SIAM J. Numer. Anal.* 47.5 (2009), pp. 3990–4020.

[DZ11] C. D'Angelo and P. Zunino. "Robust Numerical Approximation of Coupled Stokes' and Darcy's Flows Applied to Vascular Hemodynamics and Biochemical Transport". In: *ESAIM Math. Model. Numer. Anal.* 45.3 (2011), pp. 447–476.

[Dar56] H. Darcy. *Les Fontaines Publiques de la Ville de Dijon.* Histoire des Fontaines Publiques de Dijon. Appendice. - Note D. 1856.

[Das02] D. Das. "Hydrodynamic Modelling for Groundwater Flow through Permeable Reactive Barriers". In: *Hydrological Processes* 16 (2002), pp. 3393–3418.

[DMJ96] R. Dash, K. Mehta, and G. Jayaraman. "Casson Fluid Flow in a Pipe Filled with a Homogeneous Porous Medium". In: *Internat. J. Engrg. Sci.* 34.10 (1996), pp. 1145–1156.

[Dav04] T. Davis. "Algorithm 832: UMFPACK V4.3 – An Unsymmetric-Pattern Multifrontal Method". In: *ACM Trans. Math. Software* 30.2 (2004), pp. 196–199.

[DD12] F. Demengel and G. Demengel. *Functional Spaces for the Theory of Elliptic Partial Differential Equations.* Universitext. Springer, London, 2012, pp. xviii+465.

[DV+12] M. Diaz-Viera, P. Sahay, M. Coronado, and A. O. Tapia. *Mathematical and Numerical Modeling in Porous Media: Applications in Geosciences.* CRC Press, Taylor & Francis Group, Leiden, 2012.

[Die14] H.-J. Diersch. *FEFLOW. Finite Element Modeling of Flow, Mass and Heat Transport in Porous and Fractured Media*. Springer Science + Business Media; Springer Heidelberg Dordrecht London, 2014.

[DiP08] R. DiPippo. *Geothermal Power Plants: Principles, Applications, Case Studies and Environmental Impact*. Butterworth-Heinemann, 2008.

[Dis04] M. Discacciati. "Domain Decomposition Methods for the Coupling of Surface and Groundwater Flows". PhD thesis. 2004.

[Dis05] M. Discacciati. *Iterative Methods for Stokes/Darcy Coupling*. Vol. 40. Lect. Notes Comput. Sci. Engrg. Springer, Berlin, 2005, pp. 563–570.

[DMQ02] M. Discacciati, E. Miglio, and A. Quarteroni. "Mathematical and Numerical Models for Coupling Surface and Groundwater Flows". In: *Appl. Numer. Math.* 43.1-2 (2002), pp. 57–74.

[DQ09] M. Discacciati and A. Quarteroni. "Navier–Stokes/Darcy Coupling: Modeling, Analysis, and Numerical Approximation". In: *Rev. Mat. Complut.* 22.2 (2009), pp. 315–426.

[DW89] J. Douglas and J. Wang. "An Absolutely Stabilized Finite Element Method for the Stokes Problem". In: *Math. Comp.* 52.186 (1989), pp. 495–508.

[DB87] L. Durlofsky and J. Brady. "Analysis of the Brinkman Equation as a Model for Flow in Porous Media". In: *Physics of Fluids* 30.11 (1987), pp. 3329–3341.

[Dus+16] M. Dussel, E. Lüschen, R. Thomas, T. Agemar, T. Fritzer, S. Sieblitz, B. Huber, J. Birner, and R. Schulz. "Forecast for Thermal Water use from Upper Jurassic Carbonates in the Munich Region (South German Molasse Basin)". In: *Geothermics* 60 (2016), pp. 13 –30.

[Dus+18] M. Dussel, I. Moeck, M. Wolfgramm, and R. Straubinger. "Characterization of a Deep Fault Zone in Upper Jurassic Carbonates of the Northern Alpine Foreland Basin for Geothermal Production (South Germany)". In: Proceedings 43rd Workshop on Geothermal Reservoir Engineering. Stanford University, California, USA. 2018.

[EFL12] M. Ehrhardt, J. Fuhrmann, and A. Linke. "A Model of an Electrochemical Flow Cell with Porous Layer". In: Weierstrass Institute, Preprint No. 1437, 2012.

[Ehr+08] M. Ehrhardt, J. Fuhrmann, A. Linke, and E. Holzbecher. "Mathematical Modeling of Channel-Porous Layer Interfaces in PEM Fuel Cells". In: B. David B, Hissel D (eds.). Proceedings of 'FDFC2008 - Fundamentals and Developments of Fuel Cell Conference 2008'. Nancy, France, 2008.

[Ein06] A. Einstein. “Eine neue Bestimmung der Moleküldimensionen”. In: *Annalen der Physik* 324.2 (1906), pp. 289–306.

[ESW14] H. Elman, D. Silvester, and A. Wathen. *Finite Elements and Fast Iterative Solvers: With Applications in Incompressible Fluid Dynamics*. Second Edition. Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford, 2014, pp. xiv+479.

[EG04] A. Ern and J.-L. Guermond. *Theory and Practice of Finite Elements*. Vol. 159. Applied Mathematical Sciences. Springer, New York, 2004, pp. xiv+524.

[Ern10] H. Ernst. *Geothermal Energy Systems: Exploration, Development, and Utilization*. Wiley-VCH, 2010.

[EFM00] M. Espedal, A. Fasano, and A. Mikelić. *Filtration in Porous Media and Industrial Application*. Vol. 1734. Lecture Notes in Mathematics. Springer, Berlin; Centro Internazionale Matematico Estivo (C.I.M.E.), Florence, 2000, pp. viii+218.

[Eva10] L. Evans. *Partial Differential Equations*. Second Edition. Vol. 19. Graduate Studies in Mathematics. American Mathematical Society, 2010, pp. xxii+749.

[FSS16] N. Farquharson, A. Schubert, and U. Steiner. “Geothermal Energy in Munich (and Beyond). A Geothermal City Case Study”. In: Proceedings Geothermal Resources Council (GRC) Transactions, Vol. 40, Annual Meetings. 2016.

[FF92] L. Franca and S. Frey. “Stabilized Finite Element Methods. II. The Incompressible Navier–Stokes Equations”. In: *Comput. Methods Appl. Mech. Engrg.* 99.2-3 (1992), pp. 209–233.

[FH88] L. Franca and T. Hughes. “Two Classes of Mixed Finite Element Methods”. In: *Comput. Methods Appl. Mech. Engrg.* 69.1 (1988), pp. 89–129.

[FHS93] L. Franca, T. Hughes, and R. Stenberg. “Stabilized Finite Element Methods”. In: *Incompressible Computational Fluid Dynamics: Trends and Advances*. Ed. by M. Gunzburger and R. E. Nicolaides. Cambridge University Press, 1993, 87108.

[FMS16] K. Fujisawa, A. Murakami, and K. Sakai. “Numerical Analysis of Backward Erosion by Soil-Water Interface Tracking”. In: *Japanese Geotechnical Society Special Publication* 4.4 (2016), pp. 84–87.

[GK01] J. Gablonsky and C. Kelley. “A Locally-Biased Form of the DIRECT Algorithm”. In: *Journal of Global Optim.* 21.1 (2001), pp. 27–37.

[Gal11] G. Galdi. *An Introduction to the Mathematical Theory of the Navier–Stokes Equations*. Second Edition. Springer Monographs in Mathematics. Steady-State Problems. Springer, New York, 2011, pp. xiv+1018.

[Gal10] C. Gallegos. *Rheology.* Encyclopedia of Life Support Systems Bd. 1. EOLSS/UNESCO, 2010.

[Gan+16] S. Ganesan, V. John, G. Matthies, R. Meesala, S. Abdus, and U. Wilbrandt. "An Object Oriented Parallel Finite Element Scheme for Computations of PDEs: Design and Implementation". In: *2016 IEEE 23rd International Conference on High Performance Computing Workshops (HiPCW)* (2016), pp. 2–11.

[GAB05] C. Geindreau, J.-L. Auriault, and C. Boutin. "Darcy's Law, Brinkman's Law and Poor Separation of Scales". In: *Poromechanics III - Biot Centennial (1905-2005).* Taylor & Francis, 2005.

[GR09] C. Geuzaine and J.-F. Remacle. "Gmsh: A 3-D Finite Element Mesh Generator with Built-In Pre- and Post-Processing Facilities". In: *Internat. J. Numer. Methods Engrg.* 79.11 (2009), pp. 1309–1331.

[GKR13] V. Girault, G. Kanschat, and B. Rivière. "On the Coupling of Incompressible Stokes or Navier–Stokes and Darcy Flow through Porous Media". In: *Modelling and Simulation in Fluid Dynamics in Porous Media.* Vol. 28. Springer Proc. Math. Stat. Springer, New York, 2013, pp. 1–25.

[GR86] V. Girault and P.-A. Raviart. *Finite Element Methods for Navier–Stokes Equations.* Vol. 5. Springer Series in Computational Mathematics. Theory and Algorithms. Springer, Berlin, 1986, pp. x+374.

[GA94] R. Givler and S. Altobelli. "A Determination of the Effective Viscosity for the Brinkman–Forchheimer Flow model". In: *Journal of Fluid Mechanics* 258.-1 (1994), pp. 355–370.

[Gla10] W. Glassley. *Geothermal Energy: Renewable Energy and the Environment.* CRC Press, Taylor & Francis Group, 2010.

[Gre81] R. Greenkorn. "Steady Flow Through Porous Media". In: *AIChE Journal* 27.4 (1981), pp. 529–545.

[GK10] M. Griebel and M. Klitz. "Homogenization and Numerical Simulation of Flow in Geometries with Textile Microstructures". In: *Multiscale Model. Simul.* 8.4 (2010), pp. 1439–1460.

[Gri11] P. Grisvard. *Elliptic Problems in Nonsmooth Domains.* Vol. 69. Classics in Applied Mathematics. Reprint of the 1985 original. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011, pp. xx+410.

[GLK98] V. Gurau, H. Liu, and S. Kakaç. "Two-Dimensional Model for Proton Exchange Membrane Fuel Cells". In: *AIChE Journal* 44.11 (1998), pp. 2410–2422.

[GN12] J. Guzmán and M. Neilan. "A Family of Nonconforming Elements for the Brinkman Problem". In: *IMA J. Numer. Anal.* 32.4 (2012), pp. 1484–1508.

[Had07] M. Hadamard. "Les Problèmes aux Limites dans la Théorie des Équations aux Dérivées Partielles". In: *Phys. Theor. Appl.* 6 (1) (1907), pp. 202–241.

[Ham94] M. Hamdan. "Single-Phase Flow through Porous Channels: A Review of Flow Models and Channel Entry Conditions". In: *Appl. Math. Comput.* 62.2-3 (1994), pp. 203–222.

[HJS11] A. Hannukainen, M. Juntunen, and R. Stenberg. "Computations with Finite Element Methods for the Brinkman Problem". In: *Computational Geosciences* 15.1 (2011), pp. 155–166.

[HH02] A. Hansbo and P. Hansbo. "An Unfitted Finite Element Method, based on Nitsche's Method, for Elliptic Interface Problems". In: *Comput. Methods Appl. Mech. Engrg.* 191.47-48 (2002), pp. 5537–5552.

[HJ09] P. Hansbo and M. Juntunen. "Weakly Imposed Dirichlet Boundary Conditions for the Brinkman Model of Porous Media Flow". In: *Appl. Numer. Math.* 59.6 (2009), pp. 1274–1289.

[Han05] P. Hansbo. "Nitsche's Method for Interface Problems in Computational Mechanics". In: *GAMM-Mitteilungen* 28.2 (2005), pp. 183–206.

[HH92] I. Harari and T. Hughes. "What are $C$ and $h$?: Inequalities for the Analysis and Design of Finite Element Methods". In: *Comput. Methods Appl. Mech. Engrg.* 97.2 (1992), pp. 157–192.

[Har+14] A. Hartmann, N. Goldscheider, T. Wagener, J. Lange, and M. Weiler. "Karst Water Resources in a Changing World: Review of Hydrological Modeling Approaches". In: *Reviews of Geophysics* 52.3 (2014), pp. 218–242.

[Hor97] U. Hornung, ed. *Homogenization and Porous Media*. Vol. 6. Interdisciplinary Applied Mathematics. Springer, New York, 1997, pp. xvi+275.

[HF87] T. Hughes and L. Franca. "A New Finite Element Formulation for Computational Fluid Dynamics. VII. The Stokes Problem with various Well-Posed Boundary Conditions: Symmetric Formulations that Converge for all Velocity/Pressure Spaces". In: *Comput. Methods Appl. Mech. Engrg.* 65.1 (1987), pp. 85–96.

[HFB86] T. Hughes, L. Franca, and M. Balestra. "A New Finite Element Formulation for Computational Fluid Dynamics. V. Circumventing the Babuška–Brezzi Condition: A Stable Petrov–Galerkin Formulation of the Stokes Problem accommodating Equal-Order Interpolations". In: *Comput. Methods Appl. Mech. Engrg.* 62.1 (1986), pp. 85–99.

[Hug+00] T. Hughes, G. Engel, L. Mazzei, and M. Larson. "A Comparison of Discontinuous and Continuous Galerkin Methods based on Error Estimates, Conservation, Robustness and Efficiency". In: *Discontinuous Galerkin Methods (Newport, RI, 1999)*. Vol. 11. Lect. Notes Comput. Sci. Engrg. Springer, Berlin, 2000, pp. 135–146.

[HA10] W. Hwang and S. Advani. "Numerical Simulations of Stokes–Brinkman Equations for Permeability Prediction of Dual Scale Fibrous Porous Media". In: *Physics of Fluids* 22.11 (2010).

[ILW11] O. Iliev, R. Lazarov, and J. Willems. "Variational Multiscale Finite Element Method for Flows in Highly Porous Media". In: *Multiscale Model. Simul.* 9.4 (2011), pp. 1350–1372.

[IL04] O. Iliev and V. Laptev. "On Numerical Simulation of Flow through Oil Filters". In: *Computing and Visualization in Science* 6.2 (2004), pp. 139–146.

[ILW09] O. Iliev, R. Lazarov, and J. Willems. "Discontinuous Galerkin Subgrid Finite Element Method for Heterogeneous Brinkman's Equations". In: vol. 5910. Large-Scale Scientific Computing, 7th International Conference, LSSC, 2009, pp. 14–25.

[IS99] S. Ingebritsen and W. Sanford. *Groundwater in Geologic Processes.* Cambridge University Press, 1999.

[Ing11] R. Ingram. "Finite Element Approximation of Nonsolenoidal, Viscous Flows around Porous and Solid Obstacles". In: *SIAM J. Numer. Anal.* 49.2 (2011), pp. 491–520.

[JM96] W. Jäger and A. Mikelić. "On the Boundary Conditions at the Contact Interface between a Porous Medium and a Free Fluid". In: *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* 23.3 (1996), pp. 403–465.

[JM00] W. Jäger and A. Mikelić. "On the Interface Boundary Condition of Beavers, Joseph, and Saffman". In: *SIAM J. Appl. Math.* 60.4 (2000), pp. 1111–1127.

[JS89] M. Jobmann and R. Schulz. "Hydrogeothermische Energiebilanz und Grundwasserhaushalt des Malmkarstes im süddeutschen Molassebecken". In: Final report (Abschlussberricht), Band I, Archive Nr. 105040. Niedersächsisches Landesamt für Bodenforschung, Germany. 1989.

[Joh16] V. John. *Finite Element Methods for Incompressible Flow Problems.* Vol. 51. Springer Series in Computational Mathematics. Springer, Cham, 2016, pp. xiii+812.

[JKW19] V. John, P. Knobloch, and U. Wilbrandt. "Finite Element Pressure Stabilizations for Incompressible Flow Problems". In: *Weierstrass Institute, Preprint No. 2587* (2019).

[JS08] V. John and E. Schmeyer. "Finite Element Methods for Time-Dependent Convection-Diffusion-Reaction Equations with Small Diffusion". In: *Comput. Methods Appl. Mech. Engrg.* 198.3-4 (2008), pp. 475–494.

[Joh+17] V. John, A. Linke, C. Merdon, M. Neilan, and L. Rebholz. "On the Divergence Constraint in Mixed Finite Element Methods for Incompressible Flows". In: *SIAM Rev.* 59.3 (2017), pp. 492–544.

[Jon73] I. Jones. "Low Reynolds Number Flow Past a Porous Spherical Shell". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 73.01 (1973), p. 231.

[Joo+09] A. S. Joodi, S. Sizaret, S. Binet, A. Bruand, P. Albéric, and M. Lepiller. "Development of a Darcy–Brinkman Model to Simulate Water Flow and Tracer Transport in a Heterogeneous Karstic Aquifer (Val d'Orléans, France)". In: *Hydrogeology Journal* 18.2 (2009), pp. 295–309.

[JNP82] D. Joseph, D. Nield, and G. Papanicolaou. "Nonlinear Equation governing Flow in a Saturated Porous Medium". In: *Water Resources Research* 18.4 (1982), pp. 1049–1052.

[JT64] D. Joseph and L. Tao. "The Effect of Permeability on the Slow Motion of a Porous Sphere in a Viscous Liquid". In: *ZAMM - Zeitschrift für Angewandte Mathematik und Mechanik* 44.8-9 (1964), pp. 361–364.

[JS10] M. Juntunen and R. Stenberg. "Analysis of Finite Element Methods for the Brinkman Problem". In: *Calcolo* 47.3 (2010), pp. 129–147.

[JS09] M. Juntunen and R. Stenberg. "Nitsche's Method for General Boundary Conditions". In: *Math. Comp.* 78.267 (2009), pp. 1353–1374.

[Kah+19] S. Kahrobaei, R. Fonseca, C. Willems, F. Wilschut, and J. van Wees. "Regional Scale Geothermal Field Development Optimization under Geological Uncertainties". In: *European Geothermal Congress 2019*. 2019.

[KR10] G. Kanschat and B. Rivière. "A Strongly Conservative Finite Element Method for the Coupling of Stokes and Darcy Flow". In: *J. Comput. Phys.* 229.17 (2010), pp. 5933–5943.

[KLM17] G. Kanschat, R. Lazarov, and Y. Mao. "Geometric Multigrid for Darcy and Brinkman Models of Flows in Highly Heterogeneous Porous Media: A Numerical Study". In: *J. Comput. Appl. Math.* 310 (2017), pp. 174–185.

[KS08] G. Kanschat and D. Schötzau. "Energy Norm a Posteriori Error Estimation for Divergence-Free Discontinuous Galerkin Approximations of the Navier–Stokes Equations". In: *Internat. J. Numer. Methods Fluids* 57.9 (2008), pp. 1093–1113.

[Kav02] M. Kaviany. *Principles of Heat Transfer*. Wiley-Interscience publication. Wiley, 2002.

[KG06] T. Kaya and J. Goldak. "Three-Dimensional Numerical Analysis of Heat and Mass Transfer in Heat Pipes". In: *Heat and Mass Transfer* 43.8 (2006), pp. 775–785.

[KG01] N.-R. Kevlahan and J.-M. Ghidaglia. "Computation of Turbulent Flow Past an Array of Cylinders using a Spectral Method with Brinkman Penalization". In: *European Journal of Mechanics* 20.3 (2001), pp. 333–350.

[KV03] A.-R. Khaled and K. Vafai. "The Role of Porous Media in Modeling Flow and Heat Transfer in Biological Tissues". In: *International Journal of Heat and Mass Transfer* 46.26 (2003), pp. 4989–5003.

[Kne90] F. Kneubuehl. *Repititorium der Physik*. Teubner, 1990.

[KS11] J. Könnö and R. Stenberg. "$H(\mathrm{div})$-Conforming Finite Elements for the Brinkman Problem". In: *Math. Models Methods Appl. Sci.* 21.11 (2011), pp. 2227–2248.

[KS12] J. Könnö and R. Stenberg. "Numerical Computations with $H(\mathrm{div})$-Finite Elements for the Brinkman Problem". In: *Comput. Geosci.* 16.1 (2012), pp. 139–158.

[Kop83] J. Koplik. "Viscosity Renormalization in the Brinkman Equation". In: *Physics of Fluids* 26.10 (1983), p. 2864.

[Kro+11] M. Krotkiewski, I. Ligaarden, K.-A. Lie, and D. Schmid. "On the Importance of the Stokes–Brinkman Equations for Computing Effective Permeability in Karst Reservoirs". In: *Comm. in Comp. Physics* 10.5 (2011), pp. 1315–1332.

[Lam13] B. Lamichhane. "A New Finite Element Method for Darcy–Stokes–Brinkman Equations". In: *ISRN Comp. Math.* (2013), pp. 1–4.

[LL59] L. Landau and E. Lifshitz. *Fluid Mechanics*. Translated from the Russian by J.B. Sykes and W.H. Reid. Course of Theoretical Physics, Vol. 6. Pergamon Press, London-Paris-Frankfurt, 1959, pp. xii+536.

[LSY02] W. Layton, F. Schieweck, and I. Yotov. "Coupling Fluid Flow with Porous Media Flow". In: *SIAM J. Numer. Anal.* 40.6 (2002), 2195–2218 (2003).

[LBW06] M. Le Bars and M. Worster. "Interfacial Conditions between a Pure Fluid and a Porous Medium: Implications for Binary Alloy Solidification". In: *J. Fluid Mech.* 550 (2006), pp. 149–173.

[Leh16] C. Lehrenfeld. "Removing the Stabilization Parameter in Fitted and Unfitted Symmetric Nitsche Formulations". In: *Proceedings of the VII European Congress on Computational Methods in Applied Sciences and Engineering*. 2016.

[Lia65] A. Liakopoulos. "Darcy's Coefficient of Permeability as a Symmetric Tensor of Second Rank". In: *Transactions of the ASAE* 8 (1965), pp. 0216–0218.

[LYT89] C. Liu, W. Ying, and J. Tan. "Flow in the Adiabatic Section of a Heat Pipe". In: *International Communications in Heat and Mass Transfer* 16.1 (1989), pp. 79–88.

[LPN07] H. Liu, P. Patil, and U. Narusawa. "On Darcy–Brinkman Equation: Viscous Flow Between Two Parallel Plates Packed with Regular Square Arrays of Cylinders". In: *Entropy* 9.3 (2007), pp. 118–131.

[LZH15] E. Llanos, S. Zarrouk, and R. Hogarth. "Simulation of the Habanero Enhanced Geothermal System (EGS), Australia". In: *Proceedings World Geothermal Congress. Melbourne, Australia.* 2015.

[LZT06] W. Lu, C. Zhao, and S. Tassou. "Thermal Analysis on Metal-Foam Filled Heat Exchangers. Part I: Metal-Foam Filled Pipes". In: *International Journal of Heat and Mass Transfer* 49.15-16 (2006), pp. 2751–2761.

[Lub06] G. Lube. "Stabilized FEM for Incompressible Flow. Critical Review and New Trends". In: *ECCO-MAS CFD, TU Delft* (2006), pp. 1–20.

[LB16] J. Lund and T. Boyd. "Direct Utilization of Geothermal Energy 2015 Worldwide Review". In: *Geothermics* 60 (2016), pp. 66 –93.

[Lun72] T. Lundgren. "Slow Flow Through Stationary Random Beds and Suspensions of Spheres". In: *Journal of Fluid Mechanics* 51.02 (1972), pp. 273–299.

[ML98] S. Malta and A. Loula. "Numerical Analysis of Finite Element Methods for Miscible Displacements in Porous Media". In: *Numer. Methods Partial Differential Equations* 14.4 (1998), pp. 519–548.

[MTW02] K. Mardal, X.-C. Tai, and R. Winther. "A Robust Finite Element Method for Darcy–Stokes Flow". In: *SIAM J. Numer. Anal.* 40.5 (2002), pp. 1605–1631.

[MS74] G. Marsaglia and G. Styan. "Equalities and Inequalities for Ranks of Matrices". In: *Linear and Multilinear Algebra* 2 (1974/75), pp. 269–292.

[MPPM12] E. Marušić-Paloka, I. Pažanin, and S. Marušić. "Comparison between Darcy and Brinkman Laws in a Fracture". In: *Appl. Math. Comput.* 218.14 (2012), pp. 7538–7545.

[MH02] A. Masud and T. Hughes. "A Stabilized Mixed Finite Element Method for Darcy Flow". In: *Comput. Methods Appl. Mech. Engrg.* 191.39-40 (2002), pp. 4341–4370.

[MFK17] T. Matsushita, S. Fujibayashi, and T. Kokubo. "4 - Titanium Foam for Bone Tissue Engineering". In: *Metallic Foam Bone*. Ed. by C. Wen. Woodhead Publishing, 2017, pp. 111 –130.

[Mez90] T. Mezger. *The Rheology Handbook*. Vincentz Network, 1990.

[Mik00] A. Mikelić. "Homogenization Theory and Applications to Filtration through Porous Media". In: *Filtration in Porous Media and Industrial Application (Cetraro, 1998)*. Vol. 1734. Lecture Notes in Math. Springer, Berlin, 2000, pp. 127–214.

[MR09] W. Miladi and M. Racila. "Mathematical Model of Fluid Flow in an Osteon: Influence of Cardiac System". In: *Computer Methods in Biomechanics and Biomedical Engrg.* 12.sup1 (2009), pp. 187–188.

[MS17] F. Morales and R. Showalter. "A Darcy–Brinkman Model of Fractures in Porous Media". In: *J. Math. Anal. Appl.* 452.2 (2017), pp. 1332–1358.

[Mra+19] E. Mraz, M. Wolfgramm, I. Moeck, and K. Thuro. "Detailed Fluid Inclusion and Stable Isotope Analysis on Deep Carbonates from the North Alpine Foreland Basin to Constrain Paleofluid Evolution". In: *Geofluids* (2019).

[MWY14] L. Mu, J. Wang, and X. Ye. "A Stable Numerical aAlgorithm for the Brinkman Equations by Weak Galerkin Finite Element Methods". In: *J. Comput. Phys.* 273 (2014), pp. 327–342.

[MW37] M. Muskat and R. Wyckoff. *The Flow of Homogeneous Fluids Through Porous Media*. International Series in Physics. McGraw-Hill Book Company, Incorporated, 1937.

[Mut15] Y. Mutlu. *Einstellung von Volumenströmen im Bereich der Nanofluidik: Entwicklung einer Fluid-Drossel aus porösen Keramiken*. Springer Fachmedien Wiesbaden, 2015.

[NN74] G. Neale and W. Nader. "Practical Significance of Brinkman's Extension of Darcy's Law: Coupled Parallel Flows within a Channel and a Bounding Porous Medium". In: *The Canadian Journal of Chem. Engrg.* 52.4 (1974), pp. 475–478.

[New87] I. Newton. *Philosophiae naturalis principia mathematica*. J. Societatis Regiae ac Typis J. Streater, 1687.

[Nie83] D. Nield. "The Boundary Correction for the Rayleigh–Darcy Problem: Limitations of the Brinkman Equation". In: *Journal of Fluid Mech.* 128 (1983), pp. 37–46.

[NB06] D. Nield and A. Bejan. *Convection in Porous Media*. Third Edition. Springer, 2006.

[Nis+18] C. Nishad, A. Chandra, T. Karmakar, and R. Sekhar. "A Non-Primitive Boundary Element Technique for Modeling Flow through Non-Deformable Porous Medium using Brinkman Equation". In: *Meccanica* 53.9 (2018), pp. 2333–2352.

[Nit71] J. Nitsche. "Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind". In: *Abh. Math. Sem. Univ. Hamburg* 36 (1971). Collection of articles dedicated to Lothar Collatz on his sixtieth birthday, pp. 9–15.

[NS98] S. Norburn and D. Silvester. "Stabilised vs. Stable Mixed Methods for Incompressible Flow". In: *Comput. Methods Appl. Mech. Engrg.* 166.1-2 (1998), pp. 131–141.

[OTW95a] J. Ochoa-Tapia and S. Whitaker. "Momentum Transfer at the Boundary between a Porous Medium and a Homogeneous Fluid I. Theoretical Development". In: *International Journal of Heat and Mass Transfer* 38.14 (1995), pp. 2635 –2646.

[OTW95b] J. Ochoa-Tapia and S. Whitaker. "Momentum Transfer at the Boundary between a Porous Medium and a Homogeneous Fluid II. Comparison with Experiment". In: *International Journal of Heat and Mass Transfer* 38.14 (1995), pp. 2647 –2655.

[OMB70] G. Ooms, P. Mijnlieff, and H. Beckers. "Frictional Force Exerted by a Flowing Fluid on a Permeable Particle, with Particular Reference to Polymer Coils". In: *The Journal of Chem. Physics* 53.11 (1970), pp. 4123–4130.

[PVC18] S. Pandey, V. Vishal, and A. Chaudhuri. "Geothermal Reservoir Modeling in a Coupled Thermo-Hydro-Mechanical-Chemical Approach: A Review". In: *Earth-Science Reviews* 185 (2018), pp. 1157 –1169.

[Pn+17] I. Patiño, H. Power, C. Nieto-Londoño, and W. Flórez. "Stokes–Brinkman Formulation for Prediction of Void Formation in Dual-Scale Fibrous Reinforcements: A BEM/DR-BEM Simulation". In: *Comput. Mech.* 59.4 (2017), pp. 555–577.

[Pes02] C. Peskin. "The Immersed Boundary Method". In: *Acta Numer.* 11 (2002), pp. 479–517.

[Pod+14] J. Podichetty, P. Bhaskar, A. Khalf, and S. Madihally. "Modeling Pressure Drop Using Generalized Scaffold Characteristics in an Axial-Flow Bioreactor for Soft Tissue Regeneration". In: *Annals of Biomedical Engrg.* 42.6 (2014), pp. 1319–1330.

[PEQ09] P. Popov, Y. Efendiev, and G. Qin. "Multiscale Modeling and Simulations of Flows in Naturally Fractured Karst Reservoirs". In: *Commun. Comput. Phys.* 6.1 (2009), pp. 162–184.

[PLM18] F. de Prenter, C. Lehrenfeld, and A. Massing. "A Note on the Stability Parameter in Nitsche's Method for Unfitted Boundary Value Problems". In: *Comput. Math. Appl.* 75.12 (2018), pp. 4322–4336.

[QV08] A. Quarteroni and A. Valli. *Numerical Approximation of Partial Differential Equations*. Springer, Berlin, 2008.

[Ree16] S. Rees. *Advances in Ground-Source Heat Pump Systems*. Elsevier Science, 2016.

[RR04] M. Renardy and R. C. Rogers. *An Introduction to Partial Differential Equations*. Second Edition. Vol. 13. Texts in Applied Mathematics. Springer, New York, 2004, pp. xiv+434.

[RS13] L. Rios and N. Sahinidis. "Derivative-Free Optimization: A Review of Algorithms and Comparison of Software Implementations". In: *Journal of Global Optimization* 56 (2013), pp. 1247–1293.

[RWG01] B. Rivière, M. Wheeler, and V. Girault. "A Priori Error Estimates for Finite Element Methods Based on Discontinuous Approximation Spaces for Elliptic Problems". In: *SIAM J. Numer. Anal.* 39.3 (2001), pp. 902–931.

[Nlo] S. Johnson. The NLopt Nonlinear-Optimization Package. `http://github.com/stevengj/nlopt`.

[RST08] H.-G. Roos, M. Stynes, and L. Tobiska. *Robust Numerical Methods for Singularly Perturbed Differential Equations*. Second Edition. Vol. 24. Springer Series in Computational Mathematics. Springer, Berlin, 2008, pp. xiv+604.

[RJZ19] A. Rostamian, S. Jamshidi, and E. Zirbes. "The Development of a Novel Multi-Objective Optimization Framework for Non-Vertical Well Placement based on a Modified Non-Dominated Sorting Genetic Algorithm-II". In: *Comput. Geosci.* 23.5 (2019), pp. 1065–1085.

[Rub86] J. Rubinstein. "On the Macroscopic Description of Slow Viscous Flow Past a Random Array of Spheres". In: *J. Statist. Phys.* 44.5-6 (1986), pp. 849–863.

[RZ09] H. Rui and R. Zhang. "A Unified Stabilized Mixed Finite Element Method for Coupling Stokes and Darcy Flows". In: *Comput. Methods Appl. Mech. Engrg.* 198.33-36 (2009), pp. 2692–2699.

[Sae+15] S. Saeid, R. Al-Khoury, H. Nick, and M. Hicks. "A Prototype Design Model for Deep Low-Enthalpy Hydrothermal Systems". In: *Renewable Energy* 77 (2015), pp. 408–422.

[Sae+14] S. Saeid, R. Al-Khoury, H. Nick, and F. Barends. "Experimental-Numerical Study of Heat Flow in Deep Low-Enthalpy Geothermal Conditions". In: *Renewable Energy* 62 (2014), pp. 716 –730.

[Saf71] P. Saffman. "On the Boundary Condition at the Surface of a Porous Medium". In: *Studies in Appl. Math.* 50.2 (1971), pp. 93–101.

[SAD94] A. Salinger, R. Aris, and J. Derby. "Finite Element Formulations for Large-Scale, Coupled Flows in Adjacent Porous and Open Fluid Domains". In: *Internat. Journal for Numerical Methods in Fluids* 18.12 (1994), pp. 1185–1209.

[Sch61] A. Scheidegger. "General Theory of Dispersion in Porous Media". In: *Journal of Geophysical Research (1896-1977)* 66.10 (1961), pp. 3273–3278.

[Sch+16] D. Schillinger, I. Harari, M.-C. Hsu, D. Kamensky, S. Stoter, Y. Yu, and Y. Zhao. "The Non-Symmetric Nitsche Method for the Parameter-Free Imposition of Weak Boundary and Coupling Conditions in Immersed Finite Elements". In: *Comput. Methods Appl. Mech. Engrg.* 309 (2016), pp. 625–652.

[SZ90] R. Scott and S. Zhang. "Finite Element Interpolation of Nonsmooth Functions satisfying Boundary Conditions". In: *Math. Comp.* 54.190 (1990), pp. 483–493.

[ST14] W. Sobieski and A. Trykozko. "Darcy's and Forchheimer's Law in Practice. Part 1. The Experiment". In: *Technical Sciences* 17.4 (2014).

[Sog14] J. Sogn. "Stabilized Finite Element Methods for the Brinkman Equation on Fitted and Fictitious Domains". MA thesis. Springer New York, 2014.

[Ste98] R. Stenberg. "Mortaring by a Method of J. A. Nitsche". In: *Comp. Mechanics (Buenos Aires, 1998)*. Centro Internac. Métodos Numér. Ing., Barcelona, 1998, CD–ROM file.

[Ste95] R. Stenberg. "On Some Techniques for Approximating Boundary Conditions in the Finite Element Method". In: *J. Comput. Appl. Math.* 63.1-3 (1995). International Symposium on Mathematical Modelling and Computational Methods Modelling 94 (Prague, 1994), pp. 139–148.

[SF95] R. Stenberg and J. Freund. "On Weakly Imposed Boundary Conditions for Second Order Problems". In: *Proceedings of the Ninth Int. Conf. Finite Elements in Fluids, M. Morandi Cecchi et al. Eds., Venice (1995)* (1995).

[SB13] I. Stober and K. Bucher. *Geothermal Energy*. Springer Berlin Heidelberg, 2013.

[Suk+13] M. Sukop, H. Huang, P. Alvarez, E. Variano, and K. Cunningham. "Evaluation of Permeability and Non-Darcy Flow in Vuggy Macroporous Limestone Aquifer Samples with Lattice Boltzmann Methods". In: *Water Resources Research* 49.1 (2013), pp. 216–230.

[SC11] K. Sutherland and G. Chase. *Filters and Filtration Handbook*. Elsevier Science, 2011.

[TT00] S. Tada and J. Tarbell. "Interstitial Flow through the Internal Elastic Lamina Affects Shear Stress on Arterial Smooth Muscle Cells". In: *American Journal of Physiology-Heart and Circulatory Physiology* 278.5 (2000).

[Tam69] C. Tam. "The Drag on a Cloud of Spherical Particles in Low Reynolds Number Flow". In: *Journal of Fluid Mechanics* 38.03 (1969), p. 537.

[Tho06] V. Thomée. *Galerkin Finite Element Methods for Parabolic Problems*. Second Edition. Vol. 25. Springer Series in Computational Mathematics. Springer, Berlin, 2006, pp. xii+370.

[Vaf00] K. Vafai. *Handbook of Porous Media*. Taylor & Francis, 2000.

[VY09] D. Vassilev and I. Yotov. "Coupling Stokes–Darcy Flow with Transport". In: *SIAM J. Sci. Comput.* 31.5 (2009), pp. 3661–3684.

[Ver91] R. Verfürth. “Finite Element Approximation of Incompressible Navier–Stokes Equations with Slip Boundary condition. II”. In: *Numer. Math.* 59.6 (1991), pp. 615–636.

[Ver87] R. Verfürth. “Finite Element Approximation of Incompressible Navier–Stokes Equations with Slip Boundary Condition”. In: *Numer. Math.* 50.6 (1987), pp. 697–721.

[Vör+07] R. Vörös, R. Weidler, L. De Graaf, and D. Wyborn. “Thermal Modelling of Long Term Circulation of Multi-Well Development at the Cooper Basin Hot Fractured Rock (HFR) Project and current proposed Scale-Up Program”. In: Proceedings 32nd Workshop on Geothermal Reservoir Engrg. 2007.

[WWY09] J. Wang, Y. Wang, and X. Ye. “A Robust Numerical Method for Stokes Equations based on Divergence-Free $H(\mathrm{div})$ Finite Element Methods”. In: *SIAM J. Sci. Comput.* 31.4 (2009), pp. 2784–2802.

[WY07] J. Wang and X. Ye. “New Finite Element Methods in Computational Fluid Dynamics by $H(\mathrm{div})$ Elements”. In: *SIAM J. Numer. Anal.* 45.3 (2007), pp. 1269–1286.

[WBM19] J. Weber, H. Born, and I. Moeck. “Geothermal Energy Use, Country Update for Germany 2016 - 2018”. In: Proceedings European Geothermal Congress. 2019.

[Whi86] S. Whitaker. “Flow in Porous Media I: A Theoretical Derivation of Darcy’s Law”. In: *Transport in Porous Media* 1.1 (1986), pp. 3–25.

[Wil19] U. Wilbrandt. *Stokes–Darcy Equations*. Lecture Notes in Mathematical Fluid Mechanics. Analytic and Numerical Analysis. Springer, Cham, 2019, pp. viii+212.

[Wil+17a] U. Wilbrandt, C. Bartsch, N. Ahmed, N. Alia, F. Anker, L. Blank, A. Caiazzo, S. Ganesan, S. Giere, G. Matthies, R. Meesala, A. Shamim, J. Venkatesan, and V. John. “ParMooN – A Modernized Program Package based on Mapped Finite Elements”. In: *Comput. Math. Appl.* 74.1 (2017), pp. 74–88.

[Wil+17b] C. Willems, H. Nick, G. Weltje, and D. Bruhn. “An Evaluation of Interferences in Heat Production from Low Enthalpy Geothermal Doublets Systems”. In: *Energy* 135 (2017), pp. 500–512.

[Wil+17c] C. Willems, H. Nick, T. Goense, and D. Bruhn. “The Impact of Reduction of Doublet Well Spacing on the Net Present Value and the Life Time of Fluvial Hot Sedimentary Aquifer Doublets”. In: *Geothermics* 68 (2017), pp. 54 –66.

[Wil+16] C. Willems, T. Goense, H. Nick, and D. Bruhn. “The Relation Between Well Spacing and Net Present Value in Fluvial Hot Sedimentary Aquifer Geothermal Doublets: A West Netherlands Basin Case Study”. In: *41st Workshop on Geothermal Resevoir Engrg.* 2016.

[XXX08] X. Xie, J. Xu, and G. Xue. "Uniformly-Stable Finite Element Methods for Darcy–Stokes–Brinkman Models". In: *J. Comput. Math.* 26.3 (2008), pp. 437–455.

[You+10] D. Young, B. Munson, T. Okiishi, and W. Huebsch. *A Brief Introduction To Fluid Mechanics*. John Wiley & Sons, 2010.

[ZG06] Z. Zeng and R. Grigg. "A Criterion for Non-Darcy Flow in Porous Media". In: *Transport in Porous Media* 63.1 (2006), pp. 57–69.

[Zha+19] L. Zhang, Z. Deng, K. Zhang, T. Long, J. Desbordes, H. Sun, and Y. Yang. "Well-Placement Optimization in an Enhanced Geothermal System Based on the Fracture Continuum Method and 0-1 Programming". In: *Energies* 12 (2019), p. 709.

# Zusammenfassung

Die vorliegende Dissertation befasst sich mit der mathematischen Beschreibung des Verhaltens von Flüssigkeiten in porösen Medien, mittels partieller Differentialgleichungen, und dessen numerischer Approximation. Letztere ist in verschiedensten Anwendungsgebieten von Interesse, beispielsweise bei der Optimierung von industriellen Filterprozessen oder der Prognose von Grundwasserströmungen.

Die mikroskopische Komplexität poröser Strukturen und oft auch fehlende Informationen über diese, hat zur Entwicklung von makroskopischen Modellen auf Basis der Navier–Stokes-Gleichungen geführt, die skalenbezogen unterschiedliche physikalische Phänomene abbilden können.

Im Zentrum dieser Arbeit stehen die Brinkman-Gleichungen. Diese beinhalten zwei Modell-Parameter und ermöglichen eine Transformation, sowohl in ein Darcy- als auch in ein Stokes-Modell. Aufgrund der unterschiedlichen Eigenschaften der Grenzmodelle unterscheiden sich herkömmliche wohldefinierte Finite-Elemente-Ansätze zur diskreten Lösung der entsprechenden Probleme voneinander.

In der Anwendung sind Finite-Elemente-Methoden niedriger Ordnung und geringer Komplexität attraktiv, da diese den Implementierungs- und Rechenaufwand gering halten. Dies wird bei der Entwicklung der hier vorgestellten Methode berücksichtigt.

Um die Stabilität einer solchen Finite-Elemente-Diskretisierung für das gesamte Spektrum an Modell-Parametern des Brinkman-Problems zu garantieren, wird eine Kombination aus folgenden Bestandteilen vorgeschlagen und untersucht:

- stetige, elementweise lineare Finite-Elemente-Räume mit einer residualbasierten Druckstabilisierung (die unsymmetrische GLS-Methode),
- eine Grad-Div-Stabilisierung und
- die Penalty-freie unsymmetrische Nitsche Methode mit einer zusätzlichen Stabilisierung von Sprüngen der Normalengeschwindigkeit entlang des Randes im Darcy-Grenzfall.

Es wird gezeigt, dass die vorgeschlagene Methode (in zwei Raumdimensionen) unabhängig von der Wahl der Parameter in den Brinkman-Gleichungen wohldefiniert ist und optimale a priori Fehlerabschätzungen ermöglicht. Eine praktische Untersuchung im Rahmen einer umfassenden Konvergenzstudie bestätigt die Robustheit.

Darüber hinaus wird die Energiegewinnung aus Erdwärme in sedimentären Grundwasserleitern modelliert und simuliert. Der Fokus liegt hierbei auf der Untersuchung der Auswirkungen charakteristischer heterogener Permeabilitätsstrukturen und natürlicher Temperaturfelder auf die optimale Positionierung von geothermischen Mehrfachbohrungsanlagen. Dabei wird die vorgeschlagene Methode für das Brinkman Problem als Teil eines diskretisierten thermo-hydraulischen, sequentiell gekoppelten Modellierungsansatzes verwendet.

## Selbstständigkeitserklärung

Ich versichere hiermit, alle Hilfsmittel und Hilfen angegeben zu haben, und die Arbeit selbstständig und ausschlielich auf Grundlage der angegebenen Hilfsmittel und Hilfen angefertigt zu haben. Des Weiteren versichere ich, die Arbeit oder Teile der Arbeit nicht schon einmal in einem früheren Promotionsverfahren eingereicht zu haben.

Berlin, 08. November 2019

Laura Caroline Blank